An exponential FISTAR model applied to the US real effective exchange rate

Mohamed Boutahar, Imene Mootamri, Anne Peguin-Feissolle

GREQAM

An exponential FISTAR model applied to the US real effective exchange rate

**Mohamed BOUTAHAR
Imène Mootamri
Anne Péguin-Feissolle**

June 2007

DT-GREQAM

An exponential FISTAR model applied to the US real effective exchange rate[*]

Imène Mootamri[†]   Mohamed Boutahar[‡]
Anne Péguin-Feissolle[§]

June 5, 2007

Abstract

The aim of this paper is to study the dynamics of the US real effective exchange rate by capturing nonlinearity and long memory features. In this context, we used the family of fractionally integrated STAR (FISTAR) models proposed by van Dijk, Franses and Paap (2002) to the case when the transition function is an exponential function and we develop an estimation procedure. Indeed, these models can take into account processes characterized by several distinct dynamic regimes and persistence phenomena.

Keywords: fractional integration, nonlinearity, STAR models, long memory, real effective exchange rate, forecasting.

JEL classification: C22, C51, C52, F31

---

[*]*Corresponding author:* Imène MOOTAMRI, GREQAM, Centre de la Charité, 2 rue de la Charité, 13236 Marseille cedex 02, FRANCE, tel: +33.4.91.14.07.70, fax: +33.4.91.90.02.27, Email: imene.mootamri@etumel.univmed.fr

[†]GREQAM, Université de la Méditerranée, France

[‡]GREQAM, Université de la Méditerranée, France

[§]GREQAM, CNRS, France

1 Introduction

Long memory processes have received considerable attention by researchers from very diverse fields. The books by Beran (1994), Doukhan, Oppenheim and Taqqu (2003) and Robinson (2003) provide updated surveys of recent developments on this topic. The long memory processes are characterized by a long-term dependence and the presence of cycles and level changes. They were detected in economy in many fields, for example in the dynamic of exchange rates or the volatility of financial time series. In addition, we assist for the few latest years to a significant development of nonlinear modelling. For instance, in economics and finance, multiple regimes modeling becomes more and more important in order to take into account phenomena characterized for instance by recession or expansion periods, or high or low volatility periods. Consequently, a variety of models has been proposed in the literature to account for this behavior, among which Markov switching models, smooth transition autoregressive (STAR) models. The nonlinearity property of economic time series can also be justified by the existence of asymmetry in variable's dynamics, such as favorable shocks have a more important and persistent effect than the unfavorable shocks. In order to consider these possible nonlinearities, it is necessary to have econometric models able to generate different dynamic according to the cycle phase.

Therefore, this paper belongs to a literature exploring simultaneously these two key properties of economic and financial time series, namely the long-memory and nonlinear properties. Indeed, a line of papers has recently proposed that we can call "nonlinear long-memory" models. For instance, some authors provide a joint evidence of mean reversion over long horizons and nonlinear dynamics on exchange rate markets, by generalizing to the nonlinear framework the Beveridge-Nelson decomposition (see, Clarida and Taylor (2001), Sarno and Taylor (2001)). Others propose new classes of long-memory models. For instance, Franses and Paap (2002), Franses, Van der Leij and Paap (2002) introduce CLEAR and Switching CLEAR processes, which show autocorrelation at high lags with an ACF that decays at a faster rate in the beginning in comparison to the ACF of an ARFIMA model.

Along this line of research, the fractionally integrated smooth transition autoregressive (FISTAR) models have also been proposed, that offer another potential application to economic and financial data (see van Dijk, Franses and Paap (2002) and Smallwood (2005)). Van Dick, Franses and Paap (2002) present the modeling cycle for specification of these models, such as testing for nonlinearity, parameter estimation and adequacy tests, in the case where the transition function is the logistic function; they study the dynamics of monthly US unemployment rates. Smallwood (2005) extends these results

2

to consider the FISTAR model with an exponential transition function, and applies this model to the purchasing power parity puzzle by considering the real exchange rate processes for twenty countries against the United States.

In this paper, we study this class of models because these FISTAR models indeed make it possible to generate nonlinearity, since they are defined by several distinct modes in dynamics, and to take into account of the persistence phenomenon. We consider the case of an exponential transition function and propose a two-step estimation method: in the first step, we estimate the long memory parameter, then, in the second step, the STAR model parameters via nonlinear least squares estimation.

The remainder of this paper is organized as follows. In Section 2, we present the FISTAR model with an exponential transition function and the two-step estimation procedure; we describe also the out-of-sample forecasting. In Section 3, we analyze the monthly US real effective exchange rate series in order to illustrate the various elements of the modelling cycle. Finally, Section 4 concludes.

2 The FISTAR model

2.1 Presentation of the model

Let us consider a process y_t that satisfies the following long memory scheme:

$$(1 - L)^d y_t = x_t \tag{1}$$

where x_t is a covariance-stationary process. The parameter d is possibly non-integer; in which case, the time series y_t is called fractionally integrated (FI) (see, among others, Granger and Joyeux (1980) and Hosking (1981)) and

$$
\begin{aligned}
(1 - L)^d & = 1 - dL - \frac{d(1-d)}{2!}L^2 - \frac{d(1-d)(2-d)}{3!}L^3 + ... \\
& = \sum_{j=0}^{\infty} \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)} L^j,
\end{aligned}
\tag{2}
$$

where $\Gamma(\cdot)$ is the Gamma function and L is the lag operator. If $-\frac{1}{2} < d < \frac{1}{2}$, the process is stationary and invertible, and d represents the degree of long memory behavior. For $0 < d < \frac{1}{2}$, (y_t) is a stationary long memory process in the sense that autocorrelations are not absolutely summable and decays hyperbolically to zero. Finally, if $d \succ \frac{1}{2}$, y_t is nonstationary and the shocks have permanent effects.

To capture the nonlinear feature of time series, a wide variety of models can be used (see Franses and van Dijk (2000)). The smooth transition autoregressive (STAR) model is one of the more popular; it has been empirically developed by Teräsvirta (1994) and is given by:

$$y_t = \left(\varphi_{10} + \sum_{i=1}^{p} \varphi_{1i} y_{t-i} \right) + \left(\varphi_{20} + \sum_{i=1}^{p} \varphi_{2i} y_{t-i} \right) F\left(s_t, \gamma, c\right) + \varepsilon_t \qquad (3)$$

where ε_t is a white noise process, γ is the transition parameter ($\gamma > 0$) and c is the threshold parameter; the transition variable, s_t, is generally the lagged endogenous variable. In most applications, the transition function $F\left(s_t, \gamma, c\right)$ is an exponential function or a logistic function.

