A Varieties of Spurious Long Memory Process

Charfeddine Lanouar
Assistant Professor
University of Gabes
Institut Supérieur de Gestion de Gabes
Department Quantitative Methods
France
Email : lanouar_charf@yahoo.fr

Abstract

This paper introduces a large varieties of models which can generate a spurious long memory behavior. By using Monte Carlo simulations have investigated the behavior of the GPH, the Exact Local Whittle and the Wavelet methods when the data generating process is weakly dependent with changes in mean. The results show that these long memory estimation methods fail to accept the null hypothesis of short memory. A strong size distortion have been detected. The results stress the need for a test statistic which can discriminate the true long memory behavior from the spurious behavior. In the empirical applications we use stock returns of seven French companies with the CAC 40 stock market index. We show that the observed long memory behavior is a true behavior and that presence of breaks in these time series do not have a major effect on the long memory parameter.

Keywords: spurious long memory, structural change, empirical power.

JEL classification: C13;C15,C51.

1 Introduction

Modelling time series using long memory processes have largely increased since the seminal papers of Granger and Joyeux (1980) and Hosking (1981). Actually, it is well known that many financial time series exhibit strong persistence behavior. This long range dependence is generally detected in the squares or absolute values of the returns. At the same time, many empirical and theoretical works have shown that several structural changes models like the mean-plus-noise model of Chen and Tiao (1990), the stochastic permanent break model of Engle and Smith (1999), the sign model of Granger and Teräsvirta (1999), the Markov switching model of Hamilton (1989) and the Threshold Auto-Regressive (TAR) model of Lim and Tong (1980) among others can generate a long memory behavior in terms of decaying autocorrelationsand fractional value of the estimated parameter d. Only few works have examined the link between structural changes models and long memory behavior. Until now, no consensus has been reached about this problem. Diebold and Inoue (2001) argued that regime switching models and long memory behavior are intimately related. Granger and Hyung (2004) suggest that it is difficult to discriminate between these two classes of models. The existing empirical studies in testing for long-range dependence show that, when the data are weakly dependent with changes in mean, statistical tests for long-range dependence have a size distortion. They are not able to accept the null hypothesis of short memory with a high power. So, it appears that the existence of breaks in data sets can lead to a misspecified model selection.

This paper have three main objectives. First, we propose a varieties of structural change models that can generate spurious long memory behavior. Then, we study the possibility of discrimination between spurious long memory behavior and true behavior. To detect long memory behavior we use the following estimation methods: the GPH technique of Geweke and Potter-Hudak (1983), the Exact Local Whittle (hereafter ELW) proposed by Philips and Shimotsu (2005) and the wavelet method introduced by Lee (2005). Finally, we investigate the true nature of the observed long memory behavior in seven absolute returns French companies with the CAC 40 one.The remain of the paper is organized as follows: Section 2 reviews the definitions of long memory models. Section 3, presents long memory estimation methods. In Section 4, we introduces simultaneously the general form of spurious long memory models and we present the results of the simulations studies. Section 5 presents our empirical results concerning the eight French stock returns and Finally section six concludes.

2 Long memory process

In this section we present different alternatives definitions of long memory processes. These definitions will be a useful tools in the analysis of spurious long memory process. First, we introduce the typical ARFIMA model proposed by Granger and Joyeux (1980) and Hosking (1981).

Then, we present the GPH technique of Geweke and Poter-Hudak (1983), the Exact Local Whittle (ELW) method introduced in 2005 by Shimotsu & Phillips and the wavelet estimation method as in Lee (2005).

2.1 Long memory Definitions

In time domain, a series (y_t), $t = 1,...,T$, has a long memory behavior if its autocorrelation function decays hyperbolically to zero. In this case, the autocovariance function, γ_k, is defined by,

$$\gamma_k \approx c(k)\, k^{2d-1}$$

where d > 0 and c(k) is a slowly varying function at infinity. In frequency domain, long range dependence is defined in terms of rates of explosion of low-frequency spectra, this means that the spectral density tends to infinity for frequency near zero. The frequency domain definition of long memory is,

$f(\omega) \approx \infty$ as ω tends to 0

A third possible definition of long memory process is given by the $\text{Variance}(S_T) = O(T^{2d+1})$ where S_T is the partial sums of the time series (y_t), $S_T = \sum_{t=1}^{T}(y_t)$. The time domain, the frequency domain and the variance of partial sums definitions of long range dependence will be a useful tools to show that certain simple nonlinear models can generate a long memory behavior, see Diebold and Inoue (2001) for instance.

2.2 The ARFIMA model

The first long memory process introduced in the literature is the popular Auto-Regressive Fractionally Integrated Moving Average (ARFIMA) model developed independently by Granger and Joyeux (1980) and Hosking (1981).

A process $\{y_t\}_{\cdot 1}^{T}$, with $t \in N$ follows a stationary ARFIMA(p,d,q) process if it takes the following form,

$$\Phi(B)(1-B)^d(y_t - \mu) = \Theta(B)\, u_t$$

where $\Phi(.)$ and $\Theta(.)$ are polynomials of order p and q respectively, whose roots lie outside the unite circle and μ is unknown mean. (u_t) is a Gaussian strong white noise N(0, u_t^2) and B is the lag operator. If $\mathrm{d} \in (-1/2,0)$ the ARFIMA(p,d,q) model is an invertible stationary process with intermediate memory. Now, if $\mathrm{d} \in (0,1/2)$ model (1) is stationary and invertible. In the later case, the autocorrelation function $\rho(k)$ exhibits a slow decay when the lag k increases, see Beran (1994) for instance. This means that we are in presence of long memory behavior.

In the following we concentrate our analysis on the model (1) under the restriction p = q = 0 and 0 < d < 1/2. In that case the process is called a Fractionally Integrated noise FI(d).

3 Long Memory Estimation Methods

To test the presence of long memory inside a time series a large varieties of estimation methods have been proposed in the last two decades. Besides the GPH method of Geweek Poter-Hudak (1983) which is the most widely used in applications, we use also the Exact Local Whittle (ELW) method of Shimotsu & Phillips (2005) and the wavelet method (WAVE) recently by Lee (2005).

3.1 The GPH technique

The GPH technique of Geweke and Poter-Hudak (1983) is the most widely used method in empirical applications. Based on this method the estimated fractional long memory parameter d is obtained from the following least squares regression,

$$\ln\{I(\omega_j)\} = a - d\, \ln\{4\sin^2(\omega_j/2)\} + \varepsilon_t \tag{2}$$

With, $\mathrm{j} = 1,...,\mathrm{m}$, and $I(\omega_j)$ is the periodogram of the process $\{y_t\}_1^T$ at frequency $\omega_j = 2\pi j/T$, where T is the number of observations. Consistency requires that m grows with sample size, but at a slower rate. Diebold and Inoue (1999) adapt a popular rule of thumb of $m = p^T$. The ordinate least square estimator of d is asymptotically normal with standard error equal to $\pi(6m)^{(-1/2)}$, see Geweke and Poter-Hudak (1983).

3.2 The Exact Local Whittle

The second semi-parametric method, used in this paper, is the exact local whittle method of Shimotsu and Phillips (2005). This method avoids some of the strong assumptions of the Local Whittle (LW) method, see Yajima (1989).

