

Breaks or Long Memory Behaviour: An empirical Investigation

Lanouar Charfeddine * Dominique Guégan †

Abstract

Are structural breaks models true switching models or long memory processes? The answer to this question remain ambiguous. A lot of papers, in recent years, have dealt with this problem. For instance, Diebold and Inoue (2001) and Granger and Hyung (2004) show, under specific conditions, that switching models and long memory processes can be easily confused. In this paper, using several generating models like the mean-plus-noise model, the STOchastic Permanent BREAK model, the Markov switching model, The TAR model, the sign model and the Structural CHange model (SCH) and several estimation techniques like the the GPH technique, the Exact Local Whittle (ELW) and the Wavelet methods, we show that, if the answer is quite simple in some cases, it can be mitigate in other cases. Using French and American inflation rates we show that these series cannot be characterized by the same class of models. The main result of this study suggests that estimating the long memory parameter without taking account existence of breaks in the data sets may lead to misspecification and to overestimate the true parameter.

JEL classification: C13 - C32 - E3

Keywords: Structural breaks models, Spurious long memory behavior, Inflation series

*Corresponding author: *Email address:* lanouar_charf@yahoo.fr.

†PSE, CES-MSE, University Paris1 Panthéon-Sorbonne, 106 Bd de l'hopital, 75013, Paris, France, e-mail: dguegan@univ-paris1.fr

1 Introduction

In recent years, many works have provided a theoretical justification and empirical evidence that several structural change models can create spurious long memory behavior. For instance, short memory models with changing regime in mean can generate identical behaviour compared with the one derived from long memory process looking at the evolution of the autocovariance function. This behavior has been pointed out by a lot of authors: Chen and Tiao (1990) for the mean-plus-noise model, Engle and Smith (1999) for the stochastic permanent break model, Granger and Teräsvirta (1999) for the sign model, Hamilton (1989) for the Markov switching model, Lim and Tong (1980) for the Threshold Auto-Regressive (TAR) model, Gouriéroux and Jasiak (2001) for the infrequent breaks model, for instance.

In the literature, until now, no consensus has been reached about the link between presence of structural changes models and detection of long memory behavior. Diebold and Inoue (2001) argued that switching regimes and long memory behavior can be easily confused. Granger and Hyung (2004) suggested that it is difficult to discriminate between these two classes of models, say structural changes model and long memory process. Empirical studies in testing for long-range dependence show that, when the data are weakly dependent with changes in mean, statistical tests used to detect long memory behavior have a high size distortion, indeed they are not able to reject the alternative hypothesis of long memory with a high power.

This paper focus on the behaviors of the estimations for the long memory parameter under misspecification. We consider as estimation techniques the Geweke-Porter-Hudak (1989) method (GPH), the Exact Local Wittle estimate of Shimotsu and Phillips (2005) and the Wavelet estimator of Lee (2005). The generating processes are the mean-plus-noise, the Stopbreak, the Markov-Switching, the Threshold autoregressive and the Structural Changes processes.

Diebold and Inoue (2001) already provide simulation results for the performance of the GPH estimator under some data generating processes such as mean-plus-noise and stopbreak models in order to study this problem of misspecification. Here we perform some similar studies for a wider class of processes and estimators. Then, an application is provided in order to investigate the possible origin of the long range dependence behavior observed in the American and French inflation series.

All along this paper we say that a process $(Y_t)_t$ has a long memory behavior

if its autocovariance function

$$\rho_k \approx c(k)k^{2d-1}, k \rightarrow \infty,$$

where $d > 0$ and $c(k)$ is a slowly varying function at infinity. This means, that given a sample (Y_1, \dots, Y_T) , the sample correlation function $\hat{\rho}(k)$ decays with an hyperbolic rate towards zero. In the next section, we introduce stochastic processes which exhibit this kind of behavior. For more details on long memory behavior, we refer to Guégan (2005).

We say that we observe a spurious long memory behavior inside data sets, if knowing that the theoretical generating process is short memory (in this later case the autocorrelation function decreases with an exponential rate towards zero), the sample autocorrelation function decays slowly towards zero with an hyperbolic rate.

The rest of the paper is organized as follows. Section 2 recalls the analytical expressions of the processes used in this study: the long memory process and several structural changes models. Section 3 describes the estimation methods used to estimate the long memory parameter. Section 4 reports Monte Carlo simulation results, which illustrate the problem of misspecification of estimation of long memory behavior when breaks are present in the data sets. In Section 5 we investigate the behavior of French and US inflation rates. Section 6 concludes.

2 Long memory and Structural changes models

In this section, we specify the models that we study all along the paper.

2.1 FI(d) process

In 1980 Granger and Joyeux introduced the so-called Auto-Regressive Fractionally Integrated Moving Average (ARFIMA) model, developed independently by Hosking (1981). We say that a process $\{Y_t\}_t$, follows a Fractional Integrated FI(d) process if, $\forall t$,

$$(I - L)^d(Y_t - \mu) = u_t, \tag{1}$$

where $(u_t)_t$ is assumed to be a Gaussian strong white noise $N(0, \sigma_u^2)$, μ is the mean and L is the lag operator. In the following we assume that $0 < d < 1$. This means that the autocorrelation function of the process (1) exhibits a slow decay towards zero which characterizes existence of the long memory behavior. The parameter d is allowed to be greater than 1/2 which permits to work with non-stationary processes and, as soon as $d < 1$ this means that the process $(Y_t)_t$ is mean reverting.

2.2 Structural changes models

Now, we define some structural changes models. First we assume that, $\forall t$

$$Y_t = \mu_t + \epsilon_t, \quad (2)$$

where $(\epsilon_t)_t$ is a Gaussian strong white noise $N(0, \sigma_\epsilon^2)$ and $(\mu_t)_t$ is an occasional level shift process. In the following, we propose several forms for the processes $(\mu_t)_t$ and $(\epsilon_t)_t$. We can classify these models as follows:

1. We assume that the process $(\mu_t)_t$ follows a random walk denoted also I(1) process, controlled by two random variables q_t and η_t .
2. We assume that the process $(\mu_t)_t$ follows a Markov switching model, a TAR model, or a structural changes model.

All these models are known to have theoretically a short memory behavior (in the autocorrelation sense) as soon as they are stationary. We describe now these classes of models.

1. Model 1: The mean-plus-noise model. We assume that the process $(\mu_t)_t$ is such that, $\forall t$

$$\mu_t = \mu_{t-1} + q_t \eta_t, \quad (3)$$

where $(\eta_t)_t$ is a Gaussian white noise with zero mean and variance σ_η^2 and, q_t is a sequence of independent identically distributed random variables with a binomial distribution

$$q_t = \begin{cases} 0 & \text{with probability } 1-p \\ 1 & \text{with probability } p. \end{cases} \quad (4)$$

The process (2)-(4) is the so-called mean plus noise model introduced by Chen and Tiao, (1990) and Engle and Smith, (1999), see also Diebold and Inoue (2001). Given a sample (Y_1, \dots, Y_T) we provide on Figure 1, the trajectory of such a process and the sample autocorrelation function $\hat{\rho}(h)$. The breaks that we observe on the Figure 1(a) are obtained as soon as $p \times T$ is a non-zero finite constant with $0 < p < 1$ and T the sample size. The breaks arise when the sample size increases and the probability p decreases. On Figure 1(b) we can observe that the autocorrelation function decreases very slowly as the lag k increases.

