A SIMPLE FRACTIONALLY INTEGRATED MODEL WITH A TIME-VARYING LONG MEMORY PARAMETER Dt - [Document de travail n°2008 - 10]

Mohamed Boutahar, Gilles Dufrénot, Anne Peguin-Feissolle

GREQAM

Groupement de Recherche en Economie
Quantitative d'Aix-Marseille - UMR-CNRS 6579
Ecole des Hautes Etudes en Sciences Sociales
Universités d'Aix-Marseille II et III

Document de Travail
n°2008-10

A SIMPLE FRACTIONALLY INTEGRATED MODEL WITH A TIME-VARYING LONG MEMORY PARAMETER d_t

Mohamed BOUTAHAR
Gilles DUFRENOT
Anne PEGUIN-FEISSOLLE

Avril 2008

DT-GREQAM

A simple fractionally integrated model with a time-varying long memory parameter d_t [*†]

Mohamed Boutahar[‡] **Gilles Dufrénot**[§]
Anne Péguin-Feissolle[¶]

August 28, 2007

Abstract

This paper generalizes the standard long memory modeling by assuming that the long memory parameter d is stochastic and time-varying: we introduce a STAR process on this parameter characterized by a logistic function. We propose an estimation method of this model. Some simulation experiments are conducted. The empirical results suggest that this new model offers an interesting alternative competing framework to describe the persistent dynamics in modeling some financial series.

Keywords: Long-memory - Logistic function - STAR

JEL classification: C32, C51, G12

[*] *Corresponding author:* Anne PEGUIN-FEISSOLLE, GREQAM, Centre de la Charité, 2 rue de la Charite, 13236 Marseille cedex 02, FRANCE, tel: +33.4.91.14.07.35, fax: +33.4.91.90.02.27, Email: Anne.Peguin@univmed.fr

[†] Helpful comments and suggestions from a referee are gratefully acknowledged.

[‡] *GREQAM, Université de la Méditerranée, France*

[§] *ERUDITE, Université Paris 12 and GREQAM, France*

[¶] *GREQAM-CNRS, France*

1 Introduction

Long memory models have received considerable attention by researchers from various disciplines including hydrology, finance and other fields. The books by Beran (1994), Doukhan *et al.* (2003) and Robinson (2003) provide updated surveys of recent developments on this topic. The most studied long-memory process is the $ARFIMA(p,d,q)$ defined by the equation

$$\Phi(L)\,(1-L)^d\,X_t = \Theta(L)\varepsilon_t$$

where ε_t is a zero mean white noise, L is the backward shift operator, $\Phi(L)$ and $\Theta(L)$ are polynomials in L of degree p and q respectively. Under suitable conditions of the polynomial functions, the autocovariance sequence $\gamma(k)$ of the $ARFIMA(p,d,q)$ process X_t decays as k^{2d-1} and the partial sum $N^{-d-1/2}\sum_{t=1}^{N} X_t$ converges in distribution to a fractional Brownian motion.

In modeling applications, treating the long memory parameter d as a constant implies that the long range dependence structure of the underlying phenomenon persists over time. However, many phenomena exhibit a long range dependence that changes as the phenomenon itself evolves. We then need a class of processes which is based on the autocovariance function and that relates the *local* autocorrelation of the process to the values of d.

Various models with a time-varying long memory parameter have been proposed in the literature. For instance, Ray and Tsay (2002) propose an $ARFIMA$ process where the parameter d_t evolves as a random walk

$$d_t = d_{t-1} + \delta_t\beta_t, \tag{1}$$

where β_t is an i.i.d. Bernoulli sequence. They estimate d_t by using a Bayesian approach. Another paper is Gil-Alana (2004)'s where the author considers the model

$$(I-L)^{d_t}X_t = \varepsilon_t, \tag{2}$$

and tests the null hypothesis $d_t = d_0$ against the alternative that d_t evolves as $d_t = d_0 + \delta T^{1/2}$ where T is the sample size; in his paper, the fractional parameter is fixed over time, but the procedure implemented is efficient in the Pitman sense against local departures from the null. In another paper, Beine and Laurent (2001) study structural changes and long memory in the conditional variance of the exchange rate returns through a Markov-switching FIGARCH (Fractionally Integrated Generalized Autoregressive Conditional Heteroscedastic) model. They assume that the returns $(X_t)_t$ are Gaussian with a conditional variance depending on time t:

$$\sigma_t^2 = \omega_{s_t} + \beta\sigma_{t-1}^2 + (I-L)^{d_{s_t}}\epsilon_t^2, \tag{3}$$

where s_t is a Markov chain that captures different volatility regimes. Other papers relate fractional integration and structural breaks, for instance Beran and Terrin (1996), Lobato and Savin (1998), Engle and Smith (1999), Bos et al. (1999), Hyung and Franses (2006), Gil-Alana (2003).

Dufrénot et al. (2005a, 2005b and 2006) assume that d_t is a SETAR (Self-Exciting Threshold Autoregressive) process with two regimes, that is

$$(I-L)^{d_t} X_t = \varepsilon_t \quad \text{with } d_t = \begin{cases} d_{(1)} & \text{if } X_{t-1} \leq c \\ d_{(2)} & \text{otherwise} \end{cases}. \tag{4}$$

In this paper, we extend the process (4) by considering a smooth function for the long memory parameter d_t. The basic idea is to account for possible structural changes in the memory of the data. Our motivations for using such a model are twice. Firstly, as shown by Diebold and Inoue (2001) and Granger and Hyung (2004), long-memory and structural change can be easily confused so long as a small amount of regime switching occurs. One thus needs a model that allows to distinguish between the two properties in a same framework. Secondly, from an empirical aspect, volatility may exhibit different correlation patterns when it is large and when it is low. One can thus assume that the degree of persistence of the volatility is regime-changing.

The paper is organized as follows. Section 2 gives some definitions of locally stationary long memory process. This enables us to introduce the model in section 3. Section 4 proposes an estimation method for such a model and some simulations in order to examine the power of the proposed method. Section 5 contains some applications to financial time series. Section 6 concludes the paper.