Contrary to the LW method which is not consistent outside the stationary region $|d| \leq 1/2$, the ELW method have a good properties when the value of d is less than 9/2. In this case the ELW estimator is consistent and have the N(0, 1/4) limit distribution for all value of d when the optimization covers an interval of width less than 9/2. The estimated value $\widehat{d}_{ELW}$ is obtained as follow:

$$\widehat{d}_{ELW} = Arg \min_{d \in [d_1, d_2]} R(d) \tag{3}$$

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,

$$R(d) = \log \widehat{G}(d) - 2d \frac{1}{m} \sum_{j=1}^{m} \log(\omega_j) \tag{4}$$

With $\widehat{G}(d) = \frac{1}{m} \sum_{j=1}^{m} I_{\Delta_y^d}(\omega_j)$ where is the $I_{\Delta_y^d} = \frac{1}{2\pi T} \left| \sum_{t=1}^{T} \left(\Delta^d(y_t) \right) e^{it\omega} \right|^2$ is the periodogram of $\Delta^d(y_t) = (1-B)^d y_t$ .The ELW estimator satisfies the property $\sqrt{m}(\widehat{d}_{ELW} - d) \rightarrow_d N(0, 1/4)$, which induces a simple asymptotic inference.

3.3 The Wavelet method

In recent years, the use of wavelets have been extended to many areas other than physic. In econometrics, Jensen and Whitcher (2000) and Jensen (1999) have applied wavelets to estimate the fractional parameter of long memory. To start, consider a wavelet defined by

$$\psi_{a,b} = a^{-1/2} \psi\left(\frac{t-b}{a} \right) \tag{5}$$

where $a \neq 0$, a and b are respectively the dilation and translation parameter. $\psi(t)$ is called the mother wavelet which satisfies the condition $\int_{-\infty}^{+\infty} \psi(t) dt = 0$. Wavelets can be distinguished by their regularity and by their number of vanishing moments. The number of vanishing moments of a wavelet $\psi(t)$ is the largest integer υ which satisfies $\int_{-\infty}^{+\infty} t^{\upsilon} \psi(t)\, dt = 0$, for all $r = 0, ..., \upsilon - 1$. For some special values of a and b and a special choice of ψ, the wavelet $\psi_{a,b}$ constitutes an orthonormal basis for $L(R^2)$. As commonly used in the literature, we set $a = 2^j, b = k2^j$ with j and $k \in \mathrm{Z}$, then $\psi_{j,k}(t) = 2^{-j/2} \psi(2^{-j} t - k)$. The dyadic orthonormal wavelet transform of a time series (y_t), $t = 1, ..., T = 2^j - 1$ is defined as,

$$\omega_{j,k} =< y(t), \psi_{j,k}(t) >= 2^{-j/2} \int_R y(t) \psi(2^{-j} t - k) dt, \qquad (j,k) \in Z^2$$

The wavelet coefficient $\omega_{j,k}$ is the detail coefficient at scale j and position k. Our time series (y_t) can be easily reconstructed from its detail coefficients. When the scale j is large, the coefficient $\omega_{j,k}$ captures low frequency or coarse scale behavior of (y_t). In the other hand, when j is small, the wavelet coefficient captures the high-frequency or fine scale details of the signal (y_t).

The method proposed by Jensen is only robust for values of d lie inside (-1/2, 1/2), then when the expected value of *d* is in the non-stationary range we need a more robust test. Lee (2005) use a new method which is robust for a value of $d \in (0, 3/2)$, see also Kato and Masry (1999). In this paper we use the later method with the Haar wavelet and $\upsilon = 1$. The Haar wavelet $\psi(t) = 1$ if $0 < t < 1/2$ and $\psi(t) = -1$ if $1/2 < t < 1$ and 0 otherwise.

The spectral density of the wavelet transform at the scale j around zero frequency for $d \in (0, 3/2)$;

$$f_j(\lambda) = C_j |\lambda|^{-2(d-1)} g^2(\lambda) \quad as \quad \lambda \rightarrow 0 \tag{8}$$

where $C_j = c_j / 2\pi < \infty$ is a constant term with $g(t\lambda)/g(\lambda) = 1$ for all *t*, as $\lambda \rightarrow 0$ ¸ and $0 < g(0) < \infty$.. The spectral density $f_j(\lambda)$ at fixed scale j is given by,

$$I_q^{(j)} = \frac{1}{2\pi T}\sum_{k=0}^{2^j-1}\left|\omega_{j(k)\exp(i\lambda_q k)}\right|^2, \quad q = 1,2,\ldots,m \tag{9}$$

Where, $\lambda_q = 2\pi q/T$ and $m \to \infty$, $m/T \to 0$ as $T \to \infty$. From (8) a log-linear relationship is,

$$\ln(I_q) = \ln(C_j g(\lambda)^2) - 2(d-1)\ln(\lambda) \tag{10}$$

then we have the regression,

$$\ln(I_q) = \ln(C_j g(\lambda)^2) - 2(d-1)\ln(\lambda) + \varepsilon_t \quad \text{for} \quad q = 1,\ldots,m \tag{11}$$

which can be estimated by the ordinary least squares yielding the $\widehat{d}_{Wave}$ estimator consistent for $d \in (0,3/2)$.

$$m^{1/2}(\widehat{d}_{Wave} - d) \to_d N(0,\pi^2/24) \quad as \quad T \to \infty \tag{12}$$

4 Spurious Long Memory Models and Simulations Studies

In this section we introduce the general form of models that can create spurious long memory behavior in terms of linear property of the data such, the autocorrelation function, the spectrum or using the estimated value of the long memory parameter which lie, generally for this kinds of models, into (0,1). Consider the following model given by,

$$y_t = \mu_t + \varepsilon_t \tag{13}$$

where μ_t is a random variable and ²t is a strong white noise $N(0,\sigma_\varepsilon^2)$.

Assumption 1 : We assume that μ_t follows a random walk with drift in noise, $\mu_t = \mu_{t-1} + g(t)\eta_t$ with $g(t)$ is a bounded function between zero and one and η_t is a strong white noise $N(0,\sigma_\eta^2)$.

As shown in assumption 1 the occasional level shifts process μ_t is controlled by two variables, the function $g(t)$ which describes the nature of the switching process and the strong white noise η_t which describes the size of the jumps. Two trivial cases are of interests, when $g(t)$ is equal to zero the underlying process y_t is I(0) and when $g(t)$ is equal to one the processes y_t is I(1). In this paper we suppose that there are a continuous behavior of the fractional long memory parameter between zero and one for this process. Under this assumption the estimated value of *d* is expected to lies between 0 and 1. This expected value of *d* is caused by the presence of occasional shifts in the noise.

Proofs. If $g(t) = 0$, and $\mu_t = \mu_{t-1} = \mu$, then $y_t = \mu + \varepsilon_t$ which mean that y_t is I(0).

