Working Papers / Documents de travail

A Smooth Transition Long-Memory Model

Marcel Aloy
Gilles Dufrenot
Charles Lai Tong
Anne Peguin-Feissolle

WP 2012 - Nr 40

A smooth transition long-memory model

Marcel ALOY
Aix-Marseille University (Aix-Marseille School of Economics), CNRS & EHESS

Gilles DUFRENOT
Aix-Marseille University (Aix-Marseille School of Economics), CNRS & EHESS, Banque de France and CEPII

Charles LAI TONG
Aix-Marseille University (Aix-Marseille School of Economics), CNRS & EHESS

Anne PEGUIN-FEISSOLLE
Aix-Marseille University (Aix-Marseille School of Economics), CNRS & EHESS

December 2012

Abstract: This paper proposes a new fractional model with a time-varying long-memory parameter. The latter evolves nonlinearly according to a transition variable through a logistic function. We present a LR-based test that allows to discriminate between the standard fractional model and our model. We further apply the nonlinear least squares method to estimate the long memory parameter. We present an application to the unemployment rate in the United States from 1948 to 2012.

Keywords: Long-memory - nonlinearity- time varying parameter - logistic

JEL classification: C22, C51, C58

1 Introduction

This paper proposes a new model based on the commonly used fractional Gaussian noise model extended to the case in which the long-memory parameter is time-varying and varies smoothly according to a variable that cause regime shifts. Our model is a particular type of generalized autoregressive models with time-varying coefficients that have become popular in the econometrics literature. Consider the following infinite AR representation of a process X_t with time-varying coefficients:

$$X_t = \sum_{i=1}^{\infty} \beta_{it} X_{t-i} + \varepsilon_t, \quad \varepsilon_t \sim NID(0, \sigma^2) \tag{1}$$

The coefficients β_{it} are allowed to have different values for different t and we denote X_t^* the equilibrium trajectory of X_t when $t \rightarrow \infty$. This general equation has given rise to several models in the econometrics literature : Kalman filter models (Anderson and Moore (1979)), models based on simulated method of moments (Gallant et al. (1997)), Importance sampling models (Durbin and Koopman (2002)), particle filtering and particle MCMC models (Kitagawa (1996)), Markov Chain Monte Carlo based on Bayesian models (de Jong and Shepard (1995)).

Several modelling options for the time-varying coefficients are the following: breaks (with different β_{it} in different pre-defined periods), Markov-switching structures, deterministic functions of exogenous variables, stochastic functions of time.

Here, we consider the behavior of X_t when the following structure is imposed on the coefficients. We define a function $d_t = F(z_t)$ where z_t is an exogenous variable. The autoregressive coefficients obey the following structure:

$$\beta_{1t} = -d_t,\ \beta_{2t} = \frac{1}{2} d_t (1 - d_t), ..., \beta_{it} = \frac{1}{i} \beta_{i-1,t} (i - 1 - d_t), i \geq 3 \tag{2}$$

Under these assumptions, the infinite AR process (1) can be written as a time-varying long-memory model:

$$(1 - L)^{d_t} X_t = \varepsilon_{t,}\ \ d_t = F(z_t) \tag{3}$$

This process is an extension of the standard ARFIMA model with constant long-memory parameter. The formulation (3) has several advantages over the classical model (d constant).

Indeed, from an economic point of view, it is useful to model phenomena that are characterized by changing memory. In other words, it is important to specify models able to take into account the possibility of time-varying persistence. In previous papers, it has been suggested that the degree of integration of macroeconomic and financial time series is not necessarily constant over time. Instead, it may vary if the economic structure itself changes over time, or if markets are permanently hit by different shocks. Authors in the literature have formalized this idea in several ways : multi-scale or multi-fractal processes (Alvarez-Martinez et al. (2010), Engelen et al. (2011), Pasquini and Serva (1999)), support-vector machine processes (Gravishchaka and Ganguli (2003)) long-memory models with threshold transition dynamics (Boutahar et al. (2008), Dufrénot et al. (2005a, 2005b, 2008)), fractional models with regime shifts (Aloy et al. (2010)).

From an econometric point of view, Equation (3) has also some advantages in terms of better specification. Firstly, this time-varying parameter model can help hedging against misspecification when one considers time series models. Secondly, assuming that the long-memory parameter is not constant, our model allows to find which economic variable causes the changes in memory. Thirdly, in contrast to other models proposed in the literature on time-varying $ARFIMA$ models (for instance seasonal and periodic $ARFIMA$ models as in Koopman et al. (2007)), we do not consider that d varies at a regular frequency, but evolves smoothly with a dynamics described by a logistic function; this function has several attractive properties. We apply this model to the unemployment rate in the United States from 1948 to 2012.

The remainder of the paper is organized as follows. Section 2 presents the model and its main characteristics. Section 3 contains some properties of the model. Section 4 deals with statistical inference (estimation and test). Section 5 presents an empirical illustration. Finally Section 6 concludes.

2 A long-memory model with smooth transition dynamics

The model we propose can be used to describe any phenomenon for which the long-memory dynamics is subject to regime switches or smooth structural changes. Consider for instance two variables: X_t is the endogenous variable and z_t is a transition variable. Assume that the latter describes two extreme regimes corresponding to times of crises and normal periods. As is known, in times of crisis it takes more time for shocks to dissipate in the markets than during calm periods. This implies that the series show a more persistent dynamics when markets become more turbulent. A crisis can be more or less deep, thereby implying that the changes occurring in the variable z_t are not abrupt but smooth. Therefore, the degree of persistence of

X_t may vary smoothly according to the dynamics of the variable z_t as follows:

$$\Phi(L)(1-L)^{d(z_t)}X_t = \Theta(L)\varepsilon_t, \tag{4}$$

with

$$d(z_t) = \begin{cases} d_{(1)} & if \quad z_t \in R_{(1)} \\ d_{(2)} & if \quad z_t \in R_{(2)} \\ \vdots & \vdots \quad\quad \vdots \\ d_{(S)} & if \quad z_t \in R_{(S)} \end{cases}, \tag{5}$$

$$\Phi(L) = \begin{cases} \Phi_{(1)}(L) & if \quad z_t \in R_{(1)} \\ \Phi_{(2)}(L) & if \quad z_t \in R_{(2)} \\ \vdots & \vdots \quad\quad \vdots \\ \Phi_{(S)}(L) & if \quad z_t \in R_{(S)} \end{cases}, \tag{6}$$

and

$$\Theta(L) = \begin{cases} \Theta_{(1)}(L) & if \quad z_t \in R_{(1)} \\ \Theta_{(2)}(L) & if \quad z_t \in R_{(2)} \\ \vdots & \vdots \quad\quad \vdots \\ \Theta_{(S)}(L) & if \quad z_t \in R_{(S)} \end{cases}; \tag{7}$$

ε_t is a zero mean Gaussian noise with finite variance σ^2, L is the backward shift operator defined by $L^j X_t = X_{t-j}$, $\Phi_{(j)}(L)$ and $\Theta_{(j)}(L)$ are stable polynomials, i.e. their roots are strictly outside the unit circle. The $R_{(j)}$ are sub-intervals of variation of z_t and the $d_{(j)}$, $j = 1,...S$, are real numbers with $-0.5 < d_{(1)} < d_{(2)} < \ldots < d_{(S)} < 0.5$**.** The lowest value of the long-memory parameter is higher than -0.5 in order to have invertibility. The highest value is less than 0.5 in order to have local stationarity (see Section 3). The long-memory coefficient $d(z_t)$ is time-varying because it takes different values according to the regime $R_{(j)}$.