Specifically, for the Exponential STAR (ESTAR)[1] family of models, the transition function takes the following form:

$$F(s_t, \gamma, c) = 1 - \exp\left(-\frac{\gamma}{\sigma_{s_t}^2} \left(s_t - c\right)^2 \right) \qquad (4)$$

where σ_{s_t} is the standard deviation of s_t. The exponential transition function is symmetric and the parameter γ controls the degree of nonlinearity. As $\gamma \to 0$ or $\gamma \to \infty$, the exponential transition function goes to zero or 1, such that the model in (3) becomes linear.

The logistic transition function is given by:

$$F(s_t, \gamma, c) = \left(1 + \exp\left(-\frac{\gamma}{\sigma_{s_t}} \left(s_t - c\right) \right) \right)^{-1}. \qquad (5)$$

The Logistic STAR (LSTAR)[2] model describes thus an asymmetric behavior. As $\gamma \to \infty$, the logistic transition function approaches an indicator function depending on $|s_t - c|$, and the model becomes a threshold autoregressive model. When $\gamma \to 0$, $F(s_t, \gamma, c)$ becomes equal to 0.5 such that the LSTAR model reduces to a linear model.

In this paper, we consider the fractionally integrated STAR (FISTAR) model introduced by van Dijk, Franses and Paap (2002) (see also Smallwood (2005)); it combines the two representations in (1) and (3) and is given by:

$$\begin{cases} (1 - L)^d y_t = x_t \\ x_t = \left(\varphi_{10} + \sum_{i=1}^{p} \varphi_{1i} x_{t-i} \right) + \left(\varphi_{20} + \sum_{i=1}^{p} \varphi_{2i} x_{t-i} \right) F\left(s_t, \gamma, c\right) + \varepsilon_t \end{cases} \qquad (6)$$

---

[1]Paya and Peel (2006), Michael, Nobay and Peel (1997), Taylor, Peel and Sarno (2001), and Sarantis (1999) applied the ESTAR models to exchange rates.

[2]Terasvirta and Anderson (1992), for instance, applied this model in order to characterize the different dynamics from the industrial production for some OECD countries during the expansion and recession phases.

where ε_t is a martingale difference sequence with $E\left[\varepsilon_t | \Omega_{t-1}\right] = 0$ and $E\left[\varepsilon_t^2 | \Omega_{t-1}\right] = \sigma^2$ and Ω_t is the information set available at time t. The FISTAR model can be also be written as follows:

$$
(1 - L)^d y_t = \begin{array}{l} \left(\pi_{10} + \sum_{j=1}^{\infty} \pi_{1,j} y_{t-j}\right) \\ + \left(\pi_{20} + \sum_{j=1}^{\infty} \pi_{2,j} y_{t-j}\right) F\left(s_t, \gamma, c\right) + \varepsilon_t, \end{array} \tag{7}
$$

where, $\pi_{i0} = \varphi_{i0}$ and $\pi_i\left(L\right) = \varphi_i\left(L\right)\left(1 - L\right)^d$ for $i = 1, 2$ and $j = 1, ..., p$. $F(s_t, \gamma, c)$ is the transition function governing the movement from one regime to another; the model will be called Logistic FISTAR (LFISTAR) when this function is a logistic and Exponential FISTAR (EFISTAR) when it is an exponential. The transition variable is a lagged value of y_t, i.e. $s_t = y_{t-m}$ with $m > 0$, where m is the delay parameter. The fractional parameter d and the autoregressive parameters make the FISTAR model potentially useful for capturing both nonlinear and long-memory features of the time series y_t. Indeed, as noted by van Dijk, Franses, and Paap (2002), the long-run properties of y_t are restricted to be constant and these are determined by the fractional differencing parameter; however, the short-run dynamics are determined by autoregressive parameters.

We are going to present the different steps of the specification procedure for FISTAR models, such as it is proposed by van Dijk, Franses, and Paap (2002):

- Specify a linear ARFI(p) model by selecting the autoregressive order p by means of information criteria (Akaike (1974) or Schwarz (1978)).

- Test the null hypothesis of linearity against the alternative of a FISTAR model. If linearity is rejected, select the appropriate transition variable.

- Estimate the parameters in the FISTAR model.

- Evaluate the model using misspecification tests (no remaining nonlinearity, parameter constancy, no residual autocorrelation, among others).

2.2   Linearity tests

Teräsvirta (1994) developed the procedure of testing linearity against STAR models; he pointed out that this procedure is complicated by the presence of unidentified nuisance parameters under the null hypothesis. To overcome this problem, he replaced the transition function $F(s_t, \gamma, c)$ by a suitable Taylor series approximation about $\gamma = 0$; in the reparametrized equation,

5

the identification problem is no longer present, and linearity can be tested by means of a Lagrange multiplier (LM) statistic. This procedure is extended to LFISTAR models by van Dijk, Franses, and Paap (2002) and to EFISTAR by Smallwood (2005).

Our empirical results show that the EFISTAR model is more appropriate for modelling real exchange rate dynamics than the LFISTAR model. Thus, we discuss only testing linearity against the EFISTAR model given by:$x_t = \left(\varphi_{10} + \sum_{i=1}^{p} \varphi_{1i} x_{t-i}\right) + \left(\varphi_{20} + \sum_{i=1}^{p} \varphi_{2i} x_{t-i}\right) F\left(s_t, \gamma, c\right) + \varepsilon_t$

$$
\begin{cases}
(1 - L)^d y_t = x_t \\
x_t = \pi_1' w_t + \pi_2' w_t F\left(s_t, \gamma, c\right) + \varepsilon_t \\
F(s_t, \gamma, c) = 1 - \exp\left(-\frac{\gamma}{\sigma_{s_t}^2} \left(s_t - c\right)^2\right)
\end{cases}
\tag{8}
$$

where $w_t = (1, x_{t-1}, \cdots\cdots, x_{t-p})'$, $\pi_i = \left(\pi_{i0}, \pi_{i1}, ..., \pi_{ip}\right)'$ for $i = 1, 2$ and ε_t is a martingale difference sequence with $E\left[\varepsilon_t | \Omega_{t-1}\right] = 0$ and $E\left[\varepsilon_t^2 | \Omega_{t-1}\right] = \sigma^2$.