Now, if $g(t) = 1$, $\Delta y_t = \Delta\mu_t + \Delta\varepsilon_t$ and $\Delta\mu_t = \varepsilon_t$, then $\Delta y_t = \Delta\mu_t + \Delta\eta_t$ which mean that $\Delta y_t = \varepsilon_t + \Delta\eta_t$, y_t is I(1).

The only assumption that we have made about the function $g(t)$ is "' $g(t)$ is bounded between zero and one"'. Two large classes of $g(t)$ functions are to consider. Under the first class the functions $g(t)$ behave smoothly which make a gradually switching in data set. The second class of models includes processes with abrupt changes in mean.

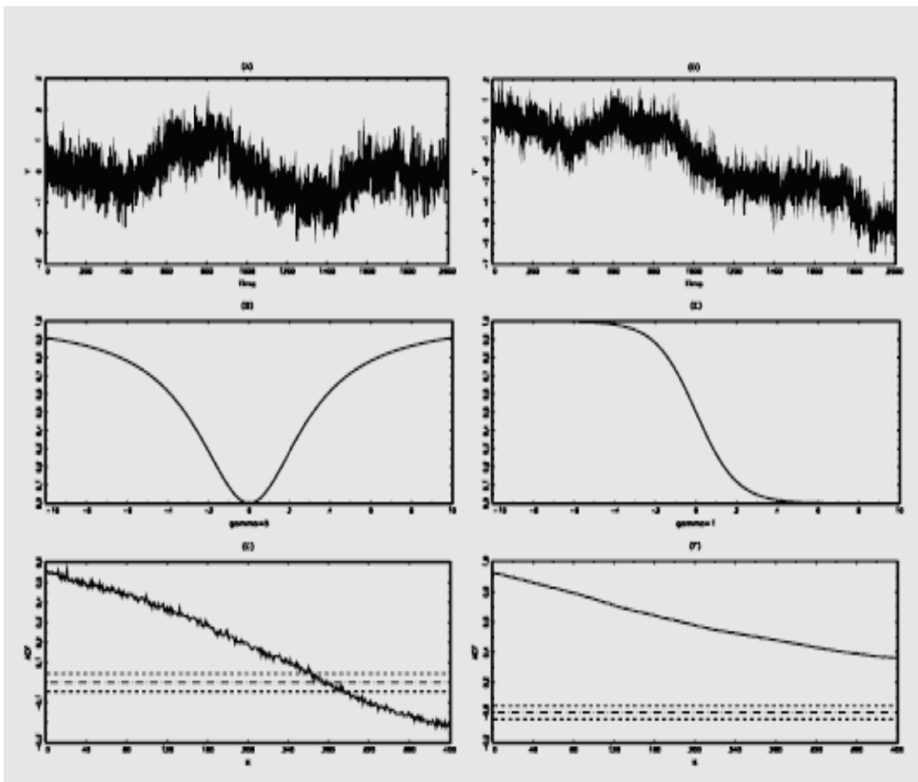
4.1 Gradually switching process

When $g(t)$ is a continuous function which decreases or increases smoothly between zero and one a gradually switching in mean occur in the series y_t. This gradually switching model can be easily confused with long memory model. Generally, when we estimate the fractional long memory parameter we obtain a value of d significantly different from zero. This value of d depends on the nature of the behavior of the function $g(t)$. We suppose here that $g(t)$ is a function of ε_t. To create a gradually switching process, many $g(t)$ functions can be proposed, Engel and Smith (1999) proposed a smoothly transition function $g(t) = \frac{\varepsilon_{t-1}^2}{\gamma + \varepsilon_{t-1}^2}$ which depends on a positive parameter γ. The nature of the behavior of the transition function depends on the parameter γ and the degree of the exponent of ε_t.

A second possibility is to use the logistic function, $g(t) = \dfrac{1}{1+\exp(\varepsilon_t)}$, which decreases between 0 and 1. We can enriched this function by adding a positive parameter γ in the denominator. Another possible function is $g(t) = \dfrac{\exp(\varepsilon_t)}{\gamma + \exp(\varepsilon_t)}$ which increases between 0 and 1. In these different cases, by varying C different behavior, especially in empirical finance, can be detected. These transition functions create a gradually switching process which induces a problem in inference statistics because model (13)-(A1).

Figure 1:

- (A), (B) and (C) are respectively the trajectory, the transition function and the ACF of the STOPBREAK model (13)-(A1) with T = 2000 and $\sigma_\varepsilon^2 = 0.25$.
- (D), (E) and (F) are respectively the trajectory, the transition function and the ACF of the STOPBREAK model (13)-(A1) with T = 2000 and $\sigma_\varepsilon^2 = 0.25$.



with $g(t)$ follows one of the transition function described above among others, can generate a long memory behavior. A graphical representations of the trajectories, the transition functions and the autocorrelation functions when we use the first and the third $g(t)$ functions are reported in figure 1. To investigate the behavior of the previous long memory estimation methods we conduct some experiment simulations3 . We generate 1000 series of the process (13)-(A1) when $g(t)$ is the transition function proposed by Engel and Smith (1999), the logistic transition function and $g(t) = \dfrac{\exp(\varepsilon_t)}{\gamma + \exp(\varepsilon_t)}$. The simulations results are reported in table 1 for different values of the parameter γ. We use a sample size T = 2000 for the GPH and the ELW methods and $T = 2^{11} = 2048$ for the wavelet method.

Table 1: Mean estimate of the long memory parameter d when the DGP is the process (13)-(A1) with $\sigma_{\varepsilon}^{2}=1$.

g(t)	λ	1	5	10	20	50	100	1000
$g(t)=\frac{\varepsilon_{t-1}^{2}}{\gamma+\varepsilon_{t-1}^{2}}$	$\widehat{d}_{GPH}$	0.994	0.998	0.987	0.949	0.815	0.633	0.097
	$\widehat{d}_{ELW}$	0.987	0.979	0.953	0.891	0.732	0.564	0.094
	$\widehat{d}_{WAVE}$	0.995	0.955	0.898	0.791	0.573	0.407	0.227
$g(t)=\frac{1}{\gamma+\exp(\varepsilon_{t})}$	$\widehat{d}_{GPH}$	1.000	0.999	0.998	0.943	0.533	0.268	0.014
	$\widehat{d}_{ELW}$	1.000	0.999	0.998	0.943	0.533	0.268	0.014
	$\widehat{d}_{WAVE}$	1.000	0.999	0.998	0.679	0.430	0.253	0.003
$g(t)=\frac{\exp(\varepsilon_{t})}{\gamma+\exp(\varepsilon_{t})}$	$\widehat{d}_{GPH}$	0.993	0.978	0.949	0.886	0.704	0.519	0.054
	$\widehat{d}_{ELW}$	0.983	0.947	0.898	0.814	0.623	0.432	0.050
	$\widehat{d}_{WAVE}$	0.964	0.884	0.783	0.663	0.460	0.344	0.211

The simulations results are consistent with the fact that when $g(t)=0$ the expected value of d is close to zero and when $g(t)=1$ the value of d is close to one. This result is observed in table 1, for example when γ tends to 1 the process y_{t} behaves like an I(0) process. The results show that all long memory estimation methods do not have the same behavior. The wavelet method have the worst performance when the value of *d* approach zero and the best performance when the value of d is expected to be higher than 0.3. The ELW method performs better than the GPH method. Nevertheless, these three methods cannot accept the null hypothesis of short memory only for large value of γ (when $g(t)\to 0$). The results reported in table 1 confirm that when the data is weakly dependent with switch in mean a spurious long memory behavior is detected. As we have noted in previous sections the behavior created by the transition function $g(t)$ has an important role in empirical applications. The intuition behind this follows from the fact that when the function $g(t)$ varies between the two extremes value the expected value of d varies, generally, between 0 and 1. This means that these models can capt shocks varying between transitory to permanent shocks.