Equation (4) can also be written as follows:

$$\Phi(L)\sum_{j=0}^{\infty}\begin{pmatrix} d(z_t) \\ j \end{pmatrix}(-1)^j X_{t-j} = \Theta(L)\varepsilon_t. \tag{8}$$

The model given by (5), (6), (7) and (8) is similar to previous models proposed in the literature. For instance, Haldrup and Nielsen (2006a and 2006b) consider the case in which z_t is an observable Markov-chain. Aloy et al. (2010), Boutahar et al. (2008), Dufrénot et al. (2005a, 2005b and 2008), Goldman and Tsurumi (2006), Lahiani and Scaillet (2009) consider the case in which z_t is a random variable. These models however present some limitations when the number of regimes becomes very high because of computational problems. They require very long time series for providing reliable estimates of all the parameters. Further, the behavior of the coefficient estimates when the number of regimes becomes very high is unknown.

In this paper, we assume the number of regimes S "visited" by $d(z_t)$ tends to infinity. In this case, $d(z_t)$ can be replaced by a continuous function. We consider a sigmoid function to model smooth transition between the different regimes:

$$d(z_t) = d_{(1)} + (d_{(S)} - d_{(1)})F(z_t), \tag{9}$$

where F is a function characterized by $F : \mathbb{R} \to [0,1]$. F must be general enough to capture different situations: there may be few regimes or an infinity between the two extreme values $d_{(1)}$ and $d_{(S)}$. The occurrence of these different cases can be captured by a parameter giving the smoothness of the "transition" between $d_{(1)}$ and $d_{(S)}$. The higher the number of regimes "visited" by $d(z_t)$, the smoother the transition between the extreme values of the long-memory parameter. Though there are several functions that can be candidates, we consider here a logistic function:

$$F(z_t) = [1 + \exp(-\gamma z_t)]^{-1}, \quad \gamma > 0. \tag{10}$$

$F(z_t)$ describes the deterministic process generating the sequence of time-varying long-memory parameter $d(z_t)$. Therefore, we have

$$d(z_t) \in \left[d_{(1)}, d_{(S)}\right]. \tag{11}$$

The variable z_t may or may not be stationary. This does not matter here because the function F maps z_t into a bounded interval, here the unit interval $[0,1]$. We do not consider the case where z_t is a lagged value of the endogenous variable since it makes the testing problem more cumbersome.

This new model is compatible with a variety of shapes of time series, some indicating regime switches and others not. Figure 1 provides some illustrations for different values of the parameters for the fractional Gaussian noise.

To simplify the exposition, we consider the simpler model with no short-term components ($\Phi_{(j)}(L) = \Theta_{(j)}(L) = 1, \forall j$):

$$(1-L)^{d(z_t)} X_t = \varepsilon_t. \tag{12}$$

The extension of our arguments to the general case is straightforward in the case where ε_t is a stationary ARMA process. In what follows, we assume that X_t is a zero-mean process.

3 Properties of the model: issues and problems

3.1 Local stationarity

Before presenting the estimation and test of our model, we briefly discuss some of its properties and some problems raised by such a model. The stationarity assumption is important for identification, estimation and forecast. In our case, it is important to distinguish between local and global stationarity. Global stationarity (or equivalently strict stationarity) implies that the distribution of $X_{t_1}, X_{t_2}, ..., X_{t_k}$ (where $t_1, t_2, ..., t_k$ is a collection of k positive integers) is invariant under time shifts. This is not the case here, since the distribution is a function of $F(z_t)$ which takes different values in time according to the observations of z_t. In a standard ARFIMA model with d constant, global stationarity also implies that the process is ergodic or satisfies mixing conditions in the sense that observations sufficiently far from each others are almost uncorrelated. Once we assume that d is time-varying, this is no longer true.

When global stationarity fails, the stationarity of a process like ours is defined only locally.

Rewriting (12) using the infinite autoregressive representation of X_t yields:

$$X_t = -\sum_{j=1}^{\infty} a_{jt} X_{t-j} + \varepsilon_t, \quad a_{jt} = \frac{\Gamma[j - d(z_t)]}{\Gamma(j+1)\Gamma[-d(z_t)]}, \tag{13}$$

where Γ is the Euler gamma function. The local stationarity of the process implies that, for a given z_t, a_{jt} is square summable ($\sum_{j=0}^{\infty} a_{jt}^2 < \infty$ for a given z_t). For examples of local stationary ARFIMA models with a time-varying parameter, the reader can refer to Beran, Bhansali and Ocker (1998), Vesilo and Chan (1996), Lavielle and Ludena (2000). This is the case if we assume that $d_{(S)}$, the highest value of $d(z_t)$, is less than 0.5.

Our model encompasses several cases of locally stationary processes. With the representation (5) $d(z_t)$ is partitioned into stationary sub-intervals. The process may be subject to abrupt changes during the transition between the different regimes, but stationary within each segment of variation of $d(z_t)$. An extreme illustration is the case where $\gamma \to \infty$ in Equation (10). In this case we simply have two regimes with the long-memory parameter defined by two values $d_{(1)}$ and $d_{(S)}$. However, when $\gamma < \infty$ and is small enough so that the slope of the logistic function becomes smooth, the process is an evolutionary process in the sense that its probabilistic characteristics (spectral density function and autocovariance functions) smoothly change over time.

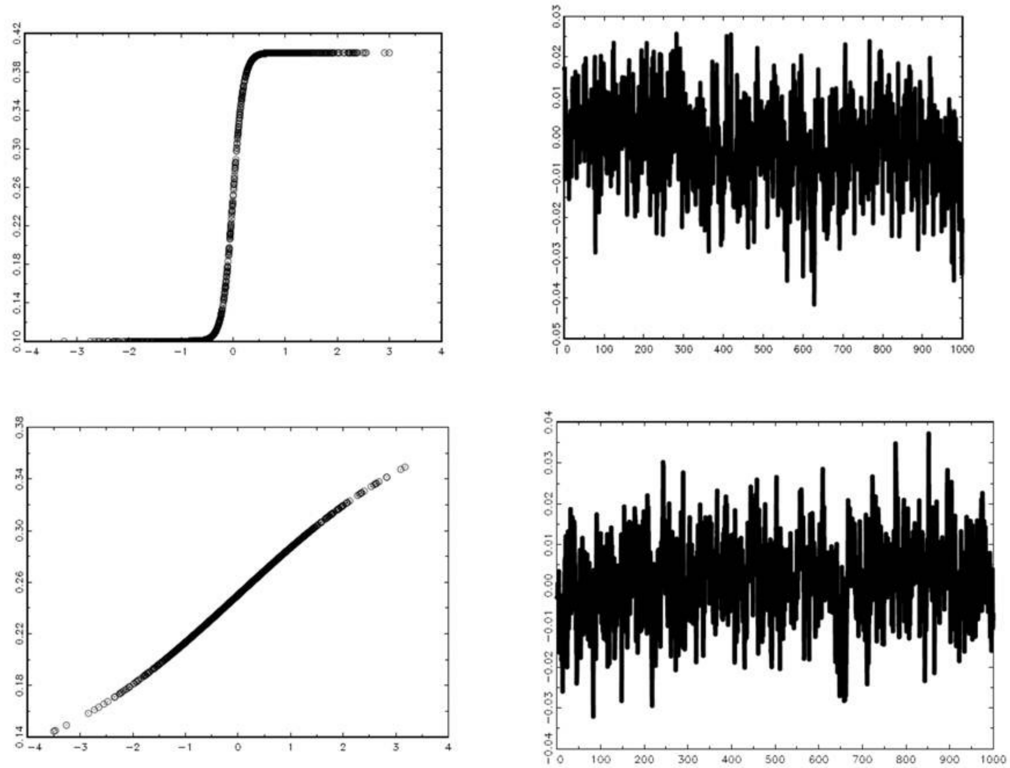

Figure 1. Examples of time series X_t, $t = 1, ..., 1000$, generated by the smooth transition long-memory model