The problem of the presence of unidentified nuisance parameters under the null hypothesis in FISTAR model is almost the same in STAR model. Thus, the first order Taylor expansion of the exponential transition function around $\gamma = 0$ is given by:

$$
\begin{aligned}
T_1\left(s_t, \gamma, c\right) &= \left. F(s_t, \gamma, c)\right|_{\gamma=0} + \gamma \left.\frac{\partial F(s_t, \gamma, c)}{\partial \gamma}\right|_{\gamma=0} + R(s_t, \gamma, c) \tag{9} \\
&= \frac{\gamma}{\sigma_{s_t}^2} \left(s_t - c\right)^2 + R(s_t, \gamma, c)
\end{aligned}
$$

where $R(s_t, \gamma, c)$ is a remainder term. Substituting $T_1\left(s_t, \gamma, c\right)$ for $F(s_t, \gamma, c)$ in the second relationship of (8) and reparametrizing terms, the auxiliary regression is given by:

$$
x_t = \theta_0' w_t + \theta_1' w_t s_t + \theta_2' w_t s_t^2 + e_t,
\tag{10}
$$

where $\theta_0 = \pi_1 + \frac{\pi_2 \gamma c^2}{\sigma_{s_t}^2}$, $\theta_1 = -2\frac{\pi_2 \gamma c}{\sigma_{s_t}^2}$, $\theta_2 = \frac{\pi_2 \gamma}{\sigma_{s_t}^2}$, and $e_t = \varepsilon_t + \beta w_t R(s_t, \gamma, c)$. Under $H_0 : \pi_2 = 0$ *or* $\gamma = 0$, we have $R(s_t, \gamma, c) = 0$ and $e_t = \varepsilon_t$. Therefore, the null hypothesis for linearity is given by: $H_0 : \theta_1 = \theta_2 = 0$.

The existence of the fractional differencing parameter, however, complicates the construction of the LM type test statistic. Following van Dijk, Franses, and Paap (2002), under the assumption that $e_t \sim N\left(0, \sigma^2\right)$, the conditional log-likelihood for observation t is given by:

$$
l_t = -\frac{1}{2} \ln 2\pi - \frac{1}{2} \ln \sigma^2 - \frac{e_t^2}{2\sigma^2}.
\tag{11}
$$

6

The remaining partial derivatives evaluated under the null hypothesis H_0 are given by:

$$\frac{\partial l_t}{\partial \theta_i} = \frac{1}{\sigma^2}\widehat{\varepsilon}_t w_t s_t^i, \ i = 0, 1, 2$$

$$\frac{\partial l_t}{\partial d} = \frac{1}{\sigma^2}e_t\frac{\partial e_t}{\partial d} = -\frac{\widehat{\varepsilon}_t}{\sigma^2}\sum_{j=1}^{t-1}\frac{\widehat{\varepsilon}_{t-j}}{j},$$

where $\widehat{\varepsilon}_t$ are the residuals obtained from the ARFI model under the null hypothesis. Therefore, the LM-type test statistic can be computed in a few steps as follows:

- Estimate an ARFI(p), obtain the set of residuals $\widehat{\varepsilon}_t$. The sum of squared errors, denoted SSR_0, is then constructed from the residuals $\widehat{\varepsilon}_t$, $SSR_0 = \sum_{t=1}^{T}\widehat{\varepsilon}_t^2$.

- Regress $\widehat{\varepsilon}_t$ on w_t, $-\sum_{j=1}^{t-1}\frac{\widehat{\varepsilon}_{t-j}}{j}$ and $w_t s_t^i$, $i = 1, 2$, and compute the sum of squared residuals SSR_1 under the alternative hypothesis.

- The χ^2 version of the LM test statistic is calculated as:

$$LM_{\chi^2} = \frac{T(SSR_0 - SSR_1)}{SSR_0} \tag{12}$$

and is distributed as $\chi^2(2(p+1))$ under H_0 (T denotes the sample size). The Fisher version[3] of the LM test statistic is calculated as:

$$LM_F = \frac{(SSR_0 - SSR_1)/2(p+1)}{SSR_1/(T - 3(p+1))} \tag{13}$$

and is distributed as an $F(2(p+1), T - 3(p+1))$ statistic under H_0.

2.3 Estimation of the FISTAR model

It is important to obtain a consistent estimate of the long memory parameter d because the test statistics for the FISTAR model depend on this estimated value. In this section, we present two approaches for estimation of the parameters in the FISTAR model: in the first one, we estimate all the parameters simultaneously (as proposed by van Dijk, Franses, and Paap (2002)), while the second method consists in performing the estimation in two steps.

---

[3]The Fisher version is preferred than χ^2 when the sample size is small and that the selected delay is important.

2.3.1 Simultaneous estimation

To estimate the parameters of the FISTAR model, van Dijk, Franses, and Paap (2002) modify "Beran's (1995) approximate maximum likelihood (AML) estimator for invertible and possibly nonstationary ARFIMA models to allow for regime switching autoregressive dynamics". This estimator minimizes the sum of squared residuals of the FISTAR model as follows:

$$S(\lambda) = \sum_{t=1}^{T} \varepsilon_t^2(\lambda), \tag{14}$$

where $\lambda = (\pi_1', \pi_2', d, \gamma, c)$ denotes the parameters of the FISTAR model (7). The residuals $\varepsilon_t(\lambda)$ are calculated as follows:

$$
\begin{aligned}
\varepsilon_t(\lambda) = \ & (1-L)^d y_t - \left(\pi_{10} + \sum_{j=1}^{t+p-1} \pi_{1,j} y_{t-j}\right) \\
& - \left(\pi_{20} + \sum_{j=1}^{t+p-1} \pi_{2,j} y_{t-j}\right) F(s_t, \gamma, c)
\end{aligned}
\tag{15}
$$

with $F(s_t, \gamma, c) = 1 - \exp\left(-\frac{\gamma}{\sigma_{s_t}^2}(s_t - c)^2\right)$. Thus, conditional upon d, γ and c, van Dijk, Franses, and Paap (2002) remark that the FISTAR model is linear in the remaining parameters; estimates of π_1 and π_2 can be thus obtained by ordinary least squares as:

$$\widehat{\mu}(d, \gamma, c)' = \left(\sum_{t=1}^{T} w_t(d, \gamma, c) \, w_t(d, \gamma, c)'\right)^{-1} \left(\sum_{t=1}^{T} w_t(d, \gamma, c) \, y_t\right), \tag{16}$$

where $w_t(d, \gamma, c) = (w_t', w_t' F(s_t, \gamma, c))'$. Therefore, the sum of squares function can be obtained by :

$$S(d, \gamma, c) = \sum_{t=1}^{T} \left(y_t - \widehat{\mu}(d, \gamma, c)' w_t(d, \gamma, c)\right)^2. \tag{17}$$

