REGULAR ARTICLE

A new unit root test against ESTAR based on a class of modified statistics

Robinson Kruse

Abstract   This paper proposes a new unit root test against a nonlinear exponential smooth transition autoregressive model. This model receives much attention in international macroeconomics as it has been successfully applied to a variety of financial time series. The new test is build upon the nonstandard testing approach of Abadir and Distaso (J Econom 140:695–718, 2007) who introduce a class of modified statistics for testing joint hypotheses when one of the alternatives is one-sided. The asymptotic properties of the suggested unit root test are derived. In a Monte Carlo study the popular Dickey–Fuller-type test proposed by Kapetanios et al. (J Econom 112:359–379, 2003) is compared to the new test. The results suggest that the new test is generally superior in terms of power. An application to a real effective exchange rate underlines its usefulness.

Keywords   Unit root test · Nonlinearities · Smooth transition · Nonstandard testing · Real effective exchange rates

1 Introduction

Nonlinear time series models like the smooth transition autoregressive (STAR) model, see Teräsvirta (1994), have become very popular in the last years. In this paper we are mainly concerned with the exponential STAR (ESTAR) model and develop a new test for the unit root hypothesis against a globally stationary ESTAR model. In particular, we focus on a prominent and widely applied specification of this model that allows for a unit root regime and two symmetric mean-reverting regimes. The time series process,

R. Kruse (✉)
CREATES, School of Economics and Management, University of Aarhus,
Building 1322, 8000 Aarhus C, Denmark
e-mail: rkruse@creates.au.dk

say y_t, behaves like a random walk if y_{t-1} was close to some location parameter c and it is mean-reverting as y_{t-1} departs from c. In the exponential smooth transition model the degree of mean-reversion depends on the squared difference between y_{t-1} and c.

When modeling real exchange rates for example, the economic intuition behind this specification is that the real exchange rate behaves like a nonstationary random walk if it was quite close its long run equilibrium value in the last period and that there are driving forces like arbitrage that leads to mean-reversion if the real exchange rate departs from its long run equilibrium. Moreover, arbitrage may not be profitable if the departure is small as arbitrageurs have to pay transaction costs for instance. Therefore, the degree of mean-reversion is small as well and vice versa. These facts make this ESTAR specification quite attractive for modeling economic time series like real exchange and interest rates, unemployment rates and log dividend yields.

There are a lot of economic theories like purchasing power parity (PPP), to name a highly debated one, that imply certain time series properties, i.e. the stationarity of real exchange rates. Often the unit root hypothesis, which contradicts PPP, is tested against stationarity with extant linear unit root tests. These tests have however not the highest power when the true data generating process exhibits nonlinearities. Therefore, many recent empirical studies make use of nonlinear unit root tests.

Regarding the ESTAR specification from above, a popular Dickey–Fuller-type test has been proposed by Kapetanios et al. (2003). However, this test assumes that the location parameter c in the smooth transition function is equal to zero. On the contrary, a lot of empirical studies on real exchange rates report significant estimates of c, cf. Michael et al. (1997), Sarantis (1999), Taylor et al (2001) and more recently, Rapach and Wohar (2006). When relaxing this assumption, we are faced with a nonstandard testing problem, i.e. a joint hypothesis where one parameter is one-sided under the alternative while all others are two-sided. Since standard inference techniques are not appropriate in this situation, we make use of the new approach by Abadir and Distaso (2007) who propose a class of modified test statistics in order to tackle such non-standard testing problems. Our aim is to derive a unit root test that permits a nonzero location parameter c and that can compete with the extant one of Kapetanios et al. (2003) in terms of size and power.

After introducing the ESTAR specification in more detail and presenting the existing test by Kapetanios et al. (2003) in Sect. 2, the inference techniques by Abadir and Distaso (2007) and the new test are discussed in Sect. 3. The nonstandard limiting distribution of the test statistic is derived and consistency of the test is proven. Moreover, we show that the limiting distribution remains unchanged if we account for potential serial correlation in the error terms by augmenting the test regression with lags of the dependent variable. By means of a Monte Carlo study in Sect. 4, we compare the small sample properties of both tests under a variety of conditions. The new test is correctly sized and quite often superior in terms of power. Both tests have lower but substantial power when the true data generating process is a logistic STAR model. However, the new test has generally higher power against logistic STAR models than the extant test. Finally, we provide an application to a monthly real effective exchange rate time series for the European Union in Sect. 5. The results suggest the validity of PPP if the new test is used and the opposite if the extant test is applied. Proofs are given in the Appendix.

2 Dickey–Fuller-type unit root test against ESTAR

The ESTAR specification we are concerned with is formally given by

$$\triangle y_t = \alpha y_{t-1} + \phi y_{t-1}(1 - \exp\{-\gamma(y_{t-1} - c)^2\}) + \varepsilon_t,$$

where $\varepsilon_t \sim \text{iid}(0, \sigma^2)$. If the smoothness parameter γ approaches zero, the ESTAR model becomes a linear AR(1) model, i.e. $\triangle y_t = \alpha y_{t-1} + \varepsilon_t$ that is stationary if $-2 < \alpha < 0$. In the following, α is set equal to zero which means that the ESTAR model becomes a random walk if $\gamma = 0$. Kapetanios et al. (2003) show that the ESTAR model under the restriction $\alpha = 0$,

$$\triangle y_t = \phi y_{t-1}(1 - \exp\{-\gamma(y_{t-1} - c)^2\}) + \varepsilon_t,$$

is globally stationary if $-2 < \phi < 0$ is true although it is locally nonstationary in the sense that it contains a partial unit root when $y_{t-1} = c$ holds. Additionally note that the random walk model can also be achieved when imposing the restriction $\