2. Model 2: The Stopbreak model. The occasional level shifts model $(\mu_t)_t$ is defined by,

$$\mu_t = \mu_{t-1} + q_{t-1} \epsilon_{t-1} \quad (5)$$

where,

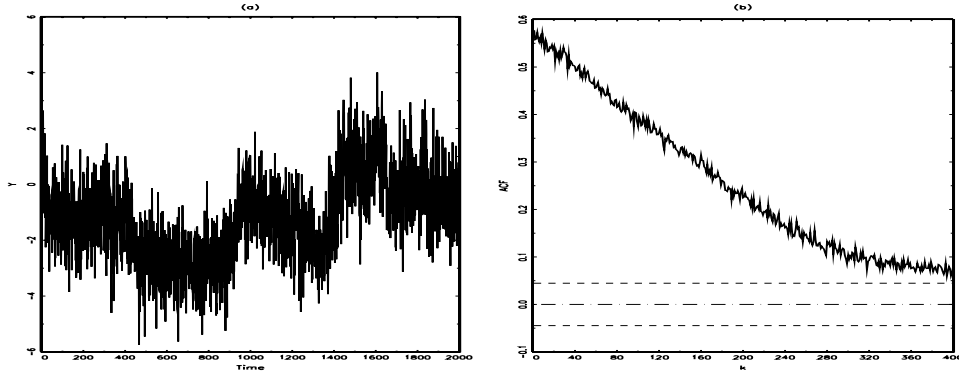


Figure 1: (a) and (b) are respectively the trajectory and the ACF of the Mean-Plus-Noise model (2)-(4) with $p = 0.01$, $T = 2000$, $\sigma_\epsilon^2 = \sigma_\eta^2 = 1$.

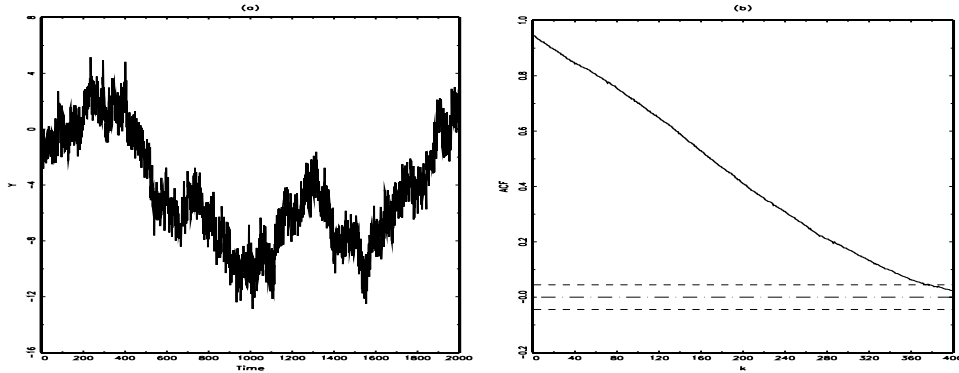


Figure 2: (a) and (b) represent respectively the trajectory and the ACF of the stochastic permanent break model (2)-(5)-(6) with $\gamma = 100$, $T = 2000$, $\sigma_\epsilon^2 = 1$.

$$q_{t-1} = \frac{\epsilon_{t-1}^2}{\gamma + \epsilon_{t-1}^2}. \quad (6)$$

The function q_{t-1} is a non decreasing function and it is bounded by zero and one. Many other smooth transition functions q_{t-1} can be proposed, for instance, the logistic function $q_{t-1} = \frac{1}{1+\exp(\epsilon_{t-1})}$ and also the functions $q_{t-1} = \frac{\exp(\epsilon_{t-1})}{\gamma+\exp(\epsilon_{t-1})}$, $q_{t-1} = \frac{1}{\gamma+\epsilon_{t-1}^2}$, or $q_{t-1} = \frac{1}{\gamma+|\epsilon_{t-1}|}$. The model (2)-(5)-(6) is the so-called Stopbreak process introduced by Engle and Smith (1999). On Figure 2(a), we report the trajectory of this model. Looking at the sample autocorrelation function, on Figure 2 (b) we observe a slow decreasing of this function.

3. Model 3. Markov switching model. We assume that the occasional level shifts process, $(\mu_t)_t$, depends on an unobserved Markov chain s_t ,

$$\mu_t = \mu_{s_t}, \quad (7)$$

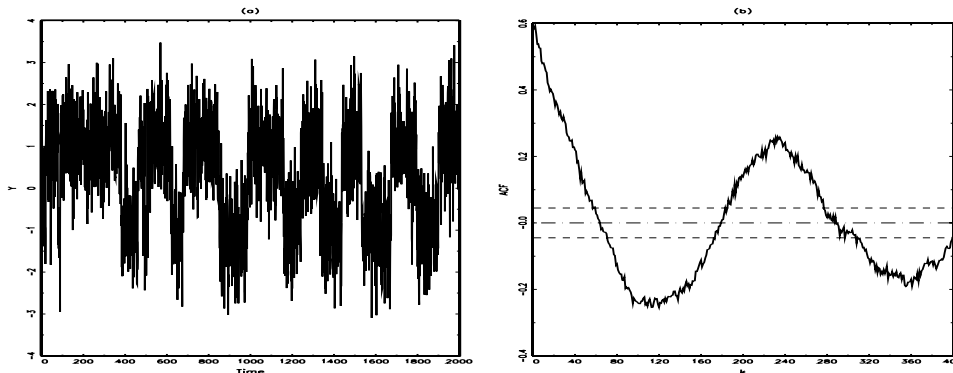


Figure 3: (a) and (b) are respectively the trajectory and the ACF of the Markov switching model defined by (2) and (7) with $T = 2000$, $\sigma_\epsilon^2 = 0.2$, $p_{00} = 0.95$, $p_{11} = 0.95$ and $\mu_0 = -\mu_1 = -1$.

where s_t is a random variable which takes values 0 and 1, characterized by a transition probability which is defined by:

$$p(s_t = j | s_{t-1} = i) = p_{ij}, \quad i, j = 0, 1,$$

with $0 \leq p_{ij} \leq 1$, $\sum_{j=0}^1 p_{ij} = 1$, $i = 0, 1$. The transition matrix is equal to

$$\mathbf{P} = \begin{pmatrix} p_{00} & 1 - p_{00} \\ 1 - p_{11} & p_{11} \end{pmatrix}.$$

Diebold and Inoue (2001) have shown that for some sets of values p_{00} and p_{11} , the autocorrelation function decreases slowly towards zero. This behavior is described on Figure 3.