According to van Dijk, Franses, and Paap (2002), it can be difficult to estimate the model parameters jointly. In particular, accurate estimation of the smoothness parameter γ is quite difficult when this parameter is large. They proposed an algorithm that is based on a grid search over d, γ and c in order to obtain starting values for the non-linear least squares procedure.

2.3.2 Two steps estimation

The properties of the process y_t depend on the value of the parameter d. Many researchers have proposed methods for estimating the long memory

parameter d. These methods can be summarized in three classes: the heuristic methods (Hurst (1951), Higuchi (1988), Lo (1991)...), the semiparametric methods (Geweke and Porter-Hudak (1983), Robinson (1994, 1995a and b), Lobato and Robinson (1996)...) and the maximum likelihood methods (Whittle (1951), Sowell (1992)...). In the first two classes, we can estimate only the long memory parameter d. However, to fit an ARFIMA (p, d, q) model, we need two steps: one filters out the long memory component and then fits an ARMA (p, q) model to the residual series. In the last class, we estimate simultaneously all the parameters.

The estimation method of the FISTAR model we propose proceeds in two steps:

- In the first step, we estimate the long memory parameter d in the simple model

$$(1 - L)^d y_t = x_t, \tag{18}$$

using a standard estimation method, such as the Geweke and Porter-Hudak (1983).

- Once we obtained $\widehat{d}_{GPH}$, in the second step, we filter out the long memory component and we estimate the STAR model parameters via nonlinear least squares estimation.

2.4 Misspecification tests of FISTAR models

Eitrheim and Teräsvirta (1996) develop misspecification tests (residual autocorrelation, remaining nonlinearity and parameter constancy) for the STAR model. Van Dijk, Franses and Paap (2002) modify these tests to obtain similar misspecification tests for the FISTAR model.

- **Testing the hypothesis of no residual autocorrelation.** Considering the FISTAR model in (7), the null hypothesis of no autocorrelation in ε_t can be tested against the alternative of serial dependence up to order q, given by:

$$\varepsilon_t = \rho_1 \varepsilon_{t-1} + \cdots + \rho_q \varepsilon_{t-q} + e_t, \ e_t \sim iid \left(0, \sigma^2\right). \tag{19}$$

Thus, the null hypothesis, given by $\rho_1 = \cdots = \rho_q = 0$, can be tested by LM test, denoted as $LM_{SI}(q)$, that is equal to TR^2, where R^2 is the coefficient of determination. The $LM_{SI}(q)$ test, is asymptotically χ^2 distributed with q degrees of freedom.

- **Testing the hypothesis of no remaining nonlinearity.** The null hypothesis can be tested against the additive third regime FISTAR model, where $F_2(s_t, \gamma_2, c_2)$ is a logistic function. We apply Van Dijk's approach to the case where $F_2(s_t, \gamma_2, c_2)$ is an exponential function i.e. to model given by:

$$
\begin{cases}
(1-L)^d y_t = x_t \\
x_t = \pi_1' w_t + \pi_2' w_t F_1(s_t, \gamma_1, c_1) + \pi_3' w_t F_2(s_t, \gamma_2, c_2) + e_t, \\
F_i(s_t, \gamma_i, c_i) = 1 - \exp\left(-\frac{\gamma_i}{\sigma_{s_t}^2}(s_t - c_i)^2\right), \quad i = 1, 2.
\end{cases}
\tag{20}
$$

The null hypothesis of no remaining non linearity is given by $\gamma_2 = 0$. This test suffers from a similar identification problem as encountered in testing linearity against a FISTAR model; the problem is solved by replacing again $F_2(s_t, \gamma_2, c_2)$ by a suitable Taylor expansion around $\gamma_2 = 0$.

- **Testing the hypothesis of parameter constancy.** The null hypothesis of the test of parameter constancy in the FISTAR model against the alternative of smoothly changing parameters is based on:

$$
\begin{cases}
x_t = \varphi_{1t}' w_t (1 - F_1(s_t, \gamma_1, c_1)) + \varphi_{2t}' w_t F_1(s_t, \gamma_1, c_1) + e_t \\
\varphi_{1t} = (1 - F_2(s_{2t}, \gamma_2, c_2))\varphi_1 + F_2(s_{2t}, \gamma_2, c_2)\varphi_3 \\
\varphi_{2t} = (1 - F_2(s_{2t}, \gamma_2, c_2))\varphi_2 + F_2(s_{2t}, \gamma_2, c_2)\varphi_4
\end{cases}
\tag{21}
$$

Lin and Teräsvirta (1994) and Lundberg and Teräsvirta (2001) supposed that the parameters have the same transition function as the endogenous variable, where the transition variable is $s_{2t} = t$. The identification problem of the parameters is solved by a suitable Taylor approximation of $F_2(t, \gamma_2, c_2)$. The LM type test statistics is denoted as LM_c.