Note:

$d_{(1)} = 0.1, d_{(S)} = 0.4, \gamma = 10$*:* d_t *as a function of* z_t *(left upper panel),* d_t *as a function of time (right upper panel)*

$d_{(1)} = 0.1, d_{(S)} = 0.4, \gamma = 0.5$*:* d_t *as a function of* z_t *(left bottom panel),* d_t *as a function of time (right bottom panel)*

This leads us to consider another definition of locally stationary process based on the evolutionary spectrum, proposed by Dalhaus (1996). One considers the generalized Cramer representation of a stationary process in which X_t is modelled as a white noise excited time-variant shaping filter:

$$X_t = \int_{-\pi}^{\pi} \exp(i\omega t) A_t^0(\omega) d\varepsilon(\omega), \tag{14}$$

where $\varepsilon(\omega)$ is a white noise with orthogonal increments on $[-\pi,\pi]$ with $\overline{\varepsilon(\omega)} = \varepsilon(-\omega)$. $A_t^0(\omega)$ is the modulation function also called transfer function. Our process is considered to be locally stationary in the Dalhaus sense if the following two conditions hold:

i) There exists a constant K and a 2π-periodic function $A : [0,1]\times R \to C$ with $\overline{A(x,\omega)} = A(-x,\omega)$ such that

$$\sup_{t,\omega} \left|A_t^0(\omega) - A(d(z_t),\omega))\right| < KT^{-1}, \ \ \forall t \tag{15}$$

and $A(x,\omega)$ is continuous in x. T is the number of observation of the sample.

ii) $A(x,\omega)$ is differentiable in x with uniformly bounded derivative $\frac{\partial^2 A}{\partial x \partial \omega}$.

In this definition, it is the smoothness of the modulation function A with respect to x that ensures the local stationarity. In our case, this implies that the modulation function must be a smooth function of z_t, which depends upon the smoothness of the function $F(z_t)$. The problem is then to compute the analytical expression of the transfer function, which is usually done by first writing the infinite MA representation of the process and taking the Fourier transform of the coefficients. An important point here is that, in our case, the solution is not

$$X_t = \sum_{j=0}^{\infty} b_{jt}\varepsilon_{t-j} \tag{16}$$

with

$$b_{jt} = \frac{\Gamma[j+d(z_t)]}{\Gamma(j+1)\Gamma[d(z_t)]}$$

but has the following more complicated formulation:

$$X_t = \sum_{i=0}^{\infty} \alpha_{it}\varepsilon_{t-i} \tag{17}$$

with

$$\begin{cases} \alpha_{0t} = & 1 \\ \alpha_{1t} = & d(z_t) \\ \alpha_{2t} = & \alpha_{1t-1} + \frac{d(z_t)(1-d(z_t))}{2} \\ \alpha_{3t} = & d(z_t)\alpha_{2t-1} + \frac{d(z_t)(1-d(z_t))}{2}\alpha_{1t-2} + \frac{d(z_t)(1-d(z_t))(2-d(z_t))}{3!} \\ \alpha_{4t} = & d(z_t)\alpha_{3t-1} + \frac{d(z_t)(1-d(z_t))}{2}\alpha_{2t-2} + \frac{d(z_t)(1-d(z_t))(2-d(z_t))}{3!}\alpha_{1t-3} \\ & + \frac{d(z_t)(1-d(z_t))(2-d(z_t))(3-d(z_t))}{4!} \\ & \vdots \end{cases}$$

Equation (16) is the inverse of (13) only when d is constant. Usually, papers dealing with a time-varying parameter d assume that $\widetilde{X}_t$ is the inverse of (13), which is not exact. To cope with this problem Philippe et al. (2008) introduce new classes of time-varying linear filters. However, the latter are not considered here since they do not match our Equation (13).

3.2 Long-memory properties

To sum up, the long-memory behavior of ARFIMA models is usually apprehended by looking at the behavior of the autocovariance function and the spectral density function. A standard approach is to consider the Wold representation (infinite moving average) of the process $(1-L)^d X_t = \varepsilon_t$ and to compute the spectral density using the transfer function. Then the autocovariance function is found using the Wiener-Kitchine representation theorem. In our case, we cannot use Equation (16) to compute the spectral density of (13), since the solution is given by Equation (17). Computing the transfer function and spectral density from the latter is quite cumbersome. The autocovariance function of X_t can be obtained by looking at the limit behavior of the sequence $\{d(z_t)\}$. As shown in other papers (for examples, see Surgailis (2008), Philippe et al. (2008), Whitcher and Jensen (2000)) studying ARFIMA models with a time-varying d, the autocovariance function verifies:

$$ACF(t, t-\tau) \propto \tau^{2\widetilde{d}-1} \text{ as } \tau \to \infty \tag{18}$$

where $\widetilde{d}$ is a constant that depends upon the limit behavior of the sequence $\{d(z_t)\}$. This function decays at an hyperbolic rate towards 0.

Now, to define the long-memory properties of our process in terms of the behavior of the spectral density, we would need an equivalent of the Wiener-Kitchine theorem for long-memory processes with a constant d parameter. The theorem states that for a stationary process the power spectrum is equal to the Fourier transform of the autocorrelation function. However, for processes that are locally (but not globally) stationary, this connection between the ACF and the spectrum does no longer hold. Indeed, as shown by Mark (1970), Priestley (1971), Rao (1979), Rees (1972), Tsao (1984), the spectral density of processes that are globally nonstationary is not the simple inverse Fourier transform of the autocovariance function. A locally hyperbolic decay of the autocovariance function does not necessarily imply an explosion of the evolutionary spectrum at the origin. The easiest way to describe the long memory properties of our model is thus to examine the behavior of the autocovariance function. The latter has a common feature with a standard $ARFIMA$ model with an ACF decaying at an hyperbolic rate. However, since d is time-varying, the rate of decay depends upon the limit behavior of the sequence $\{d_t\}$. It is thus important that this sequence converges and the limit needs not be defined by a singleton.

4 Statistical inference: estimation and tests

4.1 Estimation by nonlinear least squares

We consider a conditional sum-of-squares residual estimator which minimizes

$$RSS(\mathbf{X},\Theta) = T^{-1}\sum_{t=1}^{T}\left[X_t - M_t(\Theta)\right]^2 \tag{19}$$

with

$$M_t(\Theta) = -\sum_{j=1}^{t-1} a^*_{jt} X_{t-j} \text{ and } a^*_{jt} = \frac{\Gamma[j-d(z_t)]}{\Gamma(j+1)\Gamma[-d(z_t)]}$$

where $d(z_t)$ is defined by (9) and (10), and $\Theta = (d_{(1)}, d_{(S)}, \gamma)$. $\mathbf{X}$ is the sample which consists of T observations. The initial infinite autoregressive process is truncated by assuming that $X_0 = X_{-1} = X_{-2} = ... = 0$. The first-order conditions do not lead to an explicit solution for Θ. The value of Θ which minimizes $RSS(\mathbf{X},\Theta)$ is found using numerical methods (for instance, the Gauss-Newton algorithm). This is a standard nonlinear minimization problem. However, we need some regularity conditions in order to be able to apply a uniform law of large number to $RSS(\mathbf{X},\Theta)$. Firstly, the model is asymptotically identified if $\gamma > 0$ and $d_{(1)} \neq d_{(S)}$. Secondly, we assume that $X_t = M_t(\Theta) + \varepsilon_t$ is non-explosive for any value of z_t. This yields to impose restrictions on $d_{(1)}$ and $d_{(S)}$ since $F(z_t)$ is bounded. We assume that $-0.5 < d_{(1)} < d_{(S)} < 0.5$. Therefore, for a given z_t, the characteristic polynomial of finite order T obtained using the coefficient a^*_{jt} has all its roots outside the unit circle and the variance of X_t is bounded. Using standard arguments, we know that, when the long-memory parameter is less than 0.5, the expectation of the process is uniformly bounded.