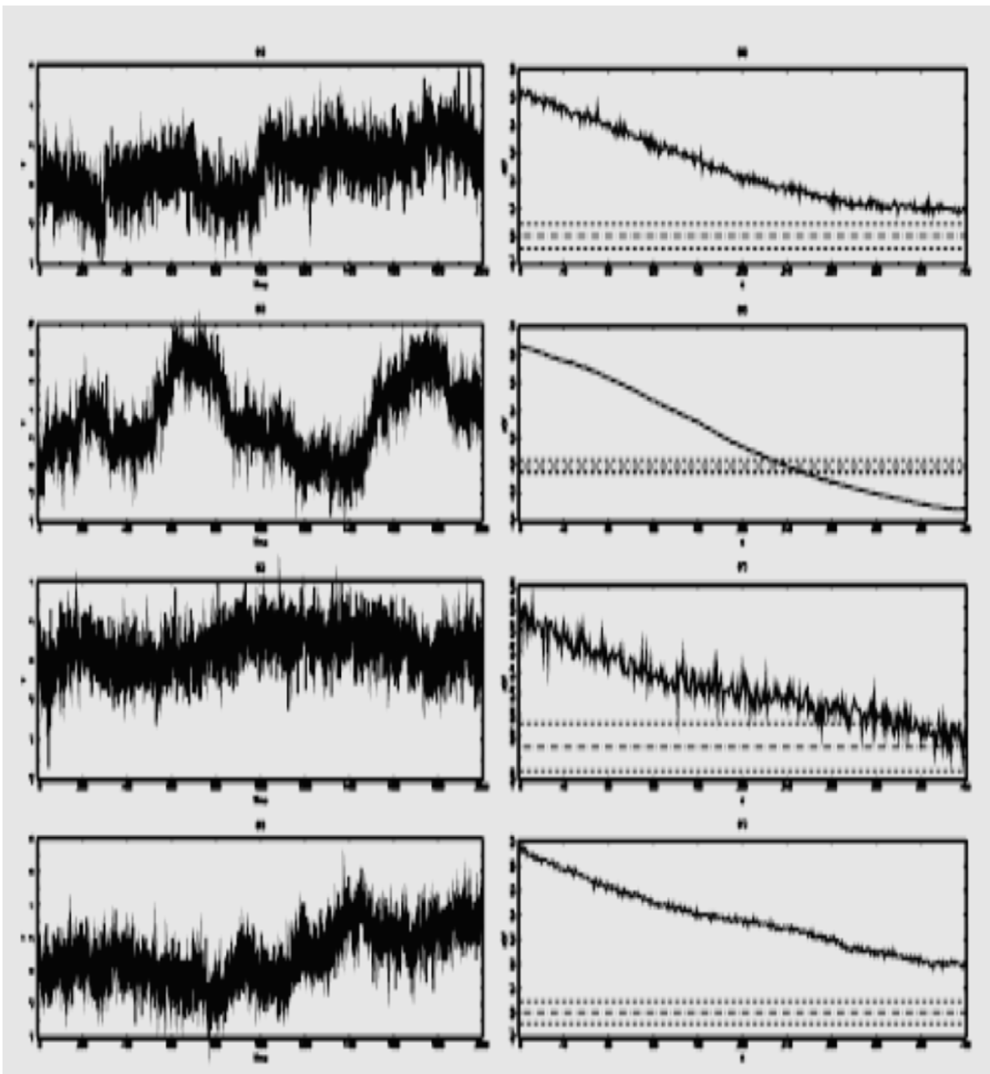
4.2 Abrupt switching process

Now, we assume that $g(t)$ is a discrete function which changes brutally between 0 and 1. Cheen and Tio (1990) proposed the binomial distribution as a process which governs the switch of the function g(t), then g(t) = 0 with probability (w.p) $1-p$ and $g(t)=1$ w.p p for more details see Cheen and Tio (1990). We can propose many others process for the function $g(t)$ like the Markov switching model developed by Hamilton (1989), the Threshold Auto-Regressive model introduced by Lim and Tong (1980) and the structural change model of Quandt (1958) among others. In the case of Markov switching model, the transition between the two regimes is governed by a Markov chain s_{t} which takes values zero and one. The probabilities of transition between the two states is defined by $\mathrm{p_{ij}=(s_{t}=i|s_{t-1}=j)}$, then $g(t)=0$ when $s_{t}=0$ and $g(t)=1$ when $s_{t}=1$. It is also possible to allow to the process which governs $g(t)$ to depend in an observed variable, generally y_{t-l}, with l is the delay parameter. For this SETAR specification, $g(t)=0$ if $y_{t-l}\leq r$ where r is the threshold parameter and $g(t)=1$ if $y_{t-l}>r$, in our simulations we set l = 1 and r = 0. Finally, we propose the structural change specification where $g(t)=0$ if $t<\pi T$ and $g(t)=1$ otherwise, where π lies into (0,1) and πT is the date of break. The process (13)-(A1) and $g(t)$ follows one of the processes described above have a same properties than long-range dependence models. This behavior can be observed in the autocorrelation function which decays slowly to zero or in the estimated long memory parameter d which is significantly different from zero. Above we give some graphical representations of the trajectories and the autocorrelation functions of models (13)-(A1) when the function $g(t)$ follows the binomial distribution, the Markov switching, the SETAR and the SCH models. Basis in the linear properties of these models different model can be proposed to describe this time series.

In order to investigate the behavior of the previous long memory estimation methods when the data generating process is (13)-(A1) with an abrupt changes we conduct some empirical simulations. We generate 1000 series of the process (13)-(A1) when $g(t)$ follows the binomial distribution, the Markov switching process (MS-AR), the TAR process and the Structural CHange (SCH) model.

Figure 2:

- (A), (C), (E) and (G) are the trajectories of model (13)-(A1) with $g(t)$ follows abrupt switching process, T = 2000 and $\sigma_\varepsilon^2 = 0.25, \gamma = 5, g(t) = \dfrac{\varepsilon_{t-1}^2}{\gamma + \varepsilon_{t-1}^2}$
- (B), (D), (F) and (H) are the ACF of model (13)-(A1) with $g(t)$ follows abrupt switching process, T = 2000 and $\sigma_\varepsilon^2 = 0.25, \gamma = 1, g(t) = \dfrac{1}{\gamma + \exp(\varepsilon_t)}$



The results are reported in table 2 and 3. The wavelet method behaves better than the GPH and the ELW methods. The estimated value of d depends on the parameters that governs each model, the probability p for the binomial distribution, the noise's variance σ_η^2 for the TAR model and the probabilities of transition (p_{00}, p_{11}) for the Markov switching model. For the SCH model, the results seem do not depend on the date of break πT. In a simulations not reported here the estimated value of d is a positive function of σ_η^2, see Charfeddine and Guégan (2006).