2 Locally stationary long memory processes

Because the process we will consider below is characterized by a time-varying long memory parameter d_t, we give some definitions and theoretical properties of this class of processes (for details, see Ayache et al. (2000)). Let (X_t) be a second order process and $\rho_t(.)$ its local correlation i.e.

$$\rho_t(k) = \frac{cov(X_t, X_{t+k})}{\sqrt{var(X_t) var(X_{t+k})}} \tag{5}$$

with

$$cov(X_t, X_{t+k}) = E\left[\left(X_t - E\left(X_t\right)\right)\left(X_{t+k} - E\left(X_{t+k}\right)\right)\right].$$

Definition 1 *(Time domain). (X_t) is a locally stationary long memory process if there exists a function $\alpha(.)$ taking values in $]-1,0[$ such that*

$$\rho_t(k) \sim C(t) k^{\alpha(t)} \text{ as } k \to \infty. \tag{6}$$

where $C(t)$ is a non negative function of t.

Definition 2 *(Frequency domain). (X_t) is a locally stationary long memory process if there exists a function $\beta(t)$ taking values in $]-1,1[$ and a function $K(t)$ such that*

$$f_t(\lambda) \sim K(t)\lambda^{-\beta(t)} \text{ as } \lambda \to 0, \tag{7}$$

where $f_t(\lambda)$ is the evolutionary spectral density of (X_t).

Some processes satisfying (6) or (7) have been introduced in the literature. For example the increment $X_t = Y_t - Y_{t-1}$ of the multifractional Brownian motion given by

$$Y_t = \int_{\mathbb{R}} \frac{e^{it\lambda} - 1}{|\lambda|^{d_t+1}} dW(\lambda), \tag{8}$$

where $d_t \in]-0.5, 0.5[$ and W is the Wiener measure, is a deterministic function. The process (X_t) is also called the generalized fractional Gaussian noise (GFGN); it satisfies (6) with

$$\alpha(t) = \max(d_t, d_{t+1}) - 0.5 \tag{9}$$

under the condition

$$\max(d_t + d_s, d_{t+1} + d_s, d_t + d_{s+1}, d_{t+1} + d_{s+1}) > 0.5$$

(see, among others, Mandelbrot and Van Ness (1968), Mandelbrot and Wallis (1968, 1969), Marinucci and Robinson (1999), Benassi et al. (1998), Kou and Xie (2004), Ayache et al. (2000)).

Another locally stationary long memory process is the $ARFIMA$ given by

$$(1-L)^{d_t}\Phi(L)X_t = \Theta(L)\varepsilon_t, \tag{10}$$

with

$$\begin{cases} \Phi(L) = 1 + \phi_1 L + ... + \phi_p L^p \\ \Theta(L) = 1 + \theta_1 L + ... + \theta_q L^q \end{cases};$$

L is the backward shift operator defined by $L^j X_t = X_{t-j}$, $\Phi(L)$ and $\Theta(L)$ are stable polynomials, i.e. their roots are strictly outside the unit circle and $d_t \in]-0.5, 0.5[\subset \mathbb{R}$.

Let $R(t,s) := cov(X_t, X_s)$ the Kernel covariance of (X_t). The Wigner spectrum of (X_t) is given by

$$f_t(\lambda) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} R(t - \tau/2, t + \tau/2)e^{-i\lambda\tau} d\tau. \tag{11}$$

If we interpret $R(t-\tau/2, t+\tau/2)$ as a form of local autocovariance function in the neighborhood of t, then f_t is the evolutionary spectral density obtained by the Fourier transform of $R(t-\tau/2, t+\tau/2)$ with respect to τ (see Priestley (1988)).

It can be shown (see Wang *et al.* (2001)) that the evolutionary spectral density of both GFGN and $ARFIMA$ can be written as

$$f_t(\lambda) = K(t)\lambda^{-2d_t}\left\{1 + o(1)\right\} \quad \text{as } \lambda \to 0$$

and thus satisfies (7) with $\beta(t) = 2d_t$; hence, if $d_t > 0$, then f_t is unbounded at $\lambda = 0$ which is indicative of long memory.

3 The model

In this paper, we consider the $ARFIMA$ process (10), and assume that d_t evolves according to a stochastic process. We extend the papers by Dufrénot *et al.* (2005a, 2005b and 2006) by assuming that d_t is a STAR process

$$d_t = d_{(1)}(1 - F_t(\gamma, c)) + d_{(2)}F_t(\gamma, c), \tag{12}$$

where $F_t(\gamma, c)$ is a smooth transition function, for instance a logistic function given by:

$$F_t(\gamma, c) = \left[1 + \exp\left(-\gamma\left(X_{t-1} - c\right)\right)\right]^{-1}. \tag{13}$$

The parameters $d_{(1)}$ and $d_{(2)}$ are the values of the long memory parameter in the two extreme regimes, that is when $F_t \to 0$ and $F_t \to 1$. The slope parameter γ indicates how rapidly the transition between two extreme regimes is; γ can be either positive or negative, depending upon whether the logistic curve is increasing or not. The parameter c is the location parameter.

If $\gamma > 0$, then the extreme regimes are obtained as follows:

- if $X_{t-1} \to -\infty$, we have $F_t \to 0$ and thus $(1-L)^{d_{(1)}}\Phi(L)Y_t = \Theta(L)\varepsilon_t$,
- if $X_{t-1} \to +\infty$, we have $F_t \to 1$ and thus $(1-L)^{d_{(2)}}\Phi(L)Y_t = \Theta(L)\varepsilon_t$.

We assume that $\Phi(L) = \Theta(L) = 1$: we consider the simplest $ARFIMA(0, d_t, 0)$, that is an $I(d)$ with a time-varying fractional parameter. Let (Ω, F, P) be the probability space, and assume that X_t and d_t are defined on Ω and given respectively by (10) and (12)-(13). The stochastic fractional filter $(1-L)^{d_t}$ can be defined as follows:

$$\forall\omega \in \Omega, (1-L)^{d(t,\omega)} = \sum_{j=0}^{\infty} \frac{\Gamma(j - d(t,\omega))}{\Gamma(j+1)\Gamma(-d(t,\omega))} L^j, \tag{14}$$

where $d(t,\omega) := d_t(\omega)$ and Γ is the gamma function. Therefore, the process (X_t) can be written as

$$X_t = \sum_{j=0}^{\infty} b_j(X_{t-1})\varepsilon_{t-j}, \tag{15}$$

where

$$b_j(x) = \frac{\Gamma(j + D(x))}{\Gamma(j+1)\Gamma(D(x))} \tag{16}$$

and

$$D(x) = d_{(1)} + (d_{(2)} - d_{(1)})\left[1 + \exp\left(-\gamma\left(x - c\right)\right)\right]^{-1}. \tag{17}$$

The above model is a random coefficient model, but instead of assuming that the driving process for d_t depends on some exogenous variables we suppose that the long-memory parameter evolves according to the lagged endogenous variable. Thus, the model is a doubly stochastic model. Such models have received a widespread attention in the literature (the interested reader may refer to Tjostheim (1986) for univariate series). Their estimation usually implies procedures that are burdensome.