Simulations of the nonlinear least-squares (NLS) estimator. In order to get insight into the distribution of the estimator and its properties, we conduct Monte Carlo simulations. We first generate z_t, the transition variable, as an AR(2) process:

$$z_t = 1.45 z_{t-1} - 0.5 z_{t-2} + \upsilon_t, \quad \upsilon_t \approx iid.N(0,1). \tag{20}$$

With these parameters, z_t has an almost periodic behavior. In this case, the sequence $\{d(z_t)\}$ is bounded and averageable with a mean value

$$\widetilde{d} = \lim_{t-s\to+\infty} \frac{1}{t-s}\sum_{u=s}^{t} d_u. \tag{21}$$

We compute $d(z_t)$, for values of $d_{(1)}$ corresponding to -0.4 and 0.1, $d_{(S)}$ corresponding to 0.3 and 0.4, γ corresponding to 2 and 10. The simulations are based on 1000 replications. For each replication, the parameters $d_{(1)}$, $d_{(S)}$, and γ are estimated by NLS method. We consider different sample sizes $T = 200, 500, 1000, 2000$. In order to reduce the effect of initial values, the first 200 observations are discarded. We report the graphs of the distributions of the parameters. We also compute the average of the 1000 estimates ($MEAN$), the bias ($BIAS$), the root mean squared error ($RMSE$) and an indicator of the distance between the true value of $d(z_t)$ and $\widehat{d}^{(i)}(z_t)$, the value of $d(z_t)$ obtained at the $i-th$ replication, $i = 1, ..., N$, given by

$$DIST = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \left(\widehat{d}^{(i)}(z_t) - d(z_t) \right)^2 . \tag{22}$$

Simulation results are shown in Table 1. We can observe the convergence of the parameters as T increases. Indeed, as is seen, the cases corresponding to $T = 2000$ produces the least biases. Therefore the number of observations should be as large as possible in order to minimize the distance between the true observations of $d(z_t)$ and their estimated values. We also see that the rate of convergence of $d_{(1)}$ and γ increases as the difference between $d_{(1)}$ and $d_{(S)}$ becomes wider (compare the case where $d_{(1)} = 0.1$ and $d_{(S)} = 0.4$ with the case where $d_{(1)} = -0.4$ and $d_{(S)} = 0.3$). This means that when the difference $\left| d_{(S)} - d_{(1)} \right|$ is less than a given value, say η, the long-memory estimator converges at a rate lower than the usual rate of $\sqrt{T}$ in one of the two extreme regimes. This happens for the following reason. When $d_{(1)}$ and $d_{(S)}$ are "close", our model behaves like an ARFIMA model with a constant d parameter. Therefore it becomes weakly identified since γ and either $d_{(1)}$ or $d_{(S)}$ are nuisance parameters. In this case, several combinations of $(d_{(1)}, d_{(S)}, \gamma)$ can yield the same minimum of the nonlinear least squares functions if the number of observations is not large enough.

Figure 2 reports the distributions of $d_{(1)}$ and $d_{(S)}$. We see that they are asymptotically normal.

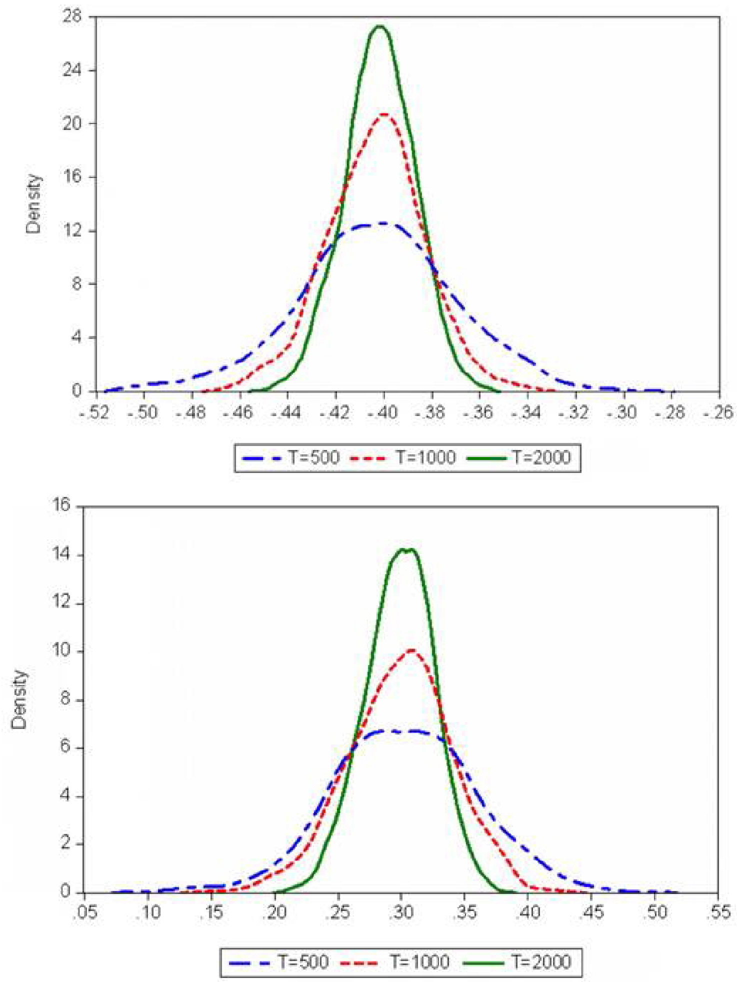

Figure 2. Simulated distributions of $d_{(1)}$ (upper panel) and $d_{(s)}$(bottom panel)

Table 1. Simulation results of the NLS estimator

$d_{(1)} = 0.1$, $d_{(s)} = 0.4$, $\gamma = 2$

T		*MEAN*	*BIAS*	*RMSE*	*DIST*
200	$d_{(1)}$	0.0463	−0.0537	0.1254	0.0060
	$d_{(s)}$	0.4308	0.0308	0.0948	
	γ	19.3356	17.3356	121.3511	
500	$d_{(1)}$	0.0687	−0.0313	0.0888	0.0026
	$d_{(s)}$	0.4189	0.0189	0.0803	
	γ	16.5262	14.5262	99.4681	
1000	$d_{(1)}$	0.0713	−0.0287	0.0789	0.0016
	$d_{(s)}$	0.4189	0.0189	0.0722	
	γ	9.0458	7.0458	58.5246	
2000	$d_{(1)}$	0.0802	−0.0198	0.0652	0.0008
	$d_{(s)}$	0.4166	0.0166	0.0613	
	γ	5.9641	3.9641	59.2108	