2.5   Out-of-sample forecasting performance

Unlike the linear model, forecasting with nonlinear models is more complicated, especially for several steps ahead (see, for instance, Granger and Teräsvirta (1993)). Let us consider the EFISTAR model (8) which can be written as:

$$
\begin{cases}
(1-L)^d y_t = x_t \\
x_t = G(w_t, \omega) + \varepsilon_t \\
F(s_t, \gamma, c) = 1 - \exp\left(-\frac{\gamma}{\sigma_{s_t}^2}(s_t - c)^2\right)
\end{cases}
\tag{22}
$$

10

where $G(w_t, \omega) = \pi'_1 w_t + \pi'_2 w_t F(s_t, \gamma, c)$ and $\omega = (\pi'_1, \pi'_2, \gamma, c)'$. The optimal one-step ahead forecast of x_t is given by:

$$x_{t+1|t} = E(x_{t+1}|\Omega_t) = G(w_{t+1}, \omega); \qquad (23)$$

this forecast can be achieved with no difficulty and can be estimated by

$$\widehat{x}_{t+1|t} = G(w_{t+1}, \widehat{\omega}) \qquad (24)$$

where $\widehat{\omega}$ is the parameter estimate. However, when the forecast horizon is larger than one period, things become more complicated because the dimension of the integral grows with the forecast horizon. For example, the two-step ahead forecast of x_t is given by:

$$\widehat{x}_{t+2|t} = E\left(G\left(\widehat{w}_{t+2|t}, \omega\right)|\Omega_t\right) = \int_{-\infty}^{\infty} G\left(\widehat{w}_{t+2|t}, \widehat{\omega}\right) f(\varepsilon)d\varepsilon \qquad (25)$$

with $\widehat{w}_{t+2/t} = \left(1, \widehat{x}_{t+1|t} + \varepsilon_{t+1}, x_t, ..., x_{t+2-p}\right)'$. The analytic expression for the integral (25) is not available. We thus need to approximate it using integration techniques. Several methods obtaining forecasts to avoid numerical integration have been developed (see Granger and Teräsvirta (1993)). In this paper, we use a bootstrap method suggested by Lundbergh and Teräsvirta (2001). This approach is based on the approximation of $E\left(G\left(\widehat{w}_{t+2|t}, \omega\right)|\Omega_t\right)$; the optimal point forecast is given by :

$$\widehat{x}_{t+2|t} = \frac{1}{k}\sum_{i=1}^{k} G\left(\widehat{w}_{t+2|t}^{(i)}, \widehat{\omega}\right), \qquad (26)$$

where k is some large number and the values of ε_{t+1} in $\widehat{w}_{t+2|t}^{(i)}$ are drawn with replacement from the residuals from the estimated model $\widehat{\varepsilon}_t$.

In general, forecasts are evaluated using the mean squared prediction error (MSPE) and the root mean squared prediction (RMSE), where m is the number of steps-ahead forecasts. Models with smaller MSPE have a better forecast performance. Further, in order to assess the accuracy of forecasts derived from two different models, we employ the Diebold and Mariano (1995) test:

$$DM = \frac{\bar{l}}{\sqrt{\frac{2\pi\widehat{f}_l(0)}{T}}}, \qquad (27)$$

where $\bar{l}$ is an average (over T observations) of a general loss differential function and $\widehat{f}_l(0)$ is a consistent estimate of the spectral density of the loss differential function at frequency zero. They show that the DM statistic is distributed as standard normal under the null hypothesis of equal forecast accuracy.

11

3 Empirical results

The fractionally integrated models[4] have been already applied in economics and finance, for instance to exchange rates (Diebold, Husted and Rush (1991), Cheung and Lai (2001), Baillie and Bollerslev (1994)), inflation (Hassler and Wolters (1995), Baillie, Chung and Tieslau (1996)) and unemployment modeling (Diebold and Rudebusch (1989), Tschernig and Zimmermann (1992), Koustas and Veloce (1996), Crato and Rothman (1996)). Therefore, the long memory models, such as the FISTAR, are not only able to study the persistence but also to capture nonlinearity features such as thresholds or asymmetries. They can be applied in various economic and financial fields, in particular the stock indexes, the exchange rates and the interest rates. Van Dijk, Franses and Paap (2002) apply the FISTAR models to US unemployment and Smallwood (2005) to the case of purchasing power parity. In this paper, we study the behavior of exchange rates and compare the forecast performances of the EFISTAR modeling compared to other more used models.

3.1 The data

We use monthly data of the US real effective exchange rate covering the period June 1978 until April 2002; these data were obtained from the *IMF International Financial Statistics*. The series is expressed in logarithm. The use of monthly data provides us with a reasonably large sample and hence meets the requirement of the linearity tests for many degrees of freedom. The series is shown in Figure 1, which demonstrates a real appreciation of the dollar during the beginning of the 1980's followed by depreciation in 1985. As noted by Smallwood (2005), consistently with the theoretical foundation of Sercu, Uppal, and van Hulle (1995), we observe four periods after 1987 in which the dollar steadily appreciates and then rapidly depreciates after reaching approximately the same value. This provides some support for the use of nonlinear models.

3.2 Linearity tests results

Application of the linearity tests models requires stationary time series. The unit root tests of Phillips and Perron (PP), Kwiatkowski, Phillips, Schmidt

---

[4]For a survey on long memory models and their applications in economics and finance, see Baillie (1996).

and Shin (KPSS)[5] and Dickey-Fuller Augmented (ADF) for the level and first difference of the real effective exchange rates, measured in logarithms, are shown in Table 1. These results indicate that the time series are integrated of order 1 at both 5% and 1% significance levels.

The selection of the maximum lag p, of the linear ARFI model was made using the AIC and BIC criteria under the non autocorrelation hypothesis. We allow for a maximum autoregressive order of $p = 6$. Both AIC and BIC indicate that an ARFI model with $p = 4$ is adequate.

The linearity tests are displayed in Table 2. In carrying out linearity tests, we have considered values for the delay parameter m over the range $[1, 6]$, and calculated the $p - values$ for the linearity test in each case; the estimate of m is chosen by the lowest $p - value$. Using 5% as a threshold $p - value$, the test classifies the US real effective exchange rates as nonlinear. Although the $p - value$ is slightly higher than 5%, we show thereafter that a nonlinear model describes the features of a time series better than a linear model[6] . Then the lowest $p - value$ corresponds to $m = 4$ $(m \leq p)$.

3.3 Estimation results

Estimation results for the ARFI and EFISTAR models are shown in table 3. The second column gives the ARFI model estimation; the estimate of d is 0.941, showing that the long memory model is nonstationary. The results of the second model are based on the specification (8) where y_t is the first difference of the US real effective exchange rates. The third column of Table 3 contains simultaneous estimation results of the parameters. In particular, the estimate of d is equal to -0.169 and belongs thus to the interval $\left]-\frac{1}{2}, 0\right[$, suggesting that the process is stationary and invertible. The autocorrelation function decreases quickly than the case where $0 < d < \frac{1}{2}$: (y_t) is an anti-persistent process. It is also interesting to note in the last column corresponding to the two-step estimation that the degree of persistence measured by the differentiation parameter increases. If the GPH estimatior is used, $\left(\widehat{d}_{GPH} = 0.218\right)$, then the process is stationary and invertible. The autocorrelation function decays hyperbolically to zero, thus y_t is a long memory process. The ratio of the standard errors for the nonlinear and linear models for the simultaneous estimation of the EFISTAR model is equal to 0.840; it's higher than for the two-step estimation (0.670). We can thus confirm that the nonlinear model improves modelling of the exchange rate process, as shown

---

[5]Contrary to ADF test, the KPSS test considers the stationnarity under the null hypothesis, and the alternative hypothesis is the presence of unit root.