Our approach differs from other recently proposed papers in the literature where the fractional parameter interact with nonlinear functions. For instance, Caporale and Gil-Alana (2006) propose the following formulation: $(1-L)^d X_t = (1-L)^d g(\theta, z_t) + u_t$, where g is a nonlinear function of the exogenous variables z_t; other papers (van Dijk et al. (2002), Smallwood (2005), Boutahar et al. (2007) and Dufrénot et al. (2008), among others) use also fractional integration and non-linear STAR models.

4 Estimation of the long memory parameter d_t and simulation experiments

In this Section, we propose a simple method to estimate the parameters of the model (10) and (12)-(13), for $t = 1,...,T$. We want to estimate the long memory parameters in the two extreme regimes $d_{(1)}$ and $d_{(2)}$ and the transition parameter γ, assuming for purpose of simplicity that the location parameter $c = 0$ and that $\Phi(L) = \Theta(L) = 1$. We thus consider the model given by:

$$\begin{cases} (1-L)^{d_t} X_t = \varepsilon_t \\ d_t = d_{(1)} + \left(d_{(2)} - d_{(1)}\right)\left[1 + \exp\left(-\gamma X_{t-1}\right)\right]^{-1} \end{cases} \tag{18}$$

for $t = 1, ..., T$.

4.1 The grounds of the proposed method

The proposed estimation method makes use of an approach from the applied literature on nonlinear models, namely "arranged regressions". An arranged regression is a regression with the endogenous and exogenous variables, say Y and X, re-ordered based on the values of a particular variable, say Z. Such regressions are widely used in physics, biometrics, medicine, psychiatry, etc. Some examples are the hierarchically arranged regressions for neural network models (see Hoya (2004)), or arranged regressions in a context of classified variables (an illustration is provided by Gaunt and Gaus (1992)). In the field of economics, such regressions have been successfully applied to estimate the parameters of nonlinear models. A seminal paper is Tsay (1989)'s, who uses arranged regressions to estimate TAR (Threshold Autoregressive) models. More specifically, when the particular variable Z is a threshold variable, the observations of Y and X are arranged according to the ascending or descending order of magnitude of the observations of Z. Tsay's approach has been generalized to other nonlinear models with an asymmetric dynamics (for instance, functional-coefficient autoregressive models as in Chen and Tsay (1993), multivariate TAR models as in Tsay (1998), nonlinear conditionally heteroscedastic models with changing parameters as in Amendola and Storti (2002), or nonlinear VAR models as in Hecq (2007)). In all these models, recursive regressions are employed on successively larger numbers of sorted observations.

Arranged regressions based on ordered variables allow more power than standard estimators, in discerning more accurately threshold or changing regimes in a series. In our case, this methodology is applied to find the values of $d_{(1)}$ and $d_{(2)}$ since the changes of d_t between its two extreme values is subject to changes according to X_{t-1}. The logistic function is used to mean that the changes in the long memory parameter occur in a continuous manner. We consider both ascending and decreasing order of X_{t-1} for purpose of robustness.

4.2 An estimation method

To expose the intuition of the estimation procedure, we begin with a simulation experiment. We generate a series of $T = 5000$ observations according to (18) with $d_{(1)} = 0.1$, $d_{(2)} = 0.4$, and $\gamma = 5$.

Figure 1 shows the graph of the arranged observations of d_t based on the increasing values of X_{t-1}. The shape of the curve is obviously a logistic shape, following (18). It is characterized by a lower regime with $d_t \approx d_{(1)}$, an upper regime with $d_t \approx d_{(2)}$, and an intermediary regime with a continuum of points between the two extreme regimes. The time breaks T_1 and T_2 are approximately the breaks that separate the three regimes (upper, inner and lower). The estimation method consists in finding first these two time breaks, from which we will deduce the estimation of $d_{(1)}$ and $d_{(2)}$, and then the estimation of γ.

In order to find the time breaks T_1 and T_2, we consider, in step 1 below, a curve as represented in figure 1 to determine a first approximation of these breaks, that we call T_1^A and T_2^A. In step 2, we consider the same curve, but in the reverse order, to determine a second approximation, called T_1^B and T_2^B. In step 3, we determine the final value of T_1 and T_2 by averaging the preceding approximations. This procedure and the computation of the averages for determining the two time breaks correspond to steps 1-3 described below. They permit to solve the problem of "*braking*" in the determination of the breaks; such a problem occurs because any of these first two steps taken alone can yield to values of the breaks that are above or under the correct values of T_1 and T_2; using the double procedure and computing the averages solve this problem. Then, we are able to estimate $d_{(1)}$, $d_{(2)}$, and γ.

More precisely, the different steps of the estimation method are as follows.

Step 1: First approximation of T_1 and T_2: T_1^A and T_2^A

Stage 1a. Consider the time series $(X_t)_t$ with T observations. We first construct the time series $(\widetilde{X}_t)_t$ of arranged observations according to the increasing values of X_{t-1}.

Stage 1b. We take the first n^* observations[1] of the vector $(\widetilde{X}_t)_t$ and we estimate the long memory parameter d. We note d_{s_1} and t_{s_1} respectively the estimator of d and the corresponding $t-ratio$, based on the sub-sample $(\widetilde{X}_t)_{1\leq t\leq n^*}$. We repeat this procedure by incrementing this vector with one observation, up to the whole sample. At each incrementation, the long-memory parameter and its $t-ratio$ are computed. At the end, we obtain a sequence of estimators of d, $\{d_{s_1}, d_{s_2}, d_{s_3}, ..., d_{s_n}\}$, and a sequence of corresponding $t-ratios$ $\{t_{s_1}, t_{s_2}, t_{s_3}, ..., t_{s_n}\}$.