$d_{(1)} = 0.1$, $d_{(s)} = 0.4$, $\gamma = 10$

T		*MEAN*	*BIAS*	*RMSE*	*DIST*
200	$d_{(1)}$	0.0620	−0.0380	0.0947	0.0063
	$d_{(s)}$	0.4107	0.0107	0.0906	
	γ	63.3535	53.3535	262.7986	
500	$d_{(1)}$	0.0812	−0.0188	0.0602	0.0029
	$d_{(s)}$	0.4123	0.0123	0.0594	
	γ	33.4306	23.4306	100.4333	
1000	$d_{(1)}$	0.0876	−0.0124	0.0439	0.0017
	$d_{(s)}$	0.4049	0.0049	0.0447	
	γ	60.4828	50.4828	443.5719	
2000	$d_{(1)}$	0.0935	−0.0065	0.0291	0.0009
	$d_{(s)}$	0.4041	0.0041	0.0318	
	γ	35.0520	25.0520	222.5527	

$d_{(1)} = -0.4$, $d_{(s)} = 0.3$, $\gamma = 2$

T		*MEAN*	*BIAS*	*RMSE*	*DIST*
200	$d_{(1)}$	−0.4151	−0.0151	0.0864	0.0062
	$d_{(s)}$	0.2906	−0.0094	0.1264	
	γ	12.2564	10.2564	111.6332	
500	$d_{(1)}$	−0.4118	−0.0118	0.0734	0.0026
	$d_{(s)}$	0.3048	0.0048	0.0983	
	γ	2.8777	0.8777	4.5415	
1000	$d_{(1)}$	−0.4143	−0.0143	0.0621	0.0013
	$d_{(s)}$	0.3080	0.0080	0.0749	
	γ	2.2451	0.2451	1.2811	
2000	$d_{(1)}$	−0.4092	−0.0092	0.0496	0.0006
	$d_{(s)}$	0.3050	0.0050	0.0536	
	γ	2.0881	0.0881	0.6042	

$d_{(1)} = -0.4$, $d_{(s)} = 0.3$, $\gamma = 10$

T		*MEAN*	*BIAS*	*RMSE*	*DIST*
200	$d_{(1)}$	−0.4121	−0.0121	0.0499	0.0063
	$d_{(s)}$	0.2942	−0.0058	0.1059	
	γ	35.9725	25.9725	173.6601	
500	$d_{(1)}$	−0.4030	−0.0030	0.0330	0.0025
	$d_{(s)}$	0.3021	0.0021	0.0561	
	γ	14.9033	4.9033	48.0806	
1000	$d_{(1)}$	−0.4034	−0.0034	0.0206	0.0012
	$d_{(s)}$	0.2988	−0.0012	0.0410	
	γ	10.8948	0.8948	5.1950	
2000	$d_{(1)}$	−0.4019	−0.0019	0.0146	0.0005
	$d_{(s)}$	0.2986	−0.0014	0.0269	
	γ	10.4179	0.4179	2.6047	

Note: The number of observations is T; the number of replications N is 1000. MEAN is the mean, BIAS is the parameter bias and RMSE is the root mean squared error. DIST is an indicator of the distance between the true value of d_t and $\widehat{d}_t^{(i)}$, the value of d_t obtained at the $i-th$ replication, $i = 1, \ldots, N$, given by $DIST = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\widehat{d}_t^{(i)} - d_t\right)^2$. z_t is generated as a $AR(2)$ process: $z_t = 1.45z_{t-1} - 0.5z_{t-2} + v_t$, $v_t \sim IIN(0,1)$; $\varepsilon_t \sim IIN(0, \sigma_\varepsilon^2)$, $\sigma_\varepsilon^2 = 0.01$.

4.2 Testing the constancy of the long-memory parameter

We consider the test of the null hypothesis of the constancy of the long memory parameter d_t:

$$H_0 : \gamma = 0 \tag{23}$$

in the model:

$$\begin{cases} (1-L)^{d(z_t)} X_t = \varepsilon_t \\ d(z_t) = d_{(1)} + (d_{(S)} - d_{(1)}) F(z_t) \\ F(z_t) = [1 + \exp(-\gamma z_t)]^{-1} - 1/2. \end{cases} \tag{24}$$

Under H_0, we thus have $F(z_t) = 0$, then $d_t = d_{(1)}$ and the model is a standard fractional Gaussian model:

$$(1-L)^{d_{(1)}} X_t = \varepsilon_t. \tag{25}$$

Under H_0, $d_{(S)}$ is not identified. As has become common practice in the literature on smooth autoregressive models (see for instance Luukkonen et al. (1988) and Teräsvirta (1994)), we replace $F(z_t)$ by its third-order Taylor approximation around $\gamma = 0$, i.e.

$$\widetilde{F}_3(z_t) = \frac{z_t}{4}\gamma - \frac{z_t^3}{48}\gamma^3. \tag{26}$$

Replacing $F(z_t)$ by $\widetilde{F}_3(z_t)$ and reparametrizing the model, we obtain the approximated model under the alternative hypothesis:

$$\begin{cases} (1-L)^{d_t} X_t = \varepsilon_t \\ d_t = \omega + \delta z_t + \phi z_t^3 \end{cases}. \tag{27}$$

with

$$\begin{cases} \omega = d_{(1)} \\ \delta = \left(d_{(S)} - d_{(1)}\right) \frac{\gamma}{4} \\ \phi = -\left(d_{(S)} - d_{(1)}\right) \frac{\gamma^3}{48}. \end{cases} \tag{28}$$

The null hypothesis of the constancy of d_t can be written as:

$$H_0' : \delta = \phi = 0. \tag{29}$$

From (28), it is clear that δ and ϕ are colinear, so that $\phi = 0$ automatically implies that $\delta = 0$ (we only have one constraint under the null hypothesis). Therefore testing $H_0' : \delta = \phi = 0$ is equivalent to testing

$$H_0'' : \phi = 0. \tag{30}$$

Under $H_0^{''}$, we thus have the standard Gaussian fractional model

$$(1-L)^{\omega} X_t = \varepsilon_t. \tag{31}$$

To test $H_0^{''}$, we use a likelihood ratio testing procedure. The test statistics is equal to the difference between the unconstrained and constrained maximum values of the log-likelihood function, say

$$LR = 2\left(L_T\left(\widehat{\theta},\widehat{\sigma}\right) - L_T\left(\widetilde{\theta},\widetilde{\sigma}\right)\right); \tag{32}$$

$L_T(\theta,\sigma)$ is the log-likelihood function of the model (27) defined by

$$L_T(\theta,\sigma) = -\frac{T}{2}\log\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2}\sum_{t=1}^{T}\left[(1-L)^{d_t}X_t\right]^2 \tag{33}$$

with $d_t = \omega + \delta z_t + \phi z_t^3$. θ is given by $\theta = (\omega,\delta,\phi)'$ and σ^2 is the variance of ε_t. Moreover,

$$\widehat{\theta} = \left(\widehat{\omega},\widehat{\delta},\widehat{\phi}\right)' \text{ and } \widetilde{\theta} = (\widetilde{\omega},0,0)'$$

(resp. $\hat{\sigma}$ and $\widetilde{\sigma}$) denote the unrestricted and restricted maximum likelihood estimates of θ (resp. σ).

The LR statistic can also be written is terms of $\widehat{\sigma}^2$and $\widetilde{\sigma}^2$ the unrestricted and restricted estimates of σ^2:

$$LR = T\log\left(\frac{\widetilde{\sigma}^2}{\widehat{\sigma}^2}\right) = T\log\left(1+\frac{\widetilde{\sigma}^2-\widehat{\sigma}^2}{\widehat{\sigma}^2}\right). \tag{34}$$

Defining $\widehat{\varepsilon}_t = (1-L)^{\widehat{d_t}}X_t$ with $\widehat{d_t} = \widehat{\omega} + \widehat{\delta}z_t + \widehat{\phi}z_t^3$, we have the consistency

$$\widehat{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T}\widehat{\varepsilon}_t^2 \xrightarrow{p} \sigma^2 \tag{35}$$

under $H_0 : \gamma = 0$ which corresponds to a standard $I(d)$ process. The existence of a distribution for LR under H_0 means that LR and

$$LR_0 = \frac{T}{\sigma^2}\left(\widetilde{\sigma}^2 - \widehat{\sigma}^2\right) \tag{36}$$

have the same nondegenerate distribution.