[6]This result is also found in Sarantis (1999).

by both estimation methods. It is worthwhile noting here the relative small value of the estimation of γ for the second estimation (2.547 compared to 12.655 for simultaneous estimation), suggesting that the transition from one regime to the other is rather slow, contrary to first estimation which assumes a slightly sharp switch. The parameter c indicates the halfway point between the different phases of the exchange rate. The value of c is negative for the first case, and not significantly different from zero in the other. These values belong to the neighborhood of the sample mean for the exchange rates. Figures 2 and 3 show the curves of the exponential transition function corresponding to the estimation of the EFISTAR model, the first one using the simultaneous estimation method and the second one the two-step method.

Table 4 gives summary statistics and misspecification tests for ARFI and FISTAR models. In particular the $p-values$ (using 5% as the threshold) reject serial correlation, the presence of ARCH, and accept normality in the residuals for both models. From the skewness and kurtosis of the series, it is evident that the US real effective exchange rate is symmetric and the frequency curve is normal, this is confirmed by the Jarque-Bera test for normality. Moreover, the test of parameter constancy against the alternative of smoothly changing parameters for $s_t = t$, give fine results, and LM test of no remaining non linearity is accepted for FISTAR model.

3.4 Forecasting performance

In this section, we evaluate the performance of EFISTAR model estimated by the two-step estimation method, by comparing its out-of-sample forecasts with the ARFI model. The final two years of data from January 2002 to April 2004 for US real effective exchange rate are used to evaluate the forecast performance of the estimated linear ARFI and EFISTAR models. To obtain the forecasts from nonlinear model, we use the method exposed in section 2.5; the forecast accuracy is evaluated using mean squared prediction error (MPSE) criterion. Further, in order to assess the accuracy of forecasts derived from two different models, we employ the Diebold and Mariano (1995) test statistic for which the null hypothesis is the hypothesis of equal accuracy of different predictive methods.

The results successfully provide evidence in favour of the predictive superiority of the EFISTAR model against the ARFI model using MPSE: the MPSE of the linear model (0.0008451) is actually greater than the MPSE of the nonlinear model (0.0004746). The statistical significance of this result is confirmed executing the Diebold and Mariano test: the Diebold and Mariano test statistic ($p-value = 0.0412$) indicates a statistically significant difference in predictive accuracy for the EFISTAR model over the ARFI model

specifications. We can thus conclude that the forecasts of the EFISTAR modeling are significantly better than those of the ARFI model.

4 Conclusion

The aim of this paper was to study the dynamic modelling of the US real effective exchange rates covering the period June 1978 until April 2002. We considered the FISTAR model, as proposed in van Dijk, Franses and Paap (2002), that can describe long memory and nonlinearity simultaneously and be used to produce out-of-sample forecasts. We used their model to the case of an exponential transition function. To this end, we employ two modelling approaches corresponding to two different estimations (simultaneous estimation or two-step estimation) of a EFISTAR model. The estimated EFISTAR models seem to provide a satisfactory description of the nonlinearity and persistency found in the US real effective exchange rates. With regards to the out-of-sample forecasting performance for US exchange rate, the tests for comparing the predictive accuracy show that the EFISTAR model seems better that the linear model.

References

[1] Akaike, H. (1974), "A new look at the statistical model identification", *IEEE Transactions on Automatic Control* 19, 6, 716–723.

[2] Baillie, R.T. (1996), "Long memory processes and fractional integration in econometrics", *Journal of Econometrics* 73, 5-59.

[3] Baillie, R.T., and T. Bollerslev (1994), "Cointegration, fractional cointegration, and exchange rate dynamics", *Journal of Finance* 49, 737-45.

[4] Baillie, R., Chung, C.F., and M. Tieslau (1996), "Analysing inflation by the fractionally integrated ARFIMA-GARCH model," *Journal of Applied Econometrics* 11, 23-40.

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

[6] Cheung, Y.W., and K.S. Lai (2001), "Long memory and nonlinear mean reversion in Japanese yen-based real exchange rates," *Journal of International Money and Finance* 20, 115-132.

[7] Clarida, R. H. and M. P. Taylor (2001), "Nonlinear permanent-temporary decomposition, with applications in macroeconomics and finance", Working Paper, University of Warwick.

[8] Crato, N. and P. Rothman (1996), "Measuring hysteresis in unemployment rates with long memory models", unpublished manuscript, East Carolina University.

[9] Diebold, F.X., Husted, S., and M. Rush (1991), "Real exchange rates under the Gold Standard," *Journal of Political Economy* 99:06, 1252-71.

[10] Diebold, F.X., and R.S. Mariano (1995), "Comparing predictive accuracy", *Journal of Business and Economic Statistics* 13: 253–263.

[11] Diebold, F.X., and G. Rudebusch (1989), "Long memory and persistence in aggregate output," *Journal of Monetary Economics* 24, 189-209.

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

[13] Eitrheim, O. and T. Teräsvirta (1996), "Testing the adequacy of smooth transition autoregressive models", *Journal of Econometrics* 74, 59-75.

[14] Franses, P. H. and R. Paap (2002), "Censored latent effects autoregression, with an application to US unemployment", *Journal of Applied Econometrics* 17, 347-366.

[15] Franses, P.H.,van Der Leij, M., and R. Paap (2002), "Modelling and forecasting level shifts in absolute returns", *Journal of Applied Econometrics* 17, 601-616.

[16] Franses, P.H., and D. van Dijk (2000), *Nonlinear Time Series Models in Empirical Finance*, Cambridge University Press: Cambridge.

[17] Geweke, J., and S. Porter-Hudak (1983), "The estimation and application of long-memory time series models", *Journal of Time Series Analysis* 4, 221-238.

[18] Granger, C.W.J., and R. Joyeux (1980), "An introduction to long-range time series models and fractional differencing", *Journal of Time Series Analysis* 1: 15–29.