Stage 1c. In order to detect two breaks in the sequence of the estimators of d (or of the $t-ratios$), we use the Bai and Perron (1998 and 2003) method[2] that

[1] n^* is chosen approximately equal to $T/5$ in the empirical illustrations. We could choose any (T/K) observations where K is an integer. However, the choice of K is subject to some constraints. If it is too small, we may miss the first break. If it is too high, we would not have enough initial points; $K = 5$ seems to be a good compromise.

[2] We use Bai and Perron's GAUSS program (avalaible on the web page of the Journal of Applied Econometrics) to estimate the break dates.

gives an efficient algorithm to estimate the break dates as global minimizers of the sum of squared residuals. Of considerable interest is the fact that this method allows very general conditions on the data and the errors. We thus consider the regression where the explanatory variables are a constant and a linear time trend; for instance, for the sequence d_{s_t} of the estimators of d, we consider the model:

$$d_{s_t} = \alpha + \beta t + u_t, \quad t = 1, 2, \ldots, n, \tag{19}$$

where u_t is an error term that is allowed to contain autocorrelated and heteroscedastic patterns. We thus determine the two time breaks denoted by T_1^A and T_2^A.

Step 2: Second approximation of T_1 and T_2: T_1^B and T_2^B

We construct now the time series $(\widetilde{X}_t)_t$ of arranged observations according to the decreasing values of X_{t-1}. We follow the same stages as in step 1 to obtain the breaks T_1^B and T_2^B.

Step 3: Final approximation of T_1 and T_2 and estimation of $d_{(1)}$ and $d_{(2)}$

Stage 3a. We define $T_1 = (T_1^A + T_1^B)/2$ and $T_2 = (T_2^A + T_2^B)/2$. The approach used here is very similar to a grid search approach. Therefore, we cannot expect to find the exact values of the breaks, but only some approximations of T_1 and T_2 . The two breaks are invariant to the order used to arrange the observations (increasing or decreasing values of X_{t-1}). The estimation of the breaks with the vector d_{s_t} arranged according to the increasing values of X_{t-1} leads to over-estimate the true – but unknown – breaks. Similarly, when the vector is arranged based on the decreasing values of X_{t-1}, we under-estimate the two breaks. Taking the average of T_1^A, T_1^B and T_2^A, T_2^B yields a value of the breaks near the true values.

Stage 3b. We construct again the time series $(\widetilde{X}_t)_t$ of arranged observations according to the increasing values of X_{t-1}.

Stage 3c. Considering the vector of the first T_1 elements of $(\widetilde{X}_t)_t$, we estimate the long memory parameter and we call the estimator $\widehat{d}_{(1)}$. Then, we obtain the estimator of the long memory parameter, that we call $\widehat{d}_{(2)}$, corresponding to the vector of the last T_2 elements of $(\widetilde{X}_t)_t$.

Step 4: Estimation of γ

We determine γ by a linear interpolation. More precisely, consider again the time series $(\widetilde{X}_t)_t$ of arranged observations according to the increasing values of X_{t-1}. Let us define $x_1 = \widetilde{X}_{T_1}$ and $x_2 = \widetilde{X}_{T_2}$; we approximate the function

$$d(x) = \widehat{d}_{(1)} + \left(\widehat{d}_{(2)} - \widehat{d}_{(1)}\right) [1 + \exp(-\gamma x)]^{-1} \tag{20}$$

in the interval (x_1,x_2) by a straight line given by $a + bx$ and passing through the points $\left(x_1, \widehat{d}_{(1)}\right)$ and $\left(x_2, \widehat{d}_{(2)}\right)$; the slope of this line is $b = \frac{\widehat{d}_{(2)} - \widehat{d}_{(1)}}{x_2 - x_1}$. We now linearize $d(x)$ in (20) by expanding the function into a first-order Taylor series around 0:

$$d(x) \simeq d(0) + x d'(0),$$

which gives:

$$d(x) \simeq \frac{\widehat{d}_{(1)} + \widehat{d}_{(2)}}{2} + x \frac{\left(\widehat{d}_{(2)} - \widehat{d}_{(1)}\right)\gamma}{4}.$$

Therefore, the approximation of γ is determined by equalizing the slopes b and $\frac{\left(\widehat{d}_{(2)} - \widehat{d}_{(1)}\right)\gamma}{4}$, that implies that

$$\widehat{\gamma} = \frac{4b}{\widehat{d}_{(2)} - \widehat{d}_{(1)}} = \frac{4}{\widetilde{X}_{T_2} - \widetilde{X}_{T_1}}. \tag{21}$$

Concerning the estimation method of the long memory parameters at the different steps, we use the feasible exact local Whittle (FELW) estimator developed by Shimotsu (2006)[3]. It is an extended version of the exact local Whittle (ELW) estimator proposed by Shimotsu and Phillips (2004, 2005 and 2006), that is a semiparametric estimator giving generally a good estimation method for the memory parameter in terms of consistency and limit distribution, except in the case where the mean is unknown. To overcome this difficulty, Shimotsu (2006) extended the ELW estimator to the FELW estimator and shows that this estimator is consistent and has a $N\left(0, \frac{1}{4}\right)$ limit distribution for $d \in \left(-\frac{1}{2}, 2\right)$.

4.3 Some simulation results

In order to assess the accuracy of the estimation method, we use artificial time series. In the simulation experiments, we focus on the model (18) by

[3]We use the code available from K. Shimotsu at
http://qed.econ.queensu.ca/faculty/shimotsu/programs/elwcode.zip

generating a sample of $T = 5000$ observations. The long memory generation is based on a polynomial hyperbolic decay with $N = 500$ lags; indeed, Bhardwaj and Swanson (2006, table 1) show that, after this lag, for most of values of the long memory parameter, the coefficients of the polynomial become smaller than 1.0e-004. $n^* = 1000$ is the initial number of observations used to generate the vector of the estimators of d or of the $t-ratios$ in the estimation of $d_{(1)}$ and $d_{(2)}$. Moreover, the initial value of the series $(X_t)_t$ is $X_0 = 0$ and more observations than needed are generated in order to eliminate the influence of this starting value. As described above, the FELW method is used to estimate the long memory parameters in the different steps, and the Bai and Perron method to determine the time breaks on the sequence of estimators of d and the sequence of corresponding $t-ratios$.