From the standard analysis of nonlinear models, the statistic LR is asymptotically distributed as a χ_1^2 under $H_0^{''} : \phi = 0$. As stated before, the assumption

$\phi = 0$ implies that $\delta = 0$. Thus, we only have one constraint under the null (it is equivalent to test $H_0^{'} : \delta = \phi = 0$ or $H_0^{''} : \phi = 0$).The parameters are estimated in the models under the null or the alternative hypothesis using the NLS estimator described in the preceding section. Tables 2 and 3 contain the results of simulations of the test[1]. The nominal size of the test is 5%. It is seen that it behaves well both in terms of power and size.

Table 2. Rejection frequencies in % of the null hypothesis $H_0 : \gamma = 0$ in the model generated under the null $(1-L)^d X_t = \varepsilon_t$

	$T = 500$	$T = 1000$
$d = 0.2$	2.80	3.90
$d = 0.4$	2.50	3.20
$d = 0.7$	1.70	3.30

Note: The number of observations is T; the number of replications N is 1000. z_t is generated as a $AR(1)$ process: $z_t = 0.8z_{t-1} + \nu_t$, $\nu_t \sim IIN(0,1)$; $\varepsilon_t \sim IIN(0,1)$.

5 An empirical illustration

We model the feedback dynamics of the unemployment rate to its natural rate in the United States from 1948 to 2012 using monthly data. As is known from the labor economics literature, the dynamics of the unemployment can be characterized by strong hysteresis stemming from labor market imperfections (labor union collective bargaining, unemployment benefits, firing restrictions and government interventions in general). This implies that the NAIRU, the level of unemployment rate consistent with stable inflation, is time-varying rather than fixed because distortions in labor markets interact with disinflation. This has lead empirical economists to propose reduced-form equations linking the current level of the NAIRU to past unemployment rate (see, among others, Gordon (1997)). It is usually thought that such an hysteresis was not observed in the US because the institutional rules in the labor markets imply more flexibility than in Europe (see, for instance Di Tella and MacCulloch (2006)). However, hysteresis in the unemployment rate can also

[1] The case $d_{(1)} > 0.5$ and $d_{(S)} > 0.5$ is examined, in addition to the stationarity case, in order to see how the test behaves in the non-stationary case. From Table 3 we see that the rejection frequencies compare well with the stationary case.

be caused by economic shocks which have persistent effects (see Blanchard and Summers (1987)). Deviations from natural rate die out more rapidly in times of expansions, but are likely to be long-lived during strong recessions. Whether or not unemployment exhibits a rapid mean-reverting dynamics of its long-run level has been a matter of debate in the applied literature. Some authors respond positively (Nelson and Plosser (1982), Perron (1988)), while others reject this view (Mitchell (1993), Breitung (1994), Leon-Ledesma (2002)). In a recent paper, Cheng et al. (2012) find evidence of long-lived mean-reverting dynamics with a half-lived point estimate of 6 to 14 years, which is long compared to the typical duration of the business cycle.

Table 3. Rejection frequencies in % of the null hypothesis $H_0 : \gamma = 0$ in the model generated under the alternative, i.e. given by:

$$\begin{cases} (1-L)^{d_t} X_t = \varepsilon_t \\ d_t = d_{(1)} + \left(d_{(S)} - d_{(1)}\right) F(\gamma, z_t) \\ F(\gamma, c, z_t) = [1 + \exp(-\gamma z_t)]^{-1} - 1/2 \end{cases}$$

	$T = 500$	$T = 1000$
$d_{(1)} = 0.1$, $d_{(S)} = 0.4$, $\gamma = 2.5$	91.50	99.90
$d_{(1)} = 0.1$, $d_{(S)} = 0.4$, $\gamma = 5$	98.70	100.00
$d_{(1)} = 0.6$, $d_{(S)} = 0.9$, $\gamma = 5$	94.90	99.90

Note: The number of observations is T; the number of replications N is 1000. z_t is generated as a $AR(1)$ process: $z_t = 0.8 z_{t-1} + \nu_t$, $\nu_t \sim IIN(0,1)$; $\varepsilon_t \sim IIN(0,1)$.

We want to contribute to the empirical literature on this topic as follows. Papers testing the hysteresis of the unemployment rate in the US usually rely on unit root models (using either aggregate or panel series based on US states). The two polar hypothesis tested are, on the one side the stationarity of the series, and, on the other side, the non-stationarity. Using here the framework of long-memory processes yields a more parsimonious description of the mean-reverting dynamics which can occur more or less rapidly. We introduce a time-dependent dynamics, in the sense that we assume that the way in which the deviations of the unemployment rate from their long-run level die out varies with firms' expectations regarding the future state of the economy.

We estimate the following $ARFIMA(1,d_t,1)$ model:

$$(1-\phi_1 L)(1-L)^{d_t}(u_t - u_t^*) = (1+\theta_1 L)\varepsilon_t, \tag{37}$$

where

$$d_t = d_{(1)} + [d_{(S)} - d_{(1)}][1+\exp\{-\gamma(z_t-\mu)\}]^{-1}; \tag{38}$$

u_t is the current unemployment rate (in logarithm), u_t^* is its natural rate computed using a HP filter. z_t is the capacity utilization at time $t-1$. L is the lag operator, ϕ_1 and θ_1 are the AR and MA coefficients capturing the short-run mean-revertion dynamics, μ is a constant capturing a threshold in the transition variable. Using the NLS method to estimate all the parameters jointly, we obtain the results given in Table 4. The series of capacity utilization is available from 1967 onwards with data on the unemployment rate between 1948 and 1966 being used as presample for the filter.

As is seen, the estimated model captures both short and long memory dynamics in the mean-reverting behavior of the unemployment rate to its long-run value. Indeed, the ARMA coefficients are significant just as are $d_{(1)}$ and $d_{(s)}$. These results hold for different values of γ above 100 (the sum of squared residuals remains unchanged), so we fix this parameter to 100.

Figure 3 displays the time variation of d_t, while Figures 4 and 5 respectively show d_t as a function of the transition variable and the evolution of the log-unemployment rate (along with its long-term component). The model shows a regime in which the deviations of the unemployment rate from its long-run level die out more rapidly than in the other one. Indeed, we obtain values for $d_{(1)}$ and $d_{(S)}$ that are statistically different (see the LR test) and respectively equal to 0.26 and 0.45.