Table 2: Mean estimate of the long memory parameter d when the DGP is the process (13)-(A2) model with $\sigma_\varepsilon^2 = 1$.

g(t)	p	0.001	0.0025	0.005	0.01	0.025	0.05 1	1
Binomial distribution	$\hat{d}_{GPH}$	0.457	0.712	0.786	0.877	0.923	0.975	0.988
	$\hat{d}_{ELW}$	0.393	0.632	0.713	0.803	0.892	0.934	0.961
	$\hat{d}_{WAVE}$	0.353	0.473	0.566	0.677	0.793	0.870	0.925
g(t)		0.001	0.005	0.01	0.05	0.1	0.4 1	1
TAR Model	$\hat{d}_{GPH}$	0.471	0.686	0.751	0.828	0.831	0.772	0.705
	$\hat{d}_{ELW}$	0.417	0.589	0.640	0.696	0.705	0.677	0.692
	$\hat{d}_{WAVE}$	0.313	0.459	0.521	0.627	0.639	0.653	0.691
g(t)		0.1	0.2	0.35	0.5	0.65	0.8	0.95
SCH Model	$\hat{d}_{GPH}$	0.760	0.907	0.966	0.994	0.999	1.005	1.009
	$\hat{d}_{ELW}$	0.875	0.933	0.960	0.971	0.978	0.986	0.987
	$\hat{d}_{WAVE}$	0.895	0.960	0.965	0.988	0.979	0.995	0.987

Table 3: Mean estimate of the long memory parameter d when the DGP is the process (13)-(A2) model with $g(t)$ follows the MS-AR model and $\sigma_\varepsilon^2 = 1$.

(p_{00}, p_{11})	(0.5, 0.95)	(0.95, 0.95)	(0.95, 0.99)	(0.99, 0.99)	(0.999, 0.999)
$\hat{d}_{GPH}$	0.983	0.999	1.01	0.990	0.824
$\hat{d}_{ELW}$	0.976	0.992	0.977	0.989	0.819
$\hat{d}_{WAVE}$	0.913	0.984	0.948	0.978	0.851

From table 2 and 3 it follows that all estimation methods of long memory have a strong size distortion. The hypothesis of long memory cannot be rejected with higher power. This means that when the data have breaks in mean a long memory behavior can be observed. Then, by using statistical inference researchers believe that the data generating process is a long memory model even if the true DGP is models with shifts in mean.

Assumption 2 We assume that $\mu_t = \mu_0(1-h(t)) + \mu_1 h(t)$ where $h(t)$ is an indicator function which takes values 0 and 1.Now, we assume that the data generating process is (13)-(A2). The process which governs the transition between μ_0 and μ_1 varies across models. When $h(t) = s_t$, with s_t is an unobserved exogenous Markov chain, the process (13)-(A2) is the Markov switching model of Hamilton (1989). In the case of Threshold Auto-Regressive model $h(t) = 0$ if $y_{t-l} < r$ and $h(t) = 1$ if $y_{t-l} \geq r$, with l is the delay parameter and r is the threshold parameter. For the Structural CHange model $h(t) = 0$ if the t^{th} observation is less than πT and $h(t) = 1$ if the t^{th} observation is higher than πT with $\pi \in (0,1)$. For these kinds of models the separation between the two regimes is more clear which means that the changes in mean is more abrupt than (13)-(A1).We use experiment simulations to study the performance of the GPH, the ELW and the wavelet methods when the data generating process is the process (13)-(A2). We generate 1000 series and we estimated the mean value of the long memory parameter d. The results are reported in table 4 and 5.

Table 4: Mean estimate of the long memory parameter d when the DGP is the process (13)-(A2) model with $h(t)$ is governed by an MS-AR model and $\sigma_\varepsilon^2 = 1$.

(p_{00}, p_{11})	(0.5, 0.95)	(0.95, 0.95)	(0.95, 0.99)	(0.99, 0.99)	(0.999, 0.999)
$\hat{d}_{GPH}$	0.001	0.139	0.238	0.552	0.566
$\hat{d}_{ELW}$	0.001	0.145	0.244	0.621	0.617
$\hat{d}_{WAVE}$	0.291	0.340	0.458	0.633	0.543

Table 5: Mean estimate of the long memory parameter d when the DGP is the process (13)-(A2) model with $\sigma_{\varepsilon}^2 = 1$.

$h(t)$ is governed by	$\sigma_{\varepsilon}^2 = 1$	´0.01	0.05	0.1	0.2	0.3	0.4	1
TAR Model	$\hat{d}_{GPH}$	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
$h(t)$ is governed by	π	0.1	0.2	0.35	0.5	0.65	0.8	0.95
SCH Model	$\hat{d}_{GPH}$	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
	$\hat{d}_{WAVE}$	0.385	0.390	0.386	0.391	0.389	0.388	0.391

The results confirm that the wavelet method is not robust for a value of the parameter $d < 0.3$. When the process $(\mu_t)_t$ follows the MS-AR model, the TAR model or the SCH model the performance of estimation methods are ambiguous. For the MS-AR model the GPH method have the best performance. In the case of TAR model, the performance of long memory estimation methods depend on σ_{η}^2. For large value of noise's variance the ELW performs better than the GPH methods excepts when σ_{η}^2 is very small (<0.2). The wavelet method have the worst performance. When the DGP are the SCH model, the performance of the long memory estimation methods depend in the location of breaks. For a value of $\sigma_{\eta}^2 < 0.5$ the ELW have the best performance. When $\sigma_{\eta}^2 > 0.5$ the wavelet performs better than the two others methods. The results reported in table 4 and 5 confirm that the behavior generated by a weakly dependent data with changes in mean can be easily confused with a long memory behavior. In the following section, we use absolute returns of several French companies and the CAC 40 stock market to investigate the true behavior of the detected long memory in the absolute returns.

5 Empirical Applications

This section is devoted to analyze the long range properties of absolute returns for seven French companies with the CAC 40 index. We use a daily data from 01/01/1994 until 31/12/2003 obtained from Datastream Base. This data provides us a large sample size of T = 2606 observations. Trajectories and autocorrelation functions of the absolute returns are reported respectively in figures 5 and 6. In all autocorrelation functions a slowly decaying are observed. For most absolute returns autocorrelations remain significantly different from zero for a lags higher than 260 (1 years) which means that absolute returns are characterized by a strong persistence behavior.

Table 6: Estimate values of *d* for the different absolute returns of the ELW and GPH methods.