4. Model 4. Threshold Auto-Regressive model. The occasional level shifts process $(\mu_t)_t$, depends on an observed variable, a threshold parameter r and a delay parameter d

$$\mu_t = \begin{cases} \mu_0 & \text{if } Y_{t-d} \leq r \\ \mu_1 & \text{if } Y_{t-d} > r. \end{cases} \quad (8)$$

The model defined by (2) and (8) is a simple form of the Threshold Auto-Regressive model introduced by Lim and Tong (1980). This model (2)-(8) is also a particular example of Markov switching model when the Markov chain $I\{Y_{t-1} \leq r\}$ is not exogenous, see Carrasco (2002). On Figures 4 (a) and (b), we report the trajectory and the autocorrelation function for this model. Again we observe a slow decreasing of the autocorrelation function.

5. Model 5. The sign model. The occasional level shifts process $(\mu_t)_t$, follows a sign model if

$$\mu_t = \text{sign}(Y_{t-1}), \quad (9)$$

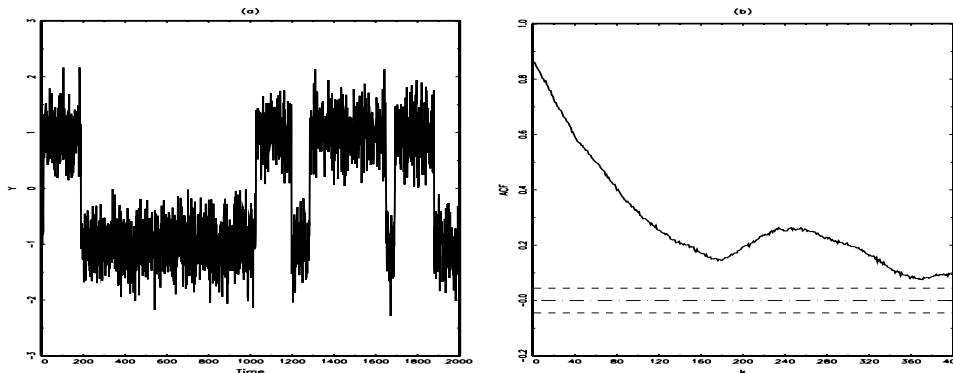


Figure 4: (a) and (b) are respectively the trajectory and the ACF of the Threshold Auto-Regressive model defined by (2) and (8) with $T = 2000$, $\sigma_\epsilon^2 = 0.2$, $r = 0$, $d = 1$ and $\mu_0 = -\mu_1 = -1$.

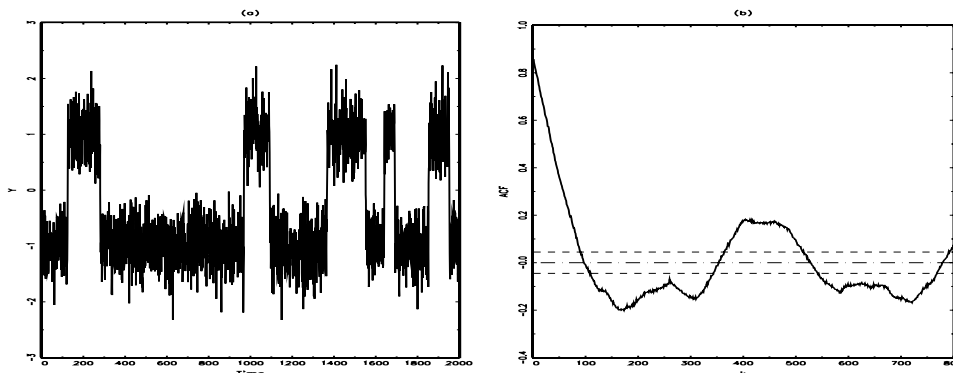


Figure 5: (a) and (b) are respectively the trajectory and the ACF of the sign model (2)-(9)-(10), $T = 2000$ and $\sigma_\epsilon^2 = 0.2$.

where the $\text{sign}(\cdot)$ function is defined by

$$\text{sign}(Y_{t-1}) = \begin{cases} 1 & \text{if } Y_{t-1} < 0 \\ 0 & \text{if } Y_{t-1} = 0 \\ -1 & \text{if } Y_{t-1} > 0, \end{cases} \quad (10)$$

This model is the so-called sign model introduced by Granger and Teräsvirta (1999). They assume that the noise $(\epsilon_t)_t$ given in (2) is such that $p = \text{Prob}(\epsilon_t < -1) = \text{Prob}(\epsilon_t > 1)$. In theory the process defined by the relationships (2), (9), and (10) still have a short memory behavior. Nevertheless, when p is small, a slow decay of the autocorrelation function can be observed. We illustrate this fact on Figure 5.

6. Model 6. The Structural CHange model. The occasional level shifts process $(\mu_t)_t$, follows a process with two regimes and an unknown time

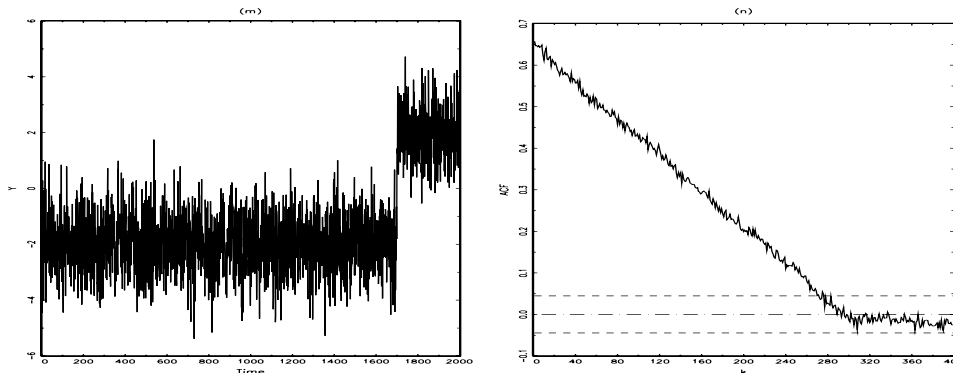


Figure 6: (a) and (b) are respectively the trajectory and the ACF of the Structural CHange model (2)-(11) with $T = 2000$, $\sigma_\varepsilon^2 = 0.19$, $\pi = 0.85$ and $\mu_0 = -\mu_1 = -1$.

break

$$\mu_t = \begin{cases} \mu_0 & \text{if } t = 1, \dots, m \\ \mu_1 & \text{if } t = m + 1, \dots, T, \end{cases} \quad (11)$$

where $m = \pi T$ and $\pi \in (0, 1)$. This model (2)-(11) has been introduced by Quandt (1958). We illustrate its behavior on Figure 6.

3 Estimation methods to detect long memory behavior

In the last two decades, many procedures have been proposed to estimate the fractional long memory parameter d in (1). In this paper, we focus on the GPH technique introduced by Geweke and Porter-Hudak (1983), the Exact Local Whittle method proposed by Philips and Shimotsu (2005) and the wavelet method introduced by Lee (2005). The testing procedure to detect the presence of long memory behavior is:

$$H_0 : d = 0 \quad \text{against} \quad H_1 : d \neq 0. \quad (12)$$

Under the null hypothesis H_0 , the process $(Y_t)_t$ has a short memory behavior and under the alternative, it has a long memory behavior. Now we specify different methods to estimate the parameter d .