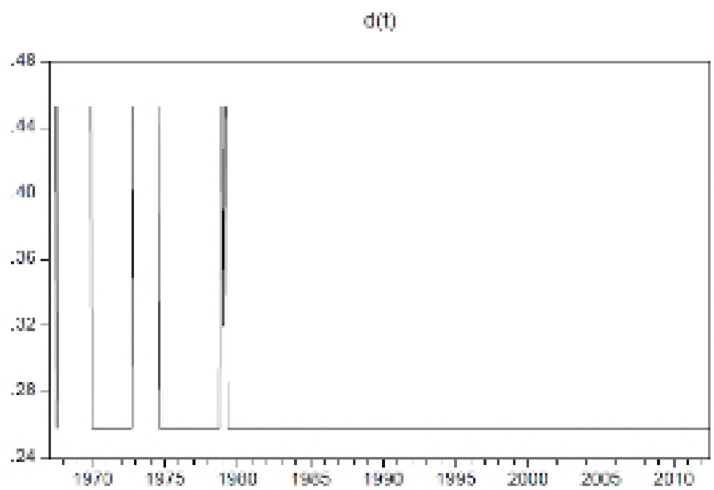

Figure 3. Evolution of d_t computed from our estimation

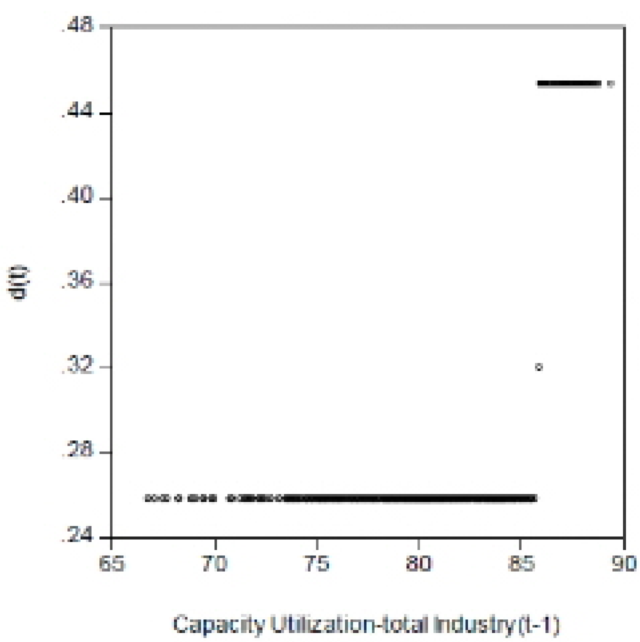

Figure 4. Scatterplot of d_t as a function of z_t

Capacity utilization provides information about future inflation and has also been found to be a leading indicator of the future stance of monetary policy in the US. As the industry resource becomes increasingly near its maximum level, enterprises anticipate a rise in future inflation. The expected rise in inflation tends to provide information regarding the future strength of monetary policy. Indeed, the FED has traditionally taken notes of changes in capacity utilization when trying to determine its policy with regard to inflation. As far as the capacity utilization is also a leading indicator of the business cycle and that the central bank policy is described by a Taylor rule, a rise in the current usage of resources is a signal a future lower output-gap and tension on prices, which may induce a reaction from the central bank through a rise of the federal fund rate. If firms thinks that the central bank could be concerned with the acceleration of inflation, then in their current hiring decisions they take into account the negative effect of a restrictive monetary policy on demand. The "pessimistic" view of the future stare of the economy can lead them to reduce hiring, thereby implying that unemployed people spend more time out of work. Our regression suggests that, if one agrees with this argument, then it may happen when the capacity utilization rate evolves above a threshold around 86% (the estimated value of the parameter μ).

Table 4. NLS estimation of the unemployment rate modeling

	coeff.	t-ratio
$\widehat{d}_{(1)}$	0.26	3.04
$\widehat{d}_{(S)}$	0.45	3.62
$\widehat{\gamma}$	100	–
$\widehat{\mu}$	85.91	1386
$\widehat{\phi}_1$	0.90	41.23
$\widehat{\theta}_1$	-0.25	-2.97
LR	16.45	$< 0.01^*$

Note : t-ratio is the Student-t ratio of the estimated coefficient. "$*$" is the $p-$value associated with the LR statistic.

Figure 3 shows that, until the second oil shock the unemployment rate changes (relative to its long-run value) in response to changes in the industry capacity utilization rate has been very sluggish with a local dynamics that was almost non-stationary (the values of d_t are near 0.5).

This finding of a local nonstationarity of the unemployment rate is in line with the results of Leon-Ledesma (2002), Cheng et al. (2012), but in their paper the nonstationarity is caused by the presence of a unit root in the unemployment rate series. In figure 5, the higher persistence in one regime relative to the other is illustrated by higher gap between the unemployment rate and its long-run component, up until the mid 1980s.

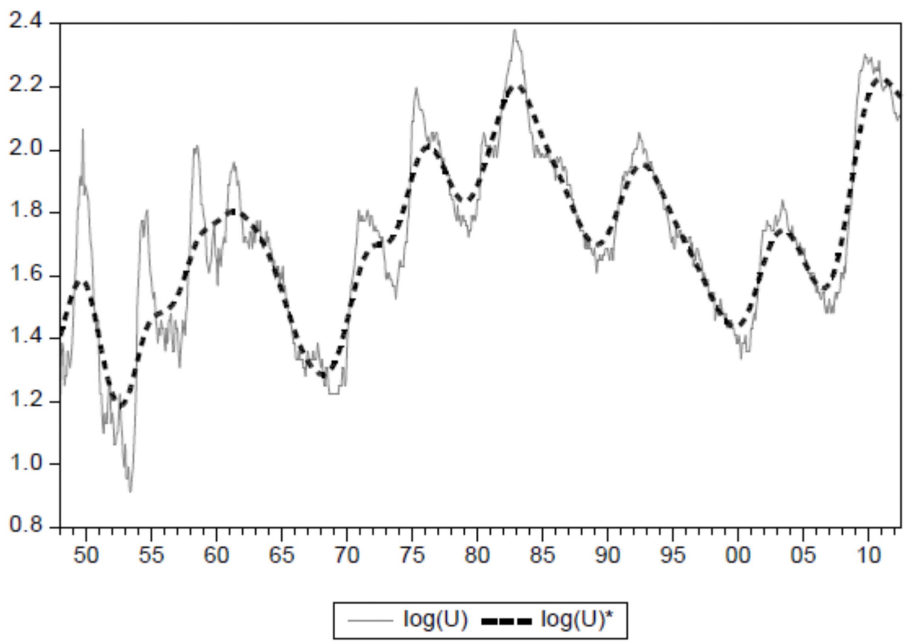

Figure 5. Log-unemployment rate and its long-run component

6 Conclusions and possible extensions

In this paper, we explore an innovative model with a time-varying long-memory parameter, evolving nonlinearly according to a transition variable through a logistic function. We investigate the main properties of this model, including the local stationarity and the specific infinite MA representation. Furthermore, we suggest an estimation method and a test of constancy of the long memory parameter. Our simulations show that the estimation method has good consistency properties in finite samples, provided that the difference between $d_{(1)}$ and $d_{(s)}$ is large enough. Moreover, our test exhibits good performances, both in terms of power and size in finite samples.

The model presented can be extended to any function describing a sigmoid transition of the long-memory parameter. Such functions are widely used in economics but also in the other sciences like physics, pharmacy, life sciences or agronomy, to describe a correlation between two variables featured by two *plateau* regions and a transition region. The function $d(z_t)$ described by Equations (9) and (10) corresponds to a particular case of situations where the memory dynamics of a process show some accumulation points around two values or multiple accumulation points around two polar values. Here are two examples:

(i) the typical five parameters logistic function with two polar values a and b for $d((z_t)$:

$$d(z_t) = a + \frac{(b-a)}{\left[1 + \left(\frac{z_t}{z_c}\right)^m\right]^n} \tag{39}$$

where a, b, m, n and z_c are constant. z_c is referred as the threshold value of z which determines the convergence trend of $d(z_t)$ when m is negative and n is a small positive number;

(ii) the generalized logistic function for which the polar values of the long-memory parameter are described by piecewise functions $g(z_t)$ and $h(z_t)$:

$$d(z_t) = g(z_t) + \frac{(h(z_t) - g(z_t))}{\left[1 + \left(\frac{z_t}{z_c}\right)^m\right]^n} \tag{40}$$

where m and n serve to control the smoothness of the two functions $g(z_t)$ and $h(z_t)$.