In Table 1, we report the results obtained when the regression in stage (1c) is applied to the sequence $\{d_{s_1}, d_{s_2}, ..., d_{s_n}\}$, and Table 2 contains the results based on the sequence of $t-ratios$ $\{t_{s_1}, t_{s_2}, ..., t_{s_n}\}$. T is the number of observations, N is the number of lags in the polynomial hyperbolic decay used to generate the long memory series; n^* is the initial number of observations. $d_{(1)}$ and $d_{(2)}$ are the "true" values used to generate the artificial series and $\widehat{d}_{(1)}$ and $\widehat{d}_{(2)}$ are their corresponding estimators. T_1 and T_2 are the time break values corresponding to the computation of $\widehat{d}_{(1)}$ and $\widehat{d}_{(2)}$. Finally, $s.e.$ are the estimated standard-errors of the Shimotsu estimator; Inf and Sup are the lower and upper bounds of the confidence intervals, respectively $\widehat{d} - 1.96s.e.$ and $\widehat{d}+1.96s.e.$ For the feasible local Whittle estimation (see Shimotsu (2006), m is chosen to be $m = T^{0.65}$ with T is the sample size; the lower and upper bounds on the long memory parameter are respectively -0.2 and 1.2.

In both tables, the true values of $d_{(1)}$ and $d_{(2)}$ are respectively equal to 0.15 and 0.35, and γ varies from 0.5 to 50. Though the true values of $d_{(1)}$ are always in the 95% confidence interval, $\widehat{d}_{(1)}$ becomes smaller when γ varies from small to high values. The more abrupt the transition from a regime to another, the smaller the value of the long memory parameter in the first regime. The simulations for the parameter $d_{(2)}$ show results that are more "unstable" in the sense that for some values of the parameter γ (when γ lies between 0.5 and 10.0), it is possible to find a 95% confidence interval that does not contain the true value $d_{(2)}$. In our example, $d_{(2)}$ is not in the interval when γ equals 0.5 and 10.0, while it lies in the 95% confidence interval when γ equals 5.0. However, when the slope parameter increases and takes high values, the risk of finding biased values of $d_{(2)}$ diminishes slightly, since $d_{(2)}$ is always in the confidence interval.

5 An application to financial time series

In this section, we apply the model to exchange rate series and to stock market indices. In the applications, we take n^* not different from $T/5$, where T is the total number of observations and n^* is the initial number of observations used to detect the structural changes in the sequence of estimators of d (we do not consider here the sequence of corresponding $t - ratios$); but slightly different values of n^* give approximately the same estimation results. We assume that the location parameter $c = 0$.

The first type of series we consider is the centered absolute growth rates of the following foreign currency exchange rate data: $AUST$ (Austria), UK (United Kingdom), CAN (Canada), JAP (Japan) and $SWIS$ (Switzerland), all expressed as number of units of the foreign currency per US dollar. The data are daily: we take 4000 observations up to December 31 1998. The series of interest (X_t) is defined by

$$X_t = |R_t| - \overline{|R_t|}$$

with

$$R_t = \ln(S_t) - \ln(S_{t-1})$$

and S_t is the nominal exchange rate.

We also consider the centered absolute returns (X_t) of different daily stocks returns up to 01/18/2002; the total number of observations is $T = 4500$. These data were used by Bhardwaj and Swanson (2006)[4], who applied $ARFIMA$ models to the series and found that the latter frequently outperform linear models in terms of prediction. We study the series (X_t) defined by

$$X_t = |R_t| - \overline{|R_t|}$$

where R_t is a stock return and $\overline{|R_t|}$ is the mean over the whole sample.

We use the absolute returns as a proxy of volatility, instead of the squared returns, as the latter measure can be noisy (see Andersen and Bollerslev (1998)). As evidenced in the literature, the absolute returns can show persistence reflected, either by a standard ARMA model, or by a long-memory process if the persistence lasts for long time intervals. Of course, a combination of both components may also account for the persistent dynamics. Here, we employ our simple model, with no ARMA components, for two reasons. Firstly, with the inclusion of ARMA components, our procedure becomes very time consuming. Indeed, the values of T_1, T_2, $\widehat{d}_{(1)}$ and $\widehat{d}_{(2)}$ vary with the values of the coefficients of the polynomials $\Phi(L)$ and $\Theta(L)$ in equation (10).

[4] The data comes from http://econweb.rutgers.edu/bhardwaj/ARFIMA/mc_4000.

So, we need a "double" grid search approach since the procedure described in section 4 needs to be considered for each potential values of $\Phi(L)$ and $\Theta(L)$. Secondly, a joint maximum likelihood estimation of all the parameters renders the Whittle estimator very complex and difficult to apply. The important point for us is to detect a dynamic with a "memory" structure that is dynamically changing in time.

The estimation results, for the exchange rate series, are shown in table 3. The whole sample shows a long memory behavior (all the estimated parameters $\widehat{d}$ are significant). The estimated parameters $\widehat{d}_{(1)}$of the lower regime are negative (except for Austria) and non significant, thereby indicating that the exchange rate volatility is characterized by a short memory process. On the opposite, the upper regime is characterized the presence of a long memory process where all the $\widehat{d}_{(2)}$ are included in the interval $[0.10, 0.21]$.

Table 4 reports the estimates for the centered absolute returns of the German Stock index (DAX), the Hong Kong Stock index (HSG), the Nikkei Stock Index (NKK) and the Standard & Poor's Index ($SP500$). We do not reject a long memory process on the whole sample for all the series as shown by the 95% confidence intervals. Moreover, the model characterized by a time-varying modelling of the long memory parameter d_t using a STAR process shows an upper regime characterized a long memory behavior with significant parameters $\widehat{d}_{(2)}$ between 0.14 and 0.38, and a lower regime characterized by a non significant long memory parameter $\widehat{d}_{(1)}$.