3.1 The GPH technique

The GPH technique developed by Geweke and Porter-Hudak (1983) and Robinson (1995) is based on the use of log-periodogram $I(w_j)$ at frequency $w_j = 2\pi j/T$. For frequency near zero, Geweke and Porter-Hudak (1983) and

Robinson (1995) show that the parameter d can be consistently estimated using the following least squares regression

$$\ln\{I(w_j)\} = a - d \ln\{4\sin^2(w_j/2)\} + \epsilon_t, \quad j = 1, \dots, m. \quad (13)$$

where $(\epsilon_t)_t$ is a white noise. If the integer m grows slowly with respect to the sample size as for instance $m = \sqrt{T}$, then the ordinate least-square estimator \hat{d} is asymptotically normal with standard error equal to $\pi(6m)^{-1/2}$.

3.2 The Exact Local Whittle (ELW)

The Exact Local Whittle (ELW) method proposed by Shimotsu and Phillips (2005) avoids some approximations in the derivation of the Local Whittle estimator proposed by Künsch (1987) and Yajima (1989). The estimated value \hat{d}_{ELW} for d obtained from this method is such that:

$$\hat{d}_{ELW} = \text{Arg } \min_{d \in [d_1, d_2]} R(d), \quad (14)$$

where d_1 and d_2 are the lower and upper bounds of the admissible values of d such that $-\infty < d_1 < d_2 < \infty$ and,

$$R(d) = \ln G(d) - 2d \frac{1}{m} \sum_{j=1}^m \ln(\omega_j), \quad (15)$$

where m is the truncation parameter, $G(d) = \frac{1}{m} \sum_{j=1}^m I(\omega_j)$, and $I(\omega_j)$ is the periodogram. Under regular assumptions, the ELW estimator \hat{d}_{ELW} is asymptotically Gaussian

$$\sqrt{m}(\hat{d}_{ELW} - d) \rightarrow_d N(0, 1/4), \quad \text{when } T \rightarrow \infty. \quad (16)$$

3.3 The Wavelet method

In recent years, the use of wavelets have been extended to a lot of areas other than physics. In econometrics, Jensen and Whitcher (2000) and Jensen (1999) propose a wavelet estimator for the long memory parameter d . In this paper, we use the wavelet method proposed by Lee (2005) which permits to estimate d inside the interval $(0, 1.5)$. The estimated value of d is obtained by applying the least squares method using the following regression,

$$\ln(I_q^{(j)}) = c - 2(d - 1)\ln(\lambda_q) + \epsilon_t \quad \text{for } q = 1, 2, \dots, m. \quad (17)$$

where, c is a constant, $\lambda_q = \frac{2\pi q}{n}$ and

$$I_q^{(j)} = \frac{1}{2\pi T} \sum_{k=0}^{2^j-1} |\omega_j(k) \exp(i\lambda_q k)|^2, \quad q = 1, 2, \dots, m. \quad (18)$$

The number of frequencies m included in the regression is such that $m \rightarrow \infty$ and $m/n \rightarrow 0$ as $n \rightarrow \infty$. $\omega_j(k)$ is the wavelet transform of a times series $Y_t, t = 1, \dots, T = 2^j - 1$,

$$w_j(k) = 2^{-j/2} \sum_t Y_t \psi(2^{-j}t - k), \quad (j, k) \in Z^2. \quad (19)$$

The wavelet coefficient $w_j(k)$ is the detail coefficient at scale j and position k . Using this approach the \hat{d}_{WAVE} estimator, for $d \in (0, 1.5)$, verifies,

$$m^{1/2}(\hat{d}_{WAVE} - d) \rightarrow_d N(0, \frac{\pi^2}{24}) \text{ as } T \rightarrow \infty. \quad (20)$$

4 Simulation experiment

In this section, using Monte Carlo simulations, we investigate the performance of the estimation methods recalled in section 3. We use 1000 Monte Carlo realizations, sample size $T = 2000$. In each Table, from Table 1 to Table 9, we provide the estimation obtained for the parameter d assuming that the generating process is theoretically a FI(d) model (1). The standard deviations and the t-statistics are not reported in the paper and can be provided under request. We specify now the sets of parameters for each generating process used in the simulations:

- The mean plus noise model: $\sigma_\eta^2 = (0.05, 0.01, 0.1, 1)$ and $p = (0.001, 0.0025, 0.005, 0.01, 0.025, 0.05, 0.1)$.
- The Stopbreak model: $\sigma_\epsilon^2 = (0.05, 0.01, 0.1, 1)$ and $\gamma = (0.001, 0.01, 0.1, 1, 10, 100, 1000)$.
- The Markov switching model: $\sigma_\epsilon^2 = (0.05, 0.01, 0.1, 1)$ and $(p_{00}, p_{11}) = ((0.5, 0.99), (0.95, 0.95), (0.95, 0.99), (0.99, 0.99), (0.999, 0.999))$.
- The TAR model: $r = 0, d = 1$ and $\sigma_\epsilon^2 = (0.05, 0.01, 0.1, 0.2, 0.3, 0.4, 1)$.
- The sign model: $p = (0.001, 0.0025, 0.005, 0.01, 0.025, 0.05, 0.1)$.
- The SCH model: $\sigma_\epsilon^2 = (0.05, 0.01, 0.1, 1)$ and $\pi = (0.1, 0.2, 0.35, 0.5, 0.65, 0.8, 0.95)$.

In all Tables, we report the mean of the estimated value of the long memory parameter d using the GPH, the Exact Local Whittle and the Wavelet methods for the six previous models. We summarize now our results.

1. The mean plus noise model. In Table 1, we observe that the estimation performance is similar for the GPH and the ELW methods. In all cases and for all methods we cannot accept the null hypothesis of

short memory. The results show also that the estimated value of the fractional long memory parameter d depends on the two parameters σ_η^2 and the product pT . We observe that the increase of pT or σ_η^2 have similar effects on the estimated parameter \hat{d} whose value increases too. Indeed, when pT increases, the number of breaks increases and the estimated value of d tends to be close to one. This makes the mean plus noise process similar to an I(1) process. Thus, we can retain the following rule:

Rule 1: When the data generating process is the mean plus noise model, if pT is a finite non zero number and if $pT\sigma_\eta^2 \sim 1$, then, the estimated value of the fractional long memory parameter d in model (1) is different of zero.