[19] Granger, C.W.J., and T. Teräsvirta (1993), *Modelling Nonlinear Economic Relationships*, Oxford University Press: Oxford.

[20] Hassler, U., and J. Wolters (1995), "Long memory in inflation rates: international evidence," *Journal of Business and Economic Statistics* 13, 37-45.

[21] Higuchi, T. (1988), "Approach to an irregular time series on the basis of the fractal theory", *Physica D* 31, 277-283.

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

[23] Hurst, H.E. (1951), "Long-term storage capacity of reservoirs", *Transactions of the American Society of Civil Engineers* 116, 770-799.

[24] Koustas, Z., and W. Veloce (1996), "Unemployment hysteresis in Canada: an approach based on long-memory time series models", *Applied Economics* 28, 823-831.

[25] Lin, C.J., and T. Teräsvirta (1994), "Testing the constancy of regression parameters against continuous structural change", *Journal of Econometrics* 62, 211-228.

[26] Lo, A. (1991), "Long-term memory in stock market prices", *Econometrica* 59, 1279-1313.

[27] Lobato, I. and P. M. Robinson (1996), "Averaged periodogram estimation of long memory", *Journal of Econometrics* 73, 303-324.

[28] Lundbergh, S., and T. Teräsvirta (2001) "Forecasting with smooth transition autoregressive models'. In *Companion to Economic Forecasting*, Clements M.P., Hendry D.F. (eds). Basil Blackwell: Oxford.

[29] Michael, P., Nobay, A.R. and D.A. Peel (1997), "Transaction costs and nonlinear adjustment in real exchange rates: an empirical investigation", *Journal of Political Economy* 105, 803-879.

[30] Paya, I., and D.A. Peel (2006), "Temporal aggregation of an ESTAR process: some implications for purchasing power parity adjustment", *Journal of Applied Econometrics* 21, 5, 655-668.

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

[32] Robinson, P.M. (1994), "Semiparametric analysis of long-memory time series", *Annals of Statistics* 22, 515-539.

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

[34] Robinson, P.M. (1995b), "Gaussian semiparametric estimation of long range dependence", *The Annals of Statistics* 23, 1630-1661.

[35] Sarantis, N. (1999), "Modeling non-linearities in real effective exchange rates", *Journal of International Money and Finance* 18: 27-45.

[36] Sarno, L., and M. Taylor (2001), "Real exchange rate dynamics in transition economies: a nonlinear analysis", *Studies in Nonlinear Dynamics and Econometrics* 5, 153-177.

[37] Schwarz, G. (1978), "Estimating the dimension of a model", *Annals of Statistics* 6, 461-464.

[38] Sercu, P., Uppal, R., and C. Van Hulle (1995), "The exchange rate in the presence of transactions costs: implications for tests of purchasing power parity," *Journal of Finance* 50, 1309-1319.

[39] 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.

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

[41] Taqqu, M. S. and Teverovsky, V. (1998), "On estimating the intensity of long-range dependence in finite and infinite variance time series", *A Practical Guide To Heavy Tails: Statistical Techniques and Applications*, Boston, 177-217.

[42] Taylor, M.P., Peel, D.A. and L. Sarno (2001), "Nonlinear mean-reversion in real exchange rates : toward a solution to the purchasing power parity puzzles", *International Economic Review* 4, 1015-1041.

[43] Teräsvirta, T. (1994), "Specification, estimation, and evaluation of smooth transition autoregressive models", *Journal of the American Statistical Association* 89: 208–218.

[44] Teräsvirta, T., and H.M. Anderson (1992), "Characterizing nonlinearities in business cycles using smooth transition autoregressive models", *Journal of Applied Econometrics* 119-136.

[45] Tschernig, R., and K.F. Zimmermann (1992), "Illusive persistence in German unemployment", *Recherches Economiques de Louvain* 58, 441-453.

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

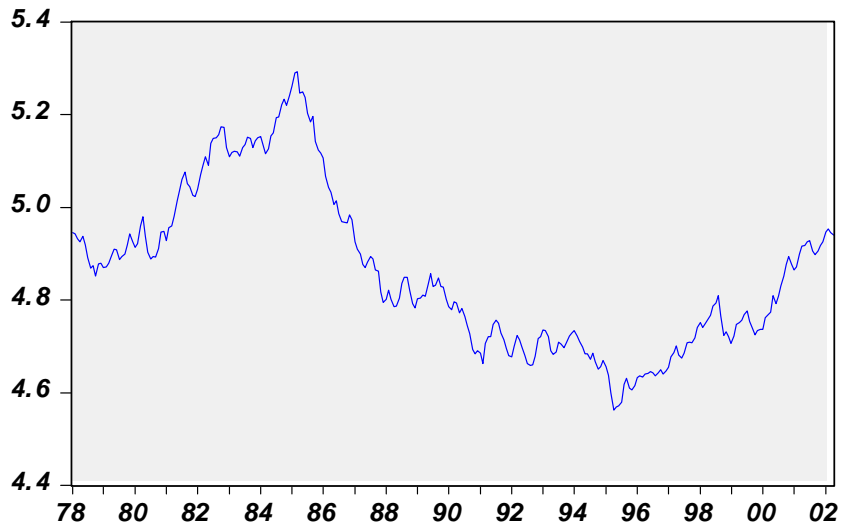[47] Whittle, P. (1951), *Hypothesis testing in time series analysis*, Hafner, New York.
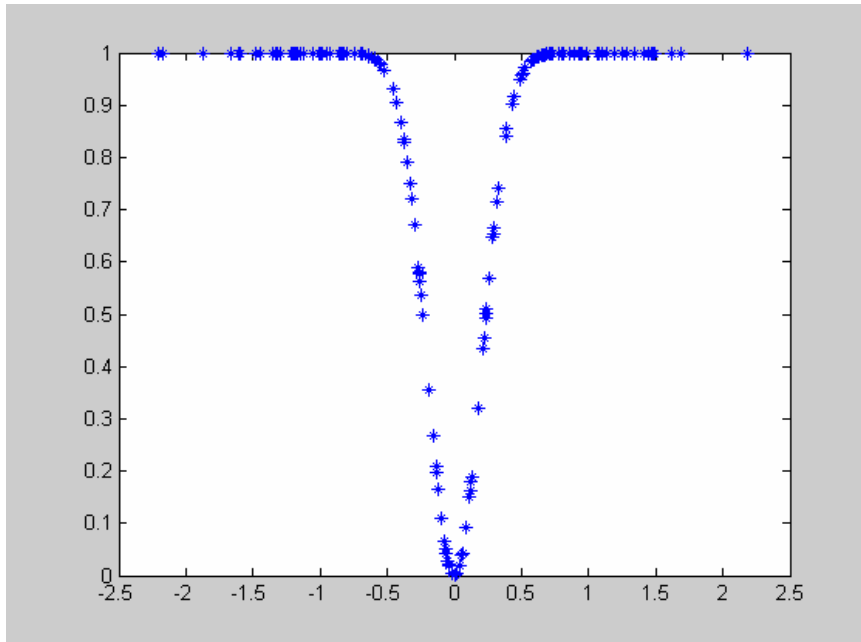
Figure 1: Monthly US real effective exchange rate (Log)