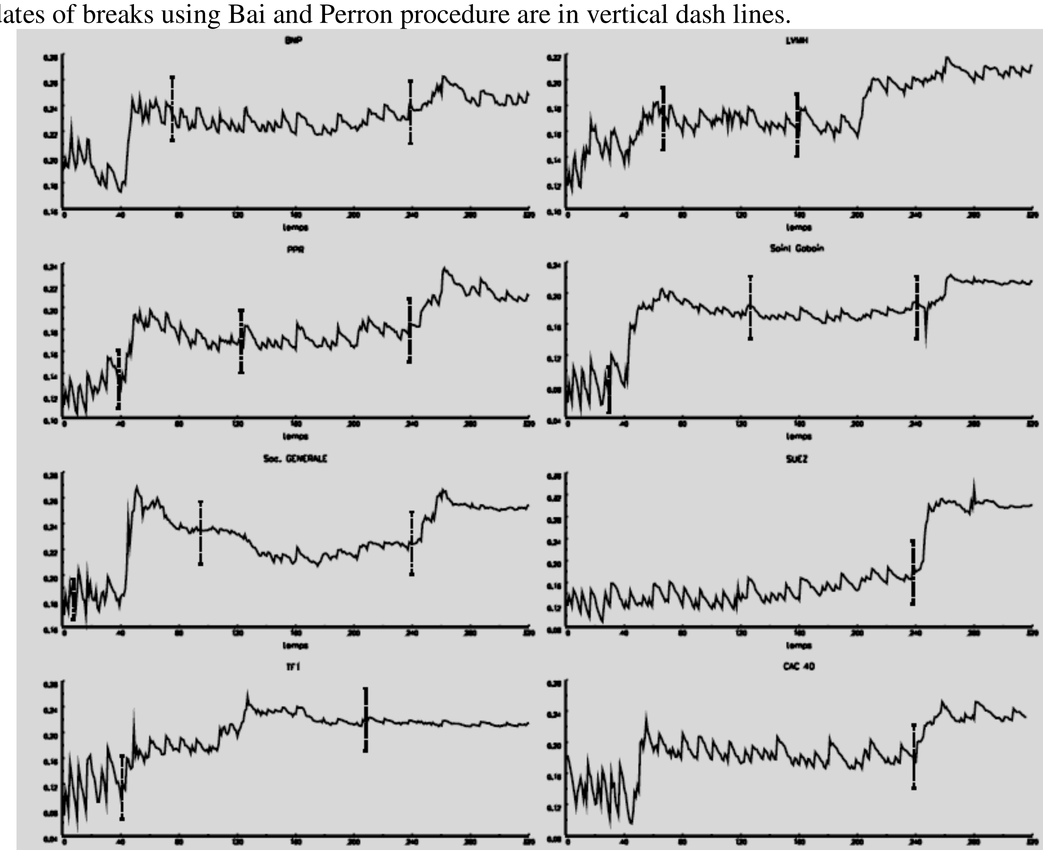
	ELW		GPH	
m	$T/2$	$T/4$	$T^{0.5}$	$T^{0.75}$
BNP	0.256 (11.28)	0.248 (7.576)	0.559 (5.067)	0.280 (7.712)
LVMH	0.247 (10.64)	0.209 (5.605)	0.435 (3.741)	0.243 (6.263)
PPR	0.233 (9.623)	0.210 (5.615)	0.552 (6.281)	0.292 (8.792)
Saint Gobain	0.238 (9.984)	0.216 (5.915)	0.493 (4.120)	0.226 (6.110)
Soc. Générale	0.244 (10.382)	0.255 (7.933)	0.559 (7.144)	0.317 (9.242)
Suez	0.314 (15.45)	0.300 (10.24)	0.572 (6.852)	0.297 (8.553)
TF1	0.261 (11.60)	0.214 (5.819)	0.546 (5.807)	0.224 (5.983)
CAC 40	0.211 (7.992)	0.232 (6.733)	0.575 (5.378)	0.277 (7.021)

t-stats are in parentheses.

In the empirical application we concentrate our analysis only on the GPH and the ELW methods. For the GPH method we use a values of *m* equal to $T^{0.5}$ and $T^{0.75}$. and T/2 and T/4 for the ELW method. The estimated values of the fractional long memory parameter are reported in table 6. The results show that all absolute returns are characterized by a long range dependance behavior. The results show that the GPH method is sensitive to the frequency m. These results are consistent with some works that have stressed that a values of m higher than T0:5 are more consistent with the hypothesis of long memory behavior. Moreover, when we use the GPH method with $m = T^{0.75}$ the obtained estimators values of d are close to those obtained from the ELW method. So, our empirical results confirm that a larger value of the frequency *m* is more consistent with the long range dependence hypothesis. Unlike the GPH method, the ELW estimator of d is not sensitive to the frequency *m*. This latter method is more efficient than the GPH one, for a details reviews see Banerje and Ungra (2004).

Figure 3:

- Plot of the Time-Varying long memory parameter d for the BNP, LVMH, PPR, Saint Gobain, Société Générale, Suez, TF1 and the CAC 40 absolute returns.
- The dates of breaks using Bai and Perron procedure are in vertical dash lines.



The question of the true nature of the long memory in absolute returns have interesting many researchers, for instance see Lobato and Savin (1998) and Granger and Hyung (2004). In order to investigate if the observed long memory is due to structural breaks or it is a true behavior generated by the data mechanisms, we propose the following strategy. First, we estimate a time varying values of *d*, we start at T=1000 observations and we increment the sample size by 5 observations (1 week) each time. This step have a two mains objectives. First, it allows us to detect the dates of shifts in the long memory parameter. Then, it allows to us to compare the dates of changes on the long memory parameter with breaks dates obtained by using the sequential procedure of Bai and Perron (1998, 2003). The idea behind this follows from the fact that if the long memory in absolute returns is due to the presence of structural breaks then the changes in long memory parameter should coincide with structural breaks dates.

In figure 3 we plot the time varying parameter d with the dates of breaks (dash line) as detected using the sequential procedure of Bai and Perron (1998, 2003). For the seven companies and the CAC 40 index the breaks dates do not coincide with the dates of changes on the long memory parameter. This means that for these absolute returns the observed long memory behavior is not due to the presence of breaks.

Moreover, we investigate the possibility that the long memory parameter is over estimated because of the presence of breaks. So, we start by filtering out the breaks and then we re-estimate another one the long memory parameter. The results are summarized in table 7. It follows from columns 2-5 that no major difference is detected compared with the results obtained from the original time series. Only a small changes in the long memory parameter have been occurred. This means that the presence of changes in these time series have a little affects on the estimated value of *d*. This can be explained by the small size of jumps. In other word, as shown in many simulations experiment the value of d depends on the size of jumps and the dates of breaks

Table 7: Estimate values of d for the different absolute returns after filtering out the breaks.

	ELW		GPH	
m	$T/2$	$T/4$	$T^{0.5}$	$T^{0.75}$
BNP	0.231 (9.486)	0.213 (5.763)	0.391 (4.387)	0.248 (6.883)
LVMH	0.235 (9.768)	0.189 (4.554)	0.447 (4.450)	0.239 (6.336)
PPR	0.209 (7.898)	0.172 (3.665)	0.265 (2.988)	0.230 (6.933)
Saint Gobain	0.211 (7.992)	0.171 (3.610)	0.262 (2.694)	0.176 (4.83)
Soc. Générale	0.228 (9.212)	0.233 (6.81)	0.387 (3.593)	0.254 (7.084)
Suez	0.291 (13.78)	0.265 (5.845)	0.259 (2.763)	0.225 (6.492)
TF1	0.241 (10.143)	0.172 (3.681)	0.342 (3.010)	0.170 (4.552)
CAC 40	0.197 (7.010)	0.215 (5.890)	0.417 (3.946)	0.232 (6.072)

t-stats are in parentheses.