The interesting feature of the model is that it allows to capture the long-memory dynamics in the volatility with heterogeneous persistence. In the upper regime, when the volatility evolves above its "normal" level (defined as the mean value), the fluctuations of the exchange rate series and the stock indices react much slower in adjusting to shocks; in the other regime, the behavior closely approximates a short-memory process. Different degrees of persistence can be viewed as a manifestation of the two facets of volatility: a low volatility clustering and a high volatility clustering. The lower regime is sometimes characterized by anti-persistence, that is a negative short-term dependence that can be related with a sub-diffusive stochastic drift in volatility: if volatility increases today, it is less likely to do so tomorrow. Anti-persistence may arise for instance because the agents' demand is bounded by their limited resources. On the opposite, the finding of a long-memory behavior in the volatility in the upper regime corroborates similar findings in the literature: pockets of predictability are usually found to exist when volatility is high.

Our findings are in line with some results previously obtained in the literature concerning the presence of long memory in volatility with a multi-scaling feature. The changing patterns of the memory can be related to different theoretical hypotheses (for instance, the fractal market hypothesis of Peters (1994) or the interacting agent hypothesis of Lux and Marchesi (1999)). Our results also corroborate the "mixture of distribution hypothesis" of Andersen and Bollerslev (1997). Here, the global long memory behaviour of the volatility series are the "sum" of short-memory and long memory processes, the presence of which may be explained by the heterogeneity of agents with different time horizons with participants having a large spectrum of dealing frequency. For instance, on one side we may have speculators with a short time horizon and on the other side institutions like central banks or pension funds investors with a longer time horizon.

The estimates of the coefficients γ show that the volatility of the exchange rate series is produced by a long-memory process with suddenly changing parameter. Indeed, the estimates are slightly greater for these series than for the stock indices.

6 Conclusions and extensions

In this paper, we have proposed an approach to model, jointly, the long-memory and changing regime in a time series. This is done by allowing the long-memory parameter in an $ARFIMA$ model to vary across time according to a stochastic process described by a STAR transition function. The model can be extended in several ways.

Firstly, the model considered in this paper could be extended by including ARMA components. In this case, the estimation of the fractional differencing parameters and the ARMA coefficients could be done using maximum likelihood approaches (for instance, Sowell (1992)), though this would be much more expensive computationally. Secondly, the long-range dependence behavior can be modelled using a Generalized $ARFIMA$ model. This allows to take into account the presence of pseudo-cycles in the data. For instance, this model could be suitable to capture heterogeneous degrees of persistence in macroeconomic time series. Thirdly, on can allow the possibility that the noise follows an heteroscedastic process, such as a $GARCH(1,1)$. In this case, a general formulation of the model is the following:

$$(I - 2uL + L^2)^{d_t} X_t = \epsilon_t \tag{22}$$

with

$$\epsilon_t \approx N(0, \sigma_t^2) \text{ and } \sigma_t^2 = a_0 + a_1\epsilon_{t-1}^2 + b\sigma_{t-1}^2. \tag{23}$$

If d_t evolves according to (12)-(13), we call this model a conditional heteroscedastic logistic Gegenbauer process. Note that this model could be extended to a k-factor Gegenbauer process. If $u = 1$, the model becomes a conditional heteroscedastic logistic long-memory process. If we further assumes that $\sigma_t^2 = a_0$, then we have a logistic long-memory process that corresponds to the model proposed in this paper. Such a model may be useful when applied to macroeconomic time series, which contain periodic components.

References

[1] Amendola, A. & Storti, G. (2002). A nonlinear time series approach to modelling asymmetry in stock market indexes. *Statistical Methods and Applications,* 11, 201-216.

[2] Andersen, T.G. & Bollerslev, T. (1997). Heterogenous information arrivals and return volatility dynamics: uncovering the long-run in high frequency data. *Journal of Finance,* 52, 975-1005.

[3] Andersen, T.G. & Bollerslev, T. (1998). Answering the skeptics: yes, standard volatility models do provide accurate forecasts. *International Economic Review,* 39, 885-905.

[4] Ayache, A., Cohen, S. & Lévy Véhel, J. (2000). The covariance structure of multifractional Brownian motion with application to long range dependence. *Proceedings of ICASSP* (IEEE International Conference on Acoustics, Speech, and Signal Processing), 6, 3810-3813.

[5] Bai, J. & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica,* 66, 47-78.

[6] Bai, J. & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics,* 18, 1-22.

[7] Beine, M. & Laurent, S. (2001) Structural changes and long memory in volatility: new evidence from daily exchange rates, in Dunis, C., A. Timmerman and J. Moody (eds). *Developments in Forecast Combination and Portfolio Choice, Wiley Series in Quantitative Analysis*, Wiley, chapter 6, 145-157.

[8] Benassi, A., Cohen, S. & Istas, J. (1998). Identifying the multifractional function of a Gaussian process. *Statistical Probability Letters,* 39, 337-345.

[9] Beran, J. (1994). *Statistics For Long-Memory Processes.* (Chapman & Hall).

[10] Beran, J. & Terrin, N. (1996). Testing for a change of the long memory parameter. *Biometrika,* 83, 627-638.

[11] Bhardwaj, G. & Swanson, N.R. (2006). An empirical investigation of the usefulness of ARFIMA models for predicting macroeconomic an financial time series. *Journal of Econometrics,* 131, 1/2, 539-578.

[12] Bos, C.S., Franses, P.H. & Ooms, M. (1999). Long-memory and level shifts: re-analyzing inflation rates. *Empirical Economics*, 24, 3, 427-449.

[13] Boutahar, M., Mootamri, I. & Péguin-Feissolle, A. (2007). An exponential FISTAR model applied to the US real effective exchange rate. Working Paper, GREQAM, France.

[14] Caporale, G.M. & Gil-Alana, L.A. (2006), Non-linearities and fractional integration in the US unemployment rate, *Oxford Bulletin of Economics and Statistics*, 69, 4, 521-544.

[15] Chen, R.C. & Tsay, R.S. (1993). Functional-Coefficient autoregressive models. *Journal of the American Economic Association,* 88, 298-308.

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

[17] Doukhan, P., Oppenheim, G. & Taqqu, M.S. (2003). *Theory and applications of long-range dependence.* (Birkhäuser: Basel).

[18] Dufrénot, G., Guégan, D. & Péguin-Feissolle, A. (2005a). Modelling squared returns using a SETAR model with long memory dynamics. *Economics Letters,* 86, 2, 237-243.