Table 1: Mean estimate of the long memory parameter d when the DGP is the Mean Plus Noise model with $T = 2000$

σ_η^2	p(pT)	0.001(2)	0.0025(5)	0.005(10)	0.01(20)	0.05(100)	0.1(200)
0.01	\hat{d}_{GPH}	0.066	0.118	0.174	0.251	0.479	0.589
	\hat{d}_{ELW}	0.058	0.117	0.176	0.240	0.432	0.524
	\hat{d}_{WAVE}	0.209	0.236	0.237	0.241	0.336	0.373
0.05	\hat{d}_{GPH}	0.142	0.268	0.359	0.472	0.728	0.819
	\hat{d}_{ELW}	0.142	0.247	0.334	0.427	0.643	0.734
	\hat{d}_{WAVE}	0.241	0.266	0.273	0.332	0.474	0.577
0.1	\hat{d}_{GPH}	0.209	0.350	0.471	0.579	0.818	0.887
	\hat{d}_{ELW}	0.188	0.314	0.416	0.518	0.736	0.819
	\hat{d}_{WAVE}	0.267	0.269	0.321	0.377	0.575	0.670
1	\hat{d}_{GPH}	0.557	0.712	0.786	0.877	0.975	0.988
	\hat{d}_{ELW}	0.393	0.632	0.713	0.803	0.934	0.961
	\hat{d}_{WAVE}	0.349	0.458	0.557	0.659	0.862	0.925

We remark also that for different triplets (p, T, σ_η^2) , which gives the same value for the product $pT\sigma_\eta^2$, then the estimated parameter of long memory d is nearly the same. In Table 2, we report results which illustrate this fact. For example, when $pT\sigma_\eta^2 = 1$ with $(p, T, \sigma_\eta^2) = (0.05, 2000, 0.01)$ or $(0.01, 2000, 0.05)$ or $(0.005, 2000, 0.1)$, \hat{d} is equal respectively to 0.479, 0.472 and 0.47 when we use the GPH technique and when we use the ELW technique, we get respectively 0.432, 0.427 and 0.416 for \hat{d} .

2. The Stopbreak model. The Table 3 reports the mean estimated values for the long memory parameter d for this model. In all cases, we cannot accept the null hypothesis, except when the value of $\sigma_\epsilon^2 T / \gamma \rightarrow 0$. When the value $\sigma_\epsilon^2 T / \gamma$ is far from 0 and does not tend to ∞ , all methods fail to reject the null hypothesis of short memory. These results are summarized in the following rule,

Table 2: Mean estimate of the fractional integration parameter d using GPH method and the corresponding value of $pT\sigma_\eta^2$

$\sigma_\eta^2 p$	0.001	0.0025	0.005	0.01	0.05	0.1
0.01	0.06	0.118	0.178	0.25	0.479	0.589
$(pT\sigma_\eta^2)$	(0.02)	(0.05)	(0.1)	(0.2)	(1)	(2)
0.05	0.142	0.268	0.360	0.472	0.727	<u>0.820</u>
$(pT\sigma_\eta^2)$	(0.1)	(0.25)	(0.5)	(1)	(5)	(10)
0.1	0.209	0.350	0.471	0.579	<u>0.818</u>	0.887
$(pT\sigma_\eta^2)$	(0.2)	(0.5)	(1)	(2)	(10)	(20)
1	0.557	0.682	<u>0.786</u>	0.877	0.970	0.988
$(pT\sigma_\eta^2)$	(2)	(5)	(10)	(20)	(100)	(200)

Rule 2: When the data generating process is the Stopbreak model (2)-(5)-(6) if $\sigma_\epsilon^2 T/\gamma \sim R$, where R is a finite non zero number, then the estimated value of the fractional long memory parameter d in model (1) is significantly different from zero.

Table 3: Mean estimate of the long memory parameter d when the DGP is the STOPBREAK model with $T = 2000$

σ_ϵ^2	γ	0.001	0.01	0.1	1	10	100	1000
0.01	\hat{d}_{GPH}	0.997	0.998	0.977	0.634	0.089	0.001	0.004
	\hat{d}_{ELW}	0.991	0.994	0.950	0.563	0.092	-0.009	-0.012
	\hat{d}_{WAVE}	0.997	0.993	0.896	0.407	0.228	0.216	0.211
0.05	\hat{d}_{GPH}	0.999	0.995	1.001	0.944	0.427	0.097	0.008
	\hat{d}_{ELW}	0.994	0.992	0.983	0.898	0.385	0.029	-0.002
	\hat{d}_{WAVE}	1.001	1.002	0.992	0.788	0.302	0.215	0.219
0.1	\hat{d}_{GPH}	0.999	0.998	0.999	0.976	0.634	0.09	0.005
	\hat{d}_{ELW}	0.990	0.989	0.987	0.949	0.559	0.090	-0.008
	\hat{d}_{WAVE}	0.996	1.003	1.001	0.901	0.406	0.232	0.198
1	\hat{d}_{GPH}	1.004	0.997	1.002	0.998	0.985	0.636	0.089
	\hat{d}_{ELW}	0.993	0.992	0.992	0.987	0.953	0.564	0.094
	\hat{d}_{WAVE}	0.992	0.987	0.982	0.964	0.783	0.404	0.211

We can illustrate this rule with the following example. If we take $\sigma_\epsilon^2 T/\gamma = 20$ for different triplets $(\sigma_\epsilon^2, T, \gamma) = (0.01, 2000, 1)$ or $(0.1, 2000, 10)$ or $(1, 2000, 100)$, \hat{d} is respectively equal to 0.634, 0.634 and 0.636, using the GPH technique. Using the ELW, we got \hat{d} equal to 0.563, 0.559 and 0.564, and equal to 0.407, 0.406 and 0.404 using the Wavelet method, see Tables 4 and 5.

3. The Markov switching (MS-AR) model (2) and (7). In Table 6 we observe that the estimated value for \hat{d} seem to depend on the values taken by the couple (p_{00}, p_{11}) . For example, when $(p_{00}, p_{11}) = (0.5, 0.95)$, \hat{d} is close to zero. We know that when p_{00} and p_{11} are far from

Table 4: Mean estimate value of the parameter d using wavelet method and its corresponding value of $T\sigma_\epsilon^2/\gamma$.

$\sigma_\epsilon^2 \gamma$	0.1	1	10	100	1000
0.01	0.950	0.563	<u>0.092</u>	-0.009	-0.012
$(T\sigma_\epsilon^2/\gamma)$	(200)	(20)	(2)	(0.2)	(0.02)
0.1	0.987	0.949	0.559	<u>0.090</u>	-0.008
$(T\sigma_\epsilon^2/\gamma)$	(2000)	(200)	(20)	(2)	(0.2)
1	0.992	0.987	0.953	0.564	<u>0.094</u>
$(T\sigma_\epsilon^2/\gamma)$	(20000)	(2000)	(200)	(20)	(2)

Table 5: Mean estimate of the fractional integration parameter d using wavelet method and its corresponding value of $T\sigma_\epsilon^2/\gamma$

$\sigma_\epsilon^2 T \gamma$	2	2.5	3.34	4	5	6.67	10	13.34	20
\hat{d}	0.087	0.117	0.163	0.202	0.241	0.314	0.432	0.501	0.634

zero, then switches between the two regimes occur frequently and the behavior of the process seems close to the one of a white noise. In the other hand, when p_{00} and p_{11} are close to one, (for instance $(p_{00}, p_{11}) = (0.99, 0.99)$), $hat{d}$ is significantly different from zero. In that case the breaks do not occur frequently inside the data sets. Thus, when the underlying process is a Markov switching process, we have no specific rules with respect of the test we consider. Sometimes we can accept H_0 , sometimes no.

Table 6: Mean estimate of the long memory parameter d when the DGP is the MS-AR model with $T = 2000$

σ_ϵ^2	(p_{00}, p_{11})	(0.5, 0.99)	(0.95, 0.95)	(0.95, 0.99)	(0.99, 0.99)	(0.999, 0.999)
0.01	\hat{d}_{GPH}	0.001	0.139	0.244	0.555	0.565
	\hat{d}_{ELW}	0.003	0.157	0.283	0.685	0.977
	\hat{d}_{WAVE}	0.315	0.395	0.538	0.783	0.974
0.05	\hat{d}_{GPH}	-0.002	0.143	0.242	0.556	0.573
	\hat{d}_{ELW}	0.004	0.151	0.267	0.684	0.932
	\hat{d}_{WAVE}	0.303	0.386	0.535	0.774	0.900
0.1	\hat{d}_{GPH}	0.006	0.140	0.244	0.561	0.577
	\hat{d}_{ELW}	0.004	0.154	0.277	0.675	0.901
	\hat{d}_{WAVE}	0.289	0.377	0.545	0.760	0.840
1	\hat{d}_{GPH}	0.0005	0.139	0.238	0.552	0.566
	\hat{d}_{ELW}	0.001	0.145	0.244	0.621	0.617
	\hat{d}_{WAVE}	0.233	0.345	0.453	0.649	0.529

- The Threshold Auto-Regressive and the Sign models. The results reported in Tables 7 and 8 for these two models show that the estimated values for d are close for both models.

More generally, in these Tables we observe that when the value of σ_ϵ^2

Table 7: Mean estimate of the long memory parameter d when the DGP is the SETAR model with $T = 2000$

σ_ϵ^2	0.01	0.05	0.1	0.2	0.3	0.4	1
\hat{d}_{GPH}	0.004	0.001	0.729	0.562	0.263	0.136	0.023
\hat{d}_{ELW}	1.024	0.874	0.866	0.602	0.255	0.127	0.009
\hat{d}_{WAVE}	0.949	0.745	0.799	0.677	0.472	0.383	0.160

varied between 0.2 and 0.6 the alternative hypothesis of long memory behavior cannot be rejected. Thus, for these models, we have no specific rule to decide when the estimated procedure detect a long memory behavior.

Table 8: Mean estimate of the long memory parameter d when the DGP is the Sign model

p (pT)	0.001 (2)	0.0025 (5)	0.005 (10)	0.01 (20)	0.025 (50)	0.05 (100)	0.1 (200)
\hat{d}_{GPH}	0.840	0.886	0.791	0.631	0.347	0.155	0.051
\hat{d}_{ELW}	0.895	0.885	0.828	0.679	0.336	0.154	0.042
\hat{d}_{WAVE}	0.837	0.865	0.832	0.748	0.553	0.366	0.218

5. The Structural change model. This model is a linear process with one break which occurs at unknown time m . We use different values of noise's variance and different values for the date of break m . In Table 9 we observe that the performance to estimate d using all methods depends on the location of the break. The results depend on the noise's variance. For example, for large value of σ_ϵ^2 , the ELW method detects nearly always the long memory behavior whatever the localization of the break.

Table 9: Mean estimate of the long memory parameter d when the DGP is the SCH model with $T = 2000$

σ_ϵ^2	π	0.1	0.2	0.35	0.5	0.65	0.8	0.95
0.01	\hat{d}_{GPH}	0.976	0.957	1.013	0.623	1.01	0.958	0.965
	\hat{d}_{ELW}	0.991	0.986	0.963	0.953	0.977	0.988	0.970
	\hat{d}_{WAVE}	0.883	0.927	0.899	0.557	0.896	0.951	0.943
0.05	\hat{d}_{GPH}	0.910	0.904	0.931	0.604	0.923	0.902	0.867
	\hat{d}_{ELW}	0.914	0.900	0.864	0.876	0.888	0.895	0.911
	\hat{d}_{WAVE}	0.757	0.774	0.765	0.540	0.762	0.781	0.793
0.1	\hat{d}_{GPH}	0.848	0.855	0.867	0.592	0.872	0.857	0.786
	\hat{d}_{ELW}	0.851	0.794	0.777	0.787	0.805	0.832	0.835
	\hat{d}_{WAVE}	0.679	0.683	0.680	0.493	0.675	0.685	0.693
1	\hat{d}_{GPH}	0.515	0.547	0.529	0.433	0.529	0.547	0.399
	\hat{d}_{ELW}	0.276	0.315	0.347	0.397	0.434	0.450	0.467
	\hat{d}_{WAVE}	0.379	0.371	0.382	0.317	0.382	0.383	0.388

This simulation exercise shows that

1. If the underlying processes is a mean-plus-noise process or a stop break process, then whatever the estimation methods we use to estimate the long memory parameter d of the model (1), we reject the null assumption and retain the presence of persistence inside the data set.
2. If the underlying processes is a Markov switching process, a TAR process, a Sign process or a structural changes model, the results are mitigate. We cannot provide rules permitting to say when these models create spurious long memory behaviors, even if in some cases, they do: the rejection of the null depends strongly on the values of specific parameters of the considered model.

5 Application to the American and the French inflation rates

Inflation time series are subject to many domestic and external shocks. International events like oil price shocks, wars, financial crisis and monetary policies play an important role on the evolution of this variable. To describe its dynamics, practitioners and researchers use mainly two classes of models. The first class includes models with changes in mean like the Markov switching model of Hamilton (1989) and the structural change model introduced in the literature by Quandt (1958, 1960) and developed recently by Bai and Perron (1998, 2003). The second class includes long memory models like the FARMA model developed by Granger and Joyeux (1980) and Hosking (1981). If we use this last class of models to explain the evolution of inflation, this means that these series are characterized by existence of persistence. This fact has already been pointed out by Hassler and Wolters (1995), Baillie, Chung and Tieslau (1996), Bos, Franses and Ooms (1999) and Lee (2005) for instance.

We know that presence of breaks in a data set can be cause of misspecification for the underlying process. We will see that the problem can arise with inflation rate, even if, in fine, it is very difficult to conclude.

We consider monthly data sets corresponding to the French and the US inflation on the period January 1959 - April 2006, sample size $T=568$. The data set is provided from Datastream Base. Inflation rates are introduced computing $y_t = 1200 * [\log(p_t) - \log(p_{t-1})]$, where p_t is the Consumer Price Indices. The trajectories and the autocorrelation functions for the two inflation rates are provided in Figures 7 and 8 (a) and (b). From these Figures, we observe a slow decaying of the autocorrelation function which indicate

existence of long memory inside the data sets. In another hand, the trajectories of the series (seem to present some breaks that could be modelled using models with structural changes. Indeed, using the sequential method of Bai and Perron (1998, 2003), we detect four breaks in the French inflation series and three breaks in the US inflation series. Dates of breaks coincide exactly with some economics and financial events. For example, we detect, for the two time series, a break that corresponds to the Vietnam war which exerts a positive effect on the US inflation rate (April 1967). This shock is transmitted ten months after to the French economic (February 1968). The second detected break corresponds to the beginning of the first oil price crisis (January 1973). Another break is detected at the end of the second oil price shock, July 1981 for the US series and, August 1983 for the French one. These two last shocks correspond to international events that have affected the two economies at the same time even if their magnitudes and their durations are not been the same. Finally the fourth break detected on the French inflation series coincides with the crisis of the European Monetary system (September 1991).

Thus, it appears that these two inflation rates time series can be modelled using long memory process (study based on the behavior of the autocorrelation functions) and using structural change models (study based on a economic approach). To detect if the long memory behavior is spurious or not, we proceed in the following way:

1. In a first step, assuming that the data sets can be modelled using a FI(d) process, we estimate the long memory parameter d using GPH method with $m = T^{0.5}$, and $m = T/2$ for the ELW method, for the whole sample.
2. In a second step, we identify existence of breaks using Bai and Perron method. We get 4 breaks for the French inflation rate and 3 breaks for the American inflation rate.
3. We divide the whole sample in four subsamples for the French series and in three subsamples for the American series.
4. On each subsample, we estimate the long memory parameter d using a FI(d) process with the same methods as in the first step.
5. We compare the values obtained for \hat{d} for each series on the whole sample and on the subsamples.
6. We consider the following rule. If there is a "true" long memory behavior, we will find it on the whole sample and on the subsamples. If this behavior exists only on the whole sample and not on the subsamples, then we conclude that the breaks create a spurious long memory behavior.

Table 10: Estimate of the number of breaks and the fractional long memory parameter d for the French inflation series.

Periods	N. of breaks	GPH		ELW	
		\hat{d}_{GPH}	t-stat	\hat{d}_{ELW}	t-stat
P1 = 1959:01-1968:02	0	-0.114	-0.262	-0.038	0.916
P2 = 1968:03-1973:01	0	-0.085	-0.174	0.039	-0.655
P3 = 1973:02-1983:08	0	0.504	1.211	0.438	5.409
P4 = 1983:09-1991:09	1	0.696	4.613	0.502	5.567
P5 = 1991:10-2006:04	0	0.414	1.581	0.141	0.758
PT = 1959:01-2006:04	4	0.928	6.613	0.338	7.99

The results are not identical for the both series, see Tables 10 and 11. For the French inflation rate, the values of \hat{d} are significantly different from zero on the whole sample and not on the subsamples. This means that we are in presence of spurious long memory and it seems preferable to model this series with a model taking into account presence of breaks. We can chose it amongst the models proposed in Section 2. For the American inflation, the values of \hat{d} are always far from zero whatever the sample we use. Thus, it seems that persistence is an important feature of this last series. Nevertheless, we observe that $\hat{d} > 1$ on the whole period and $\hat{d} < 1/2$ on the subsamples. It seems reasonable, for this series to consider a model which is able to take into account regimes with long memory behavior.

Table 11: Estimate of the number of breaks and the fractional long memory parameter d for the US inflation series.

Periods	N. of breaks	GPH		ELW	
		\hat{d}_{GPH}	t-stat	\hat{d}_{ELW}	t-stat
P1 = 1959:01-1967:04	1	0.492	3.231	0.205	1.466
P2 = 1967:05-1973:01	0	0.408	2.220	0.481	4.637
P3 = 1973:02-1981:07	2	0.734	2.004	0.540	6.160
P4 = 1981:08-2006:04	1	0.302	1.704	0.428	7.983
PT = 1959:01-2006:04	3	1.038	7.564	0.483	12.915

6 Conclusion

This paper focus on the behaviors of the estimation methods for the long memory parameter under misspecification. Using several techniques to estimate the long memory parameter of a FI(d) process, we show, through an experiment study, that it can exist misspecification between some classes of structural changes models including mainly the mean-plus-noise and the Stopbreak and the FI(d) model. For the Markov-Switching, the Threshold autoregressive and the Structural Changes processes the results are mitigate.

In this paper, we have extended the work of Diebold and Inoue (2001), to a large class of models using different techniques of estimation.

Studying French and American inflation rates on a specific period, we detect several behaviors inside the data sets and we use a procedure which permits to privilege a model with respect to another one. We show that the French inflation rate may be characterized by a spurious long memory behavior if we do not take into account existence of breaks while the US inflation series can be characterized using mixing of long memory behavior and switching process in order to take into account its main features. Thus, the structural behavior of these two series appear slightly different from this empirical study.

References

- [1] Bai, J. (1998) "A note on spurious break". *Econometric theory*, 14, 663-669.
- [2] Bai J. et P. Perron, (1998), "Estimating and testing linear models with multiple structural changes". *Econometrica* 66: 47-78.
- [3] Bai, J. et P. Perron, (2003), "Computation and analysis of multiple structural change models," *Journal of Applied Econometrics*, vol. 18(1), pages 1-22.
- [4] Beran, J. (1994) "Statistics for Long-Memory Processes". Chapman and Hall, London.
- [5] Carrasco, M. (2002). "Misspecified Structural Change, Threshold, and Markov-switching models" *Journal of Econometrics* 109, 239-237.
- [6] Charfeddine, L. and Guégan, D. (2007). "Which is the best for the US Inflation rate: A structural changes model or a long memory process?". Working paper, 12-CES-Cachan, December.
- [7] 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.
- [8] Cheung, Y.W. (1993), "Long memory in foreign-exchanges rates", *Journal of Business and Economic Statistics*, 11, 1, 93-101.
- [9] Diebold, F.X., Inoue, A., (2001) "Long memory and regime switching". *Journals of Econometrics* 105, 131-159.
- [10] Engel, R.F., Smith, A.D., (1999) "Stochastic permanent breaks". *Reviews of Economics and Statistics* 81, 553-574.

- [11] Geweke, J., Porter-Hudak, S., (1983) "The estimation and application of long memory time series models". *Journal of Time series Analysis* 4, 15-39.
- [12] Granger, C.W.J., Ding, Z., (1996). "Varieties of long memory models". *Journals of Econometrics* 73, 61-78.
- [13] Granger, C.W.J. and Hyung, N. (2004). "Occasional structural breaks and long memory with application to the S&P500 absolute stock returns". *Journal of empirical finance*, 11, 399-421.
- [14] Granger, C.W.J., Joyeux, R., (1980). "An introduction to long memory time series and fractional differencing". *Journal of Time Series Analysis* 1, 15-39.
- [15] Granger, C.W.J., Teräsvirta, T., (1999) "A simple nonlinear time series model with misleading linear properties". *Economics Letters* 62, 161-165.
- [16] Gouriéroux, C. and Jasiak, J. (2001) "Memory and infrequent breaks". *Economics Letters*, 70, 29-41.
- [17] Guégan, D. (2005) "How can we define the concept of long memory ? An econometric survey", *Econometric Reviews*, 24 (2).
- [18] Hamilton, J.D. (1989). "A new Approach to the Economic Analysis of Non stationarity Times Series and the Business Cycle," *Econometrica* 57, 357-384.
- [19] Hassler, U., and Wolters J. (1995) "Long memory in inflation rates: international evidence", *Journal of Business and Economic Statistics*, 1337-45.
- [20] Hosking, J.R.M. (1981). "Fractional differencing". *Biometrika*, 68, 165-176.
- [21] Hyung, N., and Franses, P.H. (2002) "Inflation rates: long-memory, level shifts, or both? Working paper, Econometric Institute Report 08.
- [22] Jensen, J. (1999). "An Approximate Wavelet MLE of Short- and Long-Memory Parameters", *Studies in Nonlinear Dynamics and Econometrics*, 3, 4, 239-253.
- [23] Jensen, J. et Whitcher, B. (2000). "Time-Varying Long-Memory in Volatility: Detection and Estimation with Wavelets", Technical Report, Department of Economics, University of Missouri.