Now, in order to confirm previous results we concentrate our analysis only in the CAC 40 stock market returns. We use the method proposed by Hsu (2005) and Hsu and Kuan (2000). First, we present briefly the Hsu and Kuan Local whittle methodology. Then, we applied it to absolute returns and we show that the detected breaks has a little impact on the estimated long memory parameter *d*.

To start consider the model given by (13)-(A2) where $h(t)$ is equal to zero before the observation T_0 and 1 after that. We rewrite the model as follows,

$$y_t = \begin{cases} \mu_0 + \varepsilon_t & \text{if} \quad t = 1,...,T_0 \\ \mu_1 + \varepsilon_t & \text{if} \quad t = T_0 + 1,...,T \end{cases} \tag{14}$$

Hsu and Kuan (2000) and Hsu (2005) propose the following method: First, they propose to estimate μ_0 and μ_1,

$$\widehat{\mu}_0 = \frac{1}{\tau}\sum_{t=1}^{\tau} y_t, \quad \widehat{\mu}_1 = \frac{1}{T-\tau}\sum_{t=\tau+1}^{T} y_t \tag{15}$$

Then, for each value of τ the residuals are

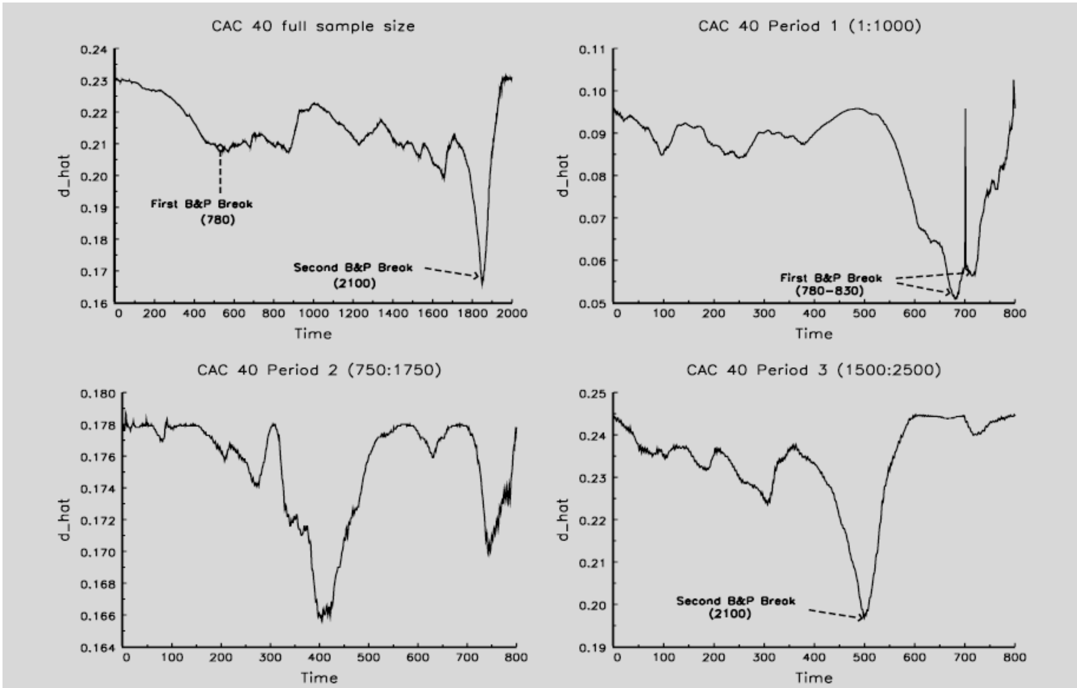
$$\widehat{\varepsilon}_t = \begin{cases} y_t - \widehat{\mu}_0 & \text{for} \quad \mathrm{t} \le \tau \\ y_t - \widehat{\mu}_1 & \text{for} \quad \mathrm{t} \ge \tau + 1 \end{cases} \tag{16}$$

And finally, to select the optimal date of change-point τ, Hsu and Kuan (2000) propose to take

$$\widehat{\tau} = ArgMin_{\tau \in [\tau_1, \tau_2]} LW(\widehat{d}_\tau, \tau) \tag{17}$$

where $[\tau_1, \tau_2] \subseteq (0,1)$ and the estimator of *d* is $\widehat{d}(\widehat{\tau})$. The authors show by using empirical simulations that the finite-sample distribution of the proposed estimator is very close to the normal one both for the short- and long-memory data, see Hsu and Kuan (2000). The principal advantage of this method lie on the jointly estimation of the long memory parameter *d* and the date of break τ . In our empirical applications we use the ELW estimator $I_{\widehat{\varepsilon}_t} = \dfrac{1}{2\pi T}\left|\sum_{t=1}^{T}(\widehat{\varepsilon}_t)e^{it\omega}\right|^2$.

Figure 4: Plots of the time-varying long memory parameter d as estimated using Hsu and Kuan (2000) method.



We applied this method to the CAC 40 absolute returns. First, we consider the full sample size (PT) and then we consider three subperiods, P1=1:1000, P2=750:1750 and P3=1500:2500 observations. We use $m = T/4, \tau_1 = 0.1*T, \tau_2 = 0.9*T$. The results reported in figure 4 show that the obtained dates of changes in the estimated long memory parameter coincide exactly with Bai and Perron breaks. Moreover, these changes are not very strong. This means that the long memory behavior in the CAC 40 absolute returns is a true behavior. Nevertheless, the presence of breaks in the CAC 40 absolute returns amplifies the long memory parameter by a value around 0.04. Note that for the second period the higher change of d is around 0.012 (0.178- 0.166). This change is not so strong to consider it as a veritably jump in the d parameter.

6 Conclusion

This paper introduces a general form of spurious long memory models. Two classes of models are proposed. They differ by the nature in which the changes occur and by the process in which it is governed. We use the most three robust methods of estimating long memory in order to study their behavior and their robustness when the data is weakly dependent with changes in mean. The results confirm that a spurious long memory behavior can be detected and this behavior can be confused with the true one. The results express the need for a new test statistic which can be robust in these cases. In the empirical applications we show that the observed long memory behavior in the absolute returns of the seven French companies with the CAC 40 stock market index is a true behavior. We show also that the detected breaks have a little effects, and it amplifies the value of d only by 0.04. Moreover, we show that the fractional long memory parameter have a time varying behavior.

References

[1] Bai J., Perron P. 1998. "Estimating and testing linear models with multiple structural changes". *Econometrica* 66: 47-78.

[2] Beran, J. (1994) "Statistics for Long-Memory Processes". Chapman and Hall, London.

[3] Banerjee A., and G. Urga (2004). "Modelling structural breaks, long memory and stock market volatility: an overview". *Journal of Econometrics*, forthcoming.