[19] Dufrénot, G., Guégan, D. & Péguin-Feissolle, A. (2005b). Long-memory dynamics in a SETAR model - Applications to stock markets. *Journal of International Financial Markets, Institutions & Money,* 15, 5, 391-406.

[20] Dufrénot, G., Guégan, D. & Péguin-Feissolle, A. (2006). Changing-regime volatility : A fractionally integrated SETAR model. *Applied Financial Economics.* (in press).

[21] Dufrénot, G., Lardic, S., Mathieu, L., Mignon, V. & Péguin-Feissolle, A. (2008). Explaining the European exchange rates deviations: long memory or nonlinear adjustment?. *Journal of International Financial Markets, Institutions & Money*, 18, 2 (in press).

[22] Engle R., & Smith, A.D. (1999). Stochastic Permanent Breaks. *The Review of Economics and Statistics.* 81, 4, 553-574.

[23] Gaunt, A.S. & Gaus, C.(1992). Serially arranged myofibers: an unappreciated variant in muscle architecture. *Cellular and Molecular Life Sciences,* 48, 864-868.

[24] Gil-Alana, L.A. (2003). Testing of unit roots and other fractionally integrated hypothesis in the presence of structural breaks. *Empirical Economics,* 28, 101-113.

[25] Gil-Alana, L.A. (2004). Testing of $I(d)$ processes in the real output. *Economics Bulletin,* 3, 32, 1-6.

[26] Granger, C.W.J. & 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.

[27] Hecq, A. (2007). Asymmetric business cycle co-movements. Working Paper, Department of Quantitative Economics, University of Maastricht.

[28] Hoya, T. (2004). Notions of intuition and attention modelled by a hierarchically arranged generalized regression neural network. *Systems, Man and Cybernetics, IEEE Transactions,* 34, 200-209.

[29] Hyung, N. & Franses, P.H. (2006). Structural breaks and long memory in US inflation rates. Do they matter for forecasting? *Research in International Business and Finance,* 20, 1, 95-110.

[30] Kou, S.C. & Xie, X.S. (2004). Generalized Langevin equations with Fractional Gaussian Noise. Subdiffusion within a single protein molecule. *Physical Review Letters,* 93, 18.

[31] Lobato, I. & Savin, N.E. (1998). Real and Spurious Long Memory Properties of Stock Market Data. *Journal of Business and Economic Statistics*, 16, 261-283.

[32] Lux, T. & Marchesi, M. (1999). Scaling and criticality in a stochastic multi-agent model of financial market. *Nature,* 397, 498-500.

[33] Mandelbrot, B. & Van Ness, J. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Review,* 10, 422-437.

[34] Mandelbrot, B. & Wallis,J. (1968). Noah, Joseph and Operational Hydrology. *Water Resources Research,* 4, 909-918.

[35] Mandelbrot, B. & Wallis,J. (1969). Computer experiments with fractional Gaussian noises. *Water Resources Research,* 5, 228-267.

[36] Marinucci, D. & . Robinson, P.M(1999). Alternative forms of fractional Brownian motion. *Journal of Statistical Planning Inference,* 80, 111-122.

[37] Peters, E. (1994). *Fractal market analysis.* John Wiley and Sons, New-York.

[38] Priestley, M.B. (1988). *Non linear and non stationary time series analysis.* (Academic Press: London).

[39] Ray, B.R. & Tsay, R.S. (2002). Bayesian methods for change-point detection in long range dependent processes. *Journal of Time Series Analysis,* 23, 687-705.

[40] Robinson, P.M. (2003). *Time series with long memory.* (Oxford University Press: Oxford).

[41] Shimotsu. K. (2006), Exact local Whittle estimation of fractional integration with unknown mean and time trend, Working Paper N. 1061, Queen's University.

[42] Shimotsu, K. & Phillips, P.C.B. (2004). Local Whittle estimation in nonstationary and unit root cases. *Annals of Statistics,* 32, 2, 656-692.

[43] Shimotsu, K. & Phillips, P.C.B. (2005). Exact local Whittle estimation of fractional integration. *Annals of Statistics,* 33, 4, 1890-1933.

[44] Shimotsu, K. & Phillips, P.C.B. (2006). Local Whittle estimation of fractional integration and some of its variants. *Journal of Econometrics,* 130, 209-233.

[45] Smallwood, A.D. (2005). Joint tests for non-linearity and long memory: the case of purchasing power parity. *Studies in Nonlinear Dynamics and Econometrics,* 9, 2, 1-28.

[46] Sowell, F. (1992). Maximum likelihood estimation of stationary univariate fractionally integrated time series models, *Journal of Econometrics*, 53, 165-188.

[47] Tjostheim, D. (1986). Some doubly stochastic time series models. *Journal of Time Series Analysis,* 7, 51-72.

[48] Tsay, R.S. (1989). Testing and modelling threshold autoregressive processes. *Journal of the American Statistical Association,* 84, 231-240.

[49] Tsay, R.S. (1998). Testing and modelling multivariate threshold models. *Journal of the American Statistical Association,* 92, 1188-1202.

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

[51] Wang. Y., Cavanaugh, J.E. & Song, C. (2001). Self-similarity index estimation via wavelets for locally self-similar processes. *Journal of Statistical Planning and Inference,* 99, 91-110.

Table 1. Simulation results
- breaks detected in the sequence of estimators of d
($T = 5000$, $N = 500$, $n^* = 1000$)

$d_{(1)}$	0.1500	0.1500	0.1500	0.1500	0.1500	0.1500
$d_{(2)}$	0.3500	0.3500	0.3500	0.3500	0.3500	0.3500
γ	0.5000	5.0000	10.0000	20.0000	30.0000	50.0000
T_1	1520	2195	2184	2519	2090	1830
$\widehat{d}_{(1)}$	0.1404	0.1277	0.1294	0.1233	0.1192	0.1021
$s.e.$	0.0462	0.0411	0.0411	0.0392	0.0418	0.0435
Inf	0.0498	0.0471	0.0488	0.0464	0.0373	0.0168
Sup	0.2309	0.2082	0.2099	0.2001	0.2012	0.1874
T_2	2938	3192	3477	3411	3390	3483
$\widehat{d}_{(2)}$	0.2269	0.2774	0.2342	0.2755	0.2744	0.2608
$s.e.$	0.0419	0.0438	0.0462	0.0456	0.0454	0.0464
Inf	0.1308	0.1915	0.1436	0.1861	0.1853	0.1698
Sup	0.3229	0.3632	0.3247	0.3648	0.3635	0.3518