Figure 2: Exponential transition function (simultaneous estimation)
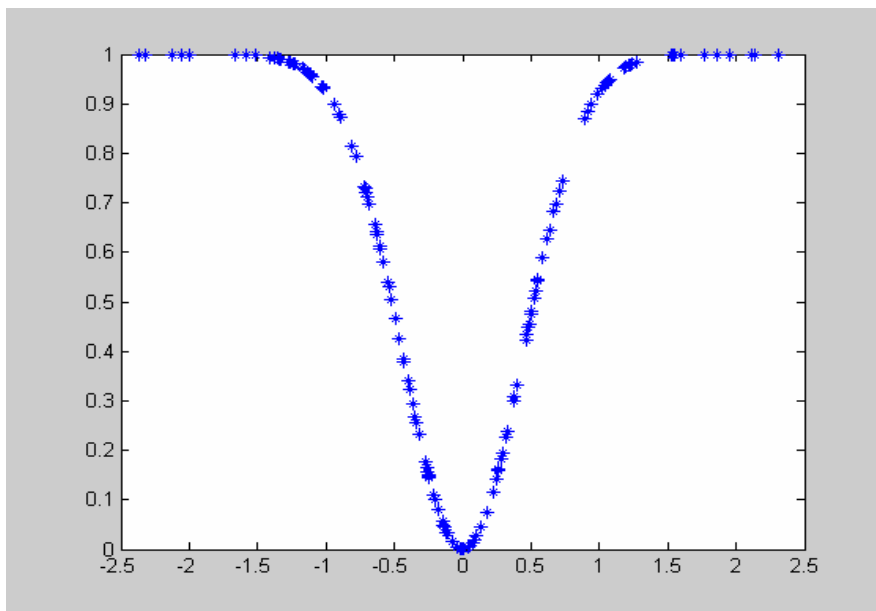
Figure 3: Exponential transition function (two-step estimation)

21

Table 1. Unit root tests

	Level	First difference
ADF	-1.118	-7.287
PP	-1.106	-12.281
KPSS	3.090	0.251

Note: The unit root tests are Phillips and Perron (PP), Kwiatkowski, Phillips, Schmidt and Shin (KPSS) and Dickey-Fuller Augmented (ADF) tests. For ADF test, the 1%, and 5% critical values are -3.455 and -2.871, respectively. For KPSS test, the 1%, and 5% critical values are 0.739 and 0.463, respectively.

Table 2. Linearity tests ($p - values$)

m	1	2	3	4	5	6
LM-test	0.868	0.346	0.087	**0.073**	0.251	0.171

Table 3. Estimation of the different models

	ARFI	EFISTAR (simultaneous estimation)	EFISTAR (two-step estimation)
π_{10}	-0.000 (0.001)	-0.063 (0.026)	-0.003 (0.014)
π_{11}	0.346 (0.059)	0.665 (0.295)	-0.175 (0.142)
π_{12}	-0.136 (0.062)	-0.167 (0.335)	0.156 (0.144)
π_{13}	0.082 (0.062)	0.194 (0.394)	0.345 (0.163)
π_{14}	-0.049 (0.059)	-0.683 (0.290)	0.217 (0.126)
π_{20}		-0.001 (0.001)	-0.004 (0.015)
π_{21}		1.256 (0.078)	0.470 (0.115)
π_{22}		-0.458 (0.122)	-0.238 (0.120)
π_{23}		0.172 (0.119)	0.121 (0.103)
π_{24}		0.035 (0.075)	-0.195 (0.121)
d	0.941 (0.134)	-0.169 (0.007)	0.218* (0.200)
γ		12.655 (8.648)	2.574 (1.190)
c		-0.101 (0.003)	0.022 (0.020)
S_E		0.840	0.670

Note: The standard errors are displayed in parentheses. * : GPH estimator. S_E is the ratio of residual variance for the nonlinear and linear models.

Table 4. Diagnostic tests

	ARFI	FISTAR
AIC	-8.195	-8.181
BIC	-7.846	-0.1327
SK	-0.166	-0.1327
K_r	3.297	3.006
JB	1.313 (0.518)	0.463 (0.793)
ARCH(1)	0.981 (0.321)	0.714 (0.398)
ARCH(2)	1.778 (0.411)	1.292 (0.524)
ARCH(3)	5.634 (0.130)	2.933 (0.402)
ARCH(4)	7.605 (0.107)	4.276 (0.370)
$LM_{SI}(2)$	0.765 (0.467)	1.764 (0.175)
$LM_{SI}(4)$	1.174 (0.325)	2.179 (0.075)
$LM_{SI}(6)$	1.280 (0.271)	2.111 (0.057)
$LM_{SI}(8)$	1.118 (0.355)	1.690 (0.106)
$LM_{SI}(31)$	0.746 (0.817)	0.965 (0.529)
LM_{NL}	-	0.937 (0.521)
LM_C	-	0.701 (0.778)

Note: The table presents selected diagnostic and misspecification tests statistics for the estimated FISTAR and ARFI models for the US real effective exchange rate; the numbers in parentheses are p-values. SK is skewness, K_r is kurtosis, JB is the Jarque–Bera test of normality of the residuals, ARCH(r) is the LM test of no autoregressive conditional heteroscedasticity up to order r, $LM_{SI}(q)$ denotes the LM test of no serial correlation in the residuals up to order q, LM_{NL} is the LM test of no remaining non linearity, and LM_C is the LM test of parameter constancy.