Finally, the model can be generalized to the case where z_t is not a centered variable (the transition function would be $z_t - c$, c being a constant).

References

[1] Aloy, M., Boutahar, M., Gente K. and A. Péguin-Feissolle, 2010, Fractional cointegration with regime shifts between equity markets. Working Paper, DEFI, GREQAM, Université d'Aix-Marseille.

[2] Alvarez-Ramirez, J., Alvarez, J. and R. Solis, 2010, Crude oil market efficiency and modeling: insights from the multiscaling autocorrelation pattern. *Energy Economics* 32, 5, 993-1000.

[3] Anderson, B. and J. Moore, 1979, Optimal filtering. Prentice Hall

[4] Beran, J., Bhansali, R.J. and D. Ocker, 1998, On unified model selection for stationary and nonstationary short- and long-memory autoregressive processes. *Biometrika*, 85(4), 921-934.

[5] Blanchard, O.J. and L.H. Summers, 1987, Hysteresis in unemployment. *European Economic Review*, 31, 288-295.

[6] Boutahar, M., Dufrénot, G. and A. Péguin-Feissolle, 2008, A simple fractionally integrated model with a time-varying long memory parameter d_t. *Computational Economics* 31, 3, 225-241.

[7] Breitung, J. 1994, Some simple tests of the moving average hypothesis. *Journal of Time Series Analysis*, 15, 351-359.

[8] Cheng, K.M., Durmaz, N., Kim, H. and M. Stern, 2012, Hysteresis vs natural rate of US unemployment. *Economic Modelling*, 29(2), 428-434.

[9] Dahlaus, R., 1996, On the Kullback-Leibler information divergence of locally stationary processes. *Stochastic Processes and their Applications*, 62, 139-168.

[10] De Jong and N. Shepard, 1995, The simulation smoother for time series models. *Biometrika*, 82, 339-350.

[11] Di Tella, R. and R. MacCulloch, 2006, Europe vs America: institutional hysteresis in a simple normative model. *Journal of Public Economics*, 90, 2161-2186.

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

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

[14] Dufrénot, G., Guégan, D. and A. Péguin-Feissolle, 2008, Changing-regime volatility : A fractionally integrated SETAR model. *Applied Financial Economics* 18, 7-9, 519–526.

[15] Durbin, J. and S. Koopman, 2002, A simple and efficient simulation smoother for state space time series analysis. *Biometrika*, 89(3), 603-615.

[16] Gallant, A., Hsieh, D. and G. Tauchen, 1997, Estimation of stochastic volatility models with diagnoctics. *Journal of Econometrics*, 81(1), 159-192.

[17] Engelen, S., Norouzzadeh, P., Dullaert, W. and C. Rahmani, 2011, Multifractal features of spot rates in the liquid petroleum gas shipping market. *Energy Economics* 33,1,88-98.

[18] Gavrishchaka, V. and B. Ganguli, 2003, Volatility forecasting from multiscale and high-dimensional market data. *Neurocomputing* 55, 1-2, 285-305.

[19] Goldman, E. and H. Tsurumi, 2006, Bayesian Comparison of Long Memory and Threshold Nonlinearity in Time Series Models. 2006 World Meeting of the International Society for Bayesian Analysis, Alicante, Spain.

[20] Gordon, R., 1997, The time-varying NAIRU and its implications for economic policy. *Journal of Economic Perspectives*, 11(1), 11-32.

[21] Haldrup, N. and M.Ø. Nielsen, 2006a, A regime switching long memory model for electricity prices. *Journal of Econometrics* 135, 349-376.

[22] Haldrup, N. and M.Ø. Nielsen, 2006b, Directional congestion in a regime switching long memory model for electricity prices. Studies in Nonlinear Dynamics and Econometrics 10, 1-24.

[23] Kitagawa, G., 1996, Monte Carlo filter and smoother for non-Gaussian nonlinear state space models. *Journal of Computational and Graphical Statistics*, 5(1), 1-25.

[24] Koopman, S., Ooms, M. and M. Camero, 2007, Periodic seasonal ARFIMA-GARCH models for daily electricity spot prices. *Journal of the American Economic Association* 102, 16-27.

[25] Lahiani, A. and O. Scaillet, 2009, Testing for threshold effect in ARFIMA models: Application to US unemployment rate data. *International Journal of Forecasting* 25, 2, 418-428.

[26] Lavielle, M. and C. Ludena, 2000, The multiple change-points problem for the spectral distribution. *Bernouilli*, 6(5), 845-869.

[27] Leon-Ledesma, M.A., 2002, Unemployment hysteresis in the US states and the EU: a panel approach. *Bulletin of Economic Research*, 54, 95-103.

[28] Luukkonen, R., Saikkonen, P. and T. Teräsvirta, 1988, Testing linearity against smooth transition autoregressive models. *Biometrika* 75, 491-499.

[29] Mark, W.D., 1970, Spectral analysis of the convolution and filtering of non-stationary stochastic processes. *Journal of Sound and Vibration*, 11,19.

[30] Mitchell, W.F., 1993, Testing for unit roots and persistence in OECD unemployment. *Applied Economics*, 25, 1489-1501.

[31] Nelson, C.R. and C.I. Plosser, 1982, Trends and random walks in macroeconomic time series: some evidence and implications. *Journal of Monetary Economics*, 10, 139-162.

[32] Nielsen, M.Ø. and P.H. Frederiksen, 2005, Finite sample comparison of parametric, semiparametric and wavelet estimators of fractional integration. *Econometric Reviews* 24, 4, 405–443.

[33] Pasquini, M. and M. Serva, 1999, Multiscale behavior of volatility autocorrelations in a financial market. *Economics Letters* 65, 3, 275-279.

[34] Perron, P., 1988, Trends and random walks in macroeconomic time series: further evidence from a new approach. *Journal of Economic Dynamic and Control*, 12, 297-332.

[35] Philippe, A., Surgailis, D., Viano, M.C. ,2008, Time-varying fractionally integrated processes with nonstationary long memory. *Theory of Probability and Applications*, 52, 4, 651-673.

[36] Priestley, M.B., 1971, Some notes on the physical interpretation of spectra of non-stationary stochastic processes. *Journal of Sound and Vibration*, 17(1),

51-54.
[37] Rao, M.M., 1979, Covariance analysis of non-stationary time series, in Developments in statistics, edited by Krishnaiah, P.R.. Academic, New-York.
[38] Rees, F.L., 1972, Spectral relationships for non-stationary stochastic processes. *Journal of Sound and Vibration*, 23(1), 15-39.
[39] Surgailis, D. ,2008, Nonhomogeneous fractional integration and multifractional processes. *Stochastic Processes and their Applications* 118, 171–198.
[40] Tanaka, K., 1999, The nonstationary fractional unit root. *Econometric Theory* 15, 9, 549-582.
[41] Teräsvirta, T., 1994, Specification, estimation and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association* 89, 425, 208-218.
[42] Tsao, Y.H., 1984, Uncertain principle in frequency-time methods. *Journal of the Acoustical Society of America*, 75, 486-498.
[43] Vesilo, R.A. and A. Chan, 1996, Detecting change points in long range dependency traffic. In: Proceedings of the Australian Telecommunication Networks & Applications Conference, Melbourne, 3-6 December 1996, 567-572.
[44] Whitcher, B.J. and M.J. Jensen, 2000, Wavelet estimation of local long-memory parameter. *Exploration Geophysics*, 31, 94-103.