Note: T is the number of observations, N is the number of lags in the polynomial hyperbolic decay used to generate the long memory series; n^* is the initial number of observations used to generate the vector of the estimators of the long memory parameters or the vector of the $t - ratios$ in the estimation of $d_{(1)}$ and $d_{(2)}$. $\widehat{d}$ is the long memory parameter estimator on the whole sample. $d_{(1)}$, $d_{(2)}$ and γ are the true generation values of the parameters; $\widehat{d}_{(1)}$and $\widehat{d}_{(2)}$ are the estimators of $d_{(1)}$and $d_{(2)}$. T_1 and T_2 are the time break values corresponding to the computation of $\widehat{d}_{(1)}$ and $\widehat{d}_{(2)}$. Finally, $s.e.$ are the estimated standard-errors of the Shimotsu estimator; Inf and Sup are the lower and upper bounds of the 95% confidence intervals. For the feasible local Whittle estimation (see Shimotsu (2006), m is chosen to be $m = T^{0.65}$ with T is the sample size; the lower and upper bounds on the long memory parameter are respectively -0.2 and 1.2. The time breaks are determined by Bai and Perron (2003) method.

Table 2. Simulation results
- breaks detected in the sequence of $t - ratios$
$(T = 5000,\ N = 500,\ n^* = 1000)$

$d_{(1)}$	0.1500	0.1500	0.1500	0.1500	0.1500	0.1500
$d_{(2)}$	0.3500	0.3500	0.3500	0.3500	0.3500	0.3500
γ	0.5000	5.0000	10.0000	20.0000	30.0000	50.0000
T_1	1506	2222	2186	2253	2110	2287
$\widehat{d}_{(1)}$	0.1323	0.1227	0.1315	0.1297	0.1178	0.1087
$s.e.$	0.0464	0.0409	0.0411	0.0406	0.0416	0.0405
Inf	0.0413	0.0425	0.0509	0.0499	0.0362	0.0293
Sup	0.2232	0.2028	0.2120	0.2094	0.1995	0.1882
T_2	2929	3161	3477	3134	3390	3485
$\widehat{d}_{(2)}$	0.2239	0.2784	0.2342	0.2952	0.2744	0.2561
$s.e.$	0.0418	0.0435	0.0462	0.0433	0.0454	0.0464
Inf	0.1298	0.1931	0.1436	0.2103	0.1853	0.1652
Sup	0.3179	0.3636	0.3247	0.3802	0.3635	0.3471

Note: see note of table 1.

Table 3. Estimation results for the exchange rates - breaks detected in the sequence of estimators of d ($T = 4000$ and $n^* = 800$)

	$AUST$	UK	CAN	JAP	$SWIS$
$\widehat{d}$	0.3016	0.3550	0.3207	0.2922	0.3202
$s.e.$	0.0337	0.0337	0.0337	0.0337	0.0337
Inf	0.2354	0.2888	0.2545	0.2260	0.2540
Sup	0.3679	0.4212	0.3869	0.3585	0.3865
T_1	1550	1797	1606	1692	1674
$\widehat{d}_{(1)}$	0.0073	-0.0024	-0.0551	-0.0378	-0.0850
$s.e.$	0.0460	0.0438	0.0454	0.0447	0.0449
Inf	-0.0828	-0.0884	-0.1442	-0.1254	-0.1730
Sup	0.0975	0.0834	0.0339	0.0498	0.0029
T_2	2750	3102	2647	2728	3080
$\widehat{d}_{(2)}$	0.1737	0.1067	0.2046	0.1256	0.2064
$s.e.$	0.0492	0.0548	0.0481	0.0490	0.0545
Inf	0.0771	-0.0007	0.1103	0.0295	0.0995
Sup	0.2702	0.2143	0.2989	0.2217	0.3133
$\widehat{\gamma}$	34.2913	22.3609	91.2237	33.4111	19.5888

Note: T is the number of observations and n^* is the initial number of observations used to generate the sequence of the estimators of the long memory parameter for the estimation of $d_{(1)}$ and $d_{(2)}$. $\widehat{d}$ is the long memory parameter estimator on the whole sample. $\widehat{d}_{(1)}$, $\widehat{d}_{(2)}$ and $\widehat{\gamma}$ are the estimators of $d_{(1)}$, $d_{(2)}$ and γ. T_1 and T_2 are the time break values corresponding to the computation of $\widehat{d}_{(1)}$ and $\widehat{d}_{(2)}$. Finally, $s.e.$ are the estimated standard-errors of the Shimotsu estimators; Inf and Sup are the lower and upper bounds of the 95% confidence intervals. For the feasible local Whittle estimation (see Shimotsu (2006), m is chosen to be $m = T^{0.65}$ with T is the sample size; the lower and upper bounds on the long memory parameter are respectively -0.2 and 1.2. The time breaks are determined by Bai and Perron (2003) method.

Table 4. Estimation results for the daily stocks indices - breaks detected in the sequence of estimators of d ($T = 4500$ and $n^* = 900$)

	DAX	HSG	NKK	$SP500$
$\widehat{d}$	0.4324	0.3638	0.4147	0.3376
$s.e.$	0.0325	0.0325	0.0325	0.0325
Inf	0.3686	0.3000	0.3509	0.2739
Sup	0.4962	0.4276	0.4785	0.4014
T_1	1411	1633	1973	1951
$\widehat{d}_{(1)}$	-0.0247	-0.1028	0.0428	0.0344
$s.e.$	0.0474	0.0452	0.0425	0.0427
Inf	-0.1177	-0.1915	-0.0405	-0.0493
Sup	0.0683	-0.0141	0.1262	0.1181
T_2	2963	3170	3030	3315
$\widehat{d}_{(2)}$	0.1494	0.3718	0.1595	0.3443
$s.e.$	0.0462	0.0483	0.0468	0.0502
Inf	0.0588	0.2771	0.0677	0.2458
Sup	0.2400	0.4666	0.2512	0.4427
$\widehat{\gamma}$	6.4883	4.7788	8.6096	8.0005

Note: see note of table 3.

Figure 1. Series of d_t ordered by increasing values of the transition variable X_{t-1}