- [24] Künch, H.R. (1987) "Statistical aspects of self-similar processes". In proceeding of the first world congress of the Bernoulli Society, Tashkent (Yu. Prohorov and V. V. Sazonov, eds.) 1, 67-74. VNU Sciences Press.
- [25] Lee, J., (2005), "Estimating memory parameter in the US inflation rate", *Economics Letters* 87, 207-210.
- [26] Lee, J. et P.M. Robinson, (1996), "Semiparametric Exploration of Long Memory in Stock Prices". *Journal of Statistical Planning and inference*, 50, 155-174.
- [27] Lim, K.S. et H. Tong, (1980), "Threshold autoregressions, limit cycles, and data," *Journal of the Royal Statistical Society*, B 42, pp. 245-92.
- [28] Quandt, R. (1958) "The Estimation of the Parameter of a linear Regression System Obeying Two Separate Regimes", *Journal of the American Statistical Association* 53, 873-880.
- [29] Quandt, R. (1960) "Tests of the Hypothesis that a Linear Regression System Obeys Two Separate Regimes", *Journal of the American Statistical Associations* 55, 324-330.
- [30] Robinson P.M. (1995) "Log-periodogram regression of time series with long range dependence", *The annals of Statistics*, 23, 1048-1072.
- [31] Shimotsu, K. and Peter C.B.P. (2005) "Exact Local Whittle Estimation of Fractional Integration", *Annals of Statistics*, 33(4), pp. 1890-1933.
- [32] Yajima, Y. (1989). "A central limit theorem of Fourier transforms of strongly dependent stationary processes", *Journal of Time series analysis*, 10, 375-383.

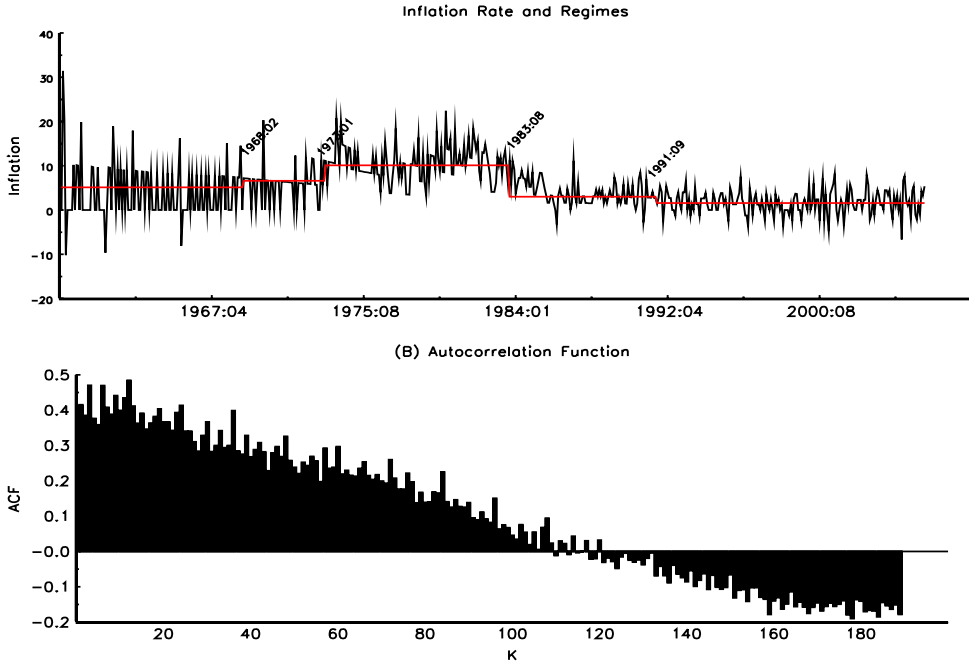


Figure 7: (a) and (b) are respectively the trajectory and the ACF of the French inflation time series.

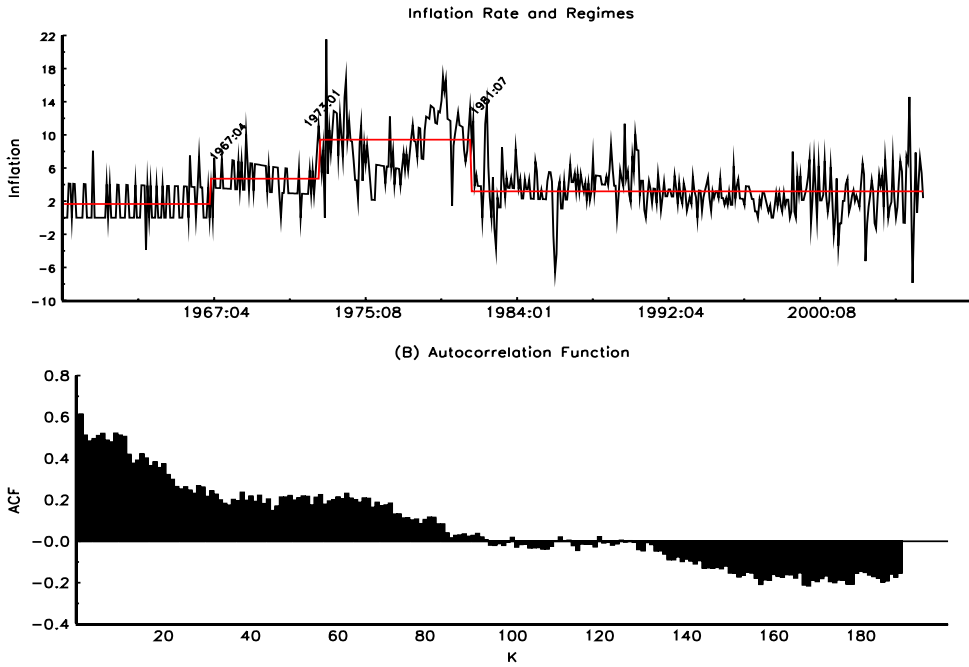


Figure 8: (a) and (b) are respectively the trajectory and the ACF of the US inflation time series.