[4] Charfeddine, L. and Guégan, D. (2006). "Is it possible to discriminate between structural breaks model and long memory processes: Empirical evidence in the French and US Inflation rates?". Working paper, 12-CES-Cachan, December. 18.

[5] 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.

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

[7] Engel, R.F., Smith, A.D., (1999) "Stochastic permanent breaks". *Reviews of Economics and Statistics* 81, 553-574.

[8] Geweke, J., Porter-Hudak, S., (1983) "The estimation and application of long memory time series models". *Journal of Time series Analysis* 4, 15-39.

[9] 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.

[10] 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.

[11] Granger, C.W.J., Teräsvirta, T., (1999) " A simple nonlinear time series model with misleading linear properties". *Economics Letters* 62, 161-165.

[12] Hamilton, J.D. (1989). "A new Approach to the Economic Analysis of Non stationarity Times Series and the Business Cycle," Econometrica 57, 357-384.

[13] Hosking, J.R.M. (1981). "Fractional differencing". *Biometrika*, 68, 165-176.

[14] Hsu, C.C and Kuan C.M., (2000). "Long memory or structural Change: Testing method and empirical examination".

[15] Jensen, J. (1999). "Using wavelets to obtain a consistent ordinary least squares estimator of long memory parameter". Journal of Forecasting 18, 17-32.

[16] Jensen, J. and Whitcher, B. (2000). "Time-Varying Long-Memory in Volatility: Detection and Estimation with Wavelets". Technical Report, Department of Economics, University of Missouri.

[17] Kato, T., Masry, E., (1999). "On the spectral density of the wavelet transform of fractional Brownian motion". *Journal of Time Series Analysis* 20, 559-563.

[18] Lim, K.S. and Tong, H. (1980). "Threshold autoregressions, limit cycles, and data," *Journal of the Royal Statistical Sociaty*, B 42, 245-92 (with discussion).

[19] Lee, J., (2005). "Estimating memory parameter in the US in_ation rate". *Economics Letters* 87, 207-210.

[20] Lee, J. and Robinson, P.M., (1996). "Semiparametric Exploration of Long Memory in Stock Prices". *Journal of Statistical Planning and inference* 50,155-174.

[21] Lobato I.N., Savin N.E. (1998). "Real and spurious long memory properties of stock market data". *Journal of Business and Economic Statistics* 16: 261-268.

[22] P. 2003. "Computation and analysis of multiple structural change models". *Journal of Applied Econometrics* 18: 1-22.

[23] Robinson P.M. (1995) "Log-periodogram regression of time series with long range dependence". *The annals of Statistics*, 23, (1995), 1048-1072.

[24] 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.

[25] Shimotsu, K. and Phillips C.B.P. (2005) "Exact Local Whittle Estimation of Fractional Integration", *Annals of Statistics*, 33(4), pp. 1890-1933.

[26] Yajima, Y. (1989). "A central limit theorem of Fourier transforms of strongly dependent stationary processes". *Journal of Time series analysis*, 10, 375-383.

Figure 5: Trajectories of the absolute stock return for the BNP, LVMH, PPR, Saint Gobain, Société Générale, Suez and TF1 companies and the CAC 40 stock market index.

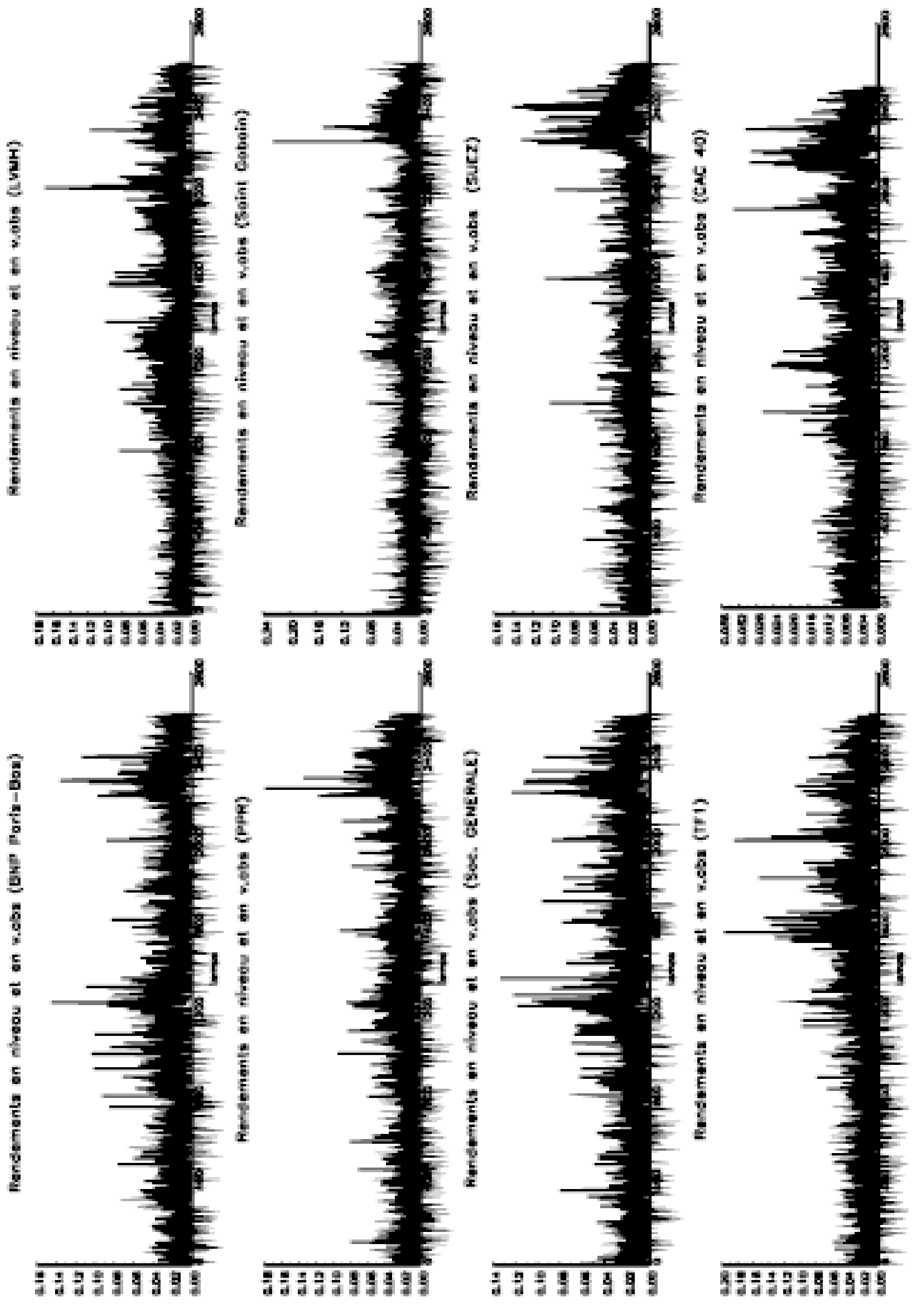


Figure 6: Autocorrelation functions of the level (dash line) and the absolute (solid line) returns for the BNP, LVMH, PPR, Saint Gobain, Soc. Générale, Suez and TF1 companies and the CAC 40 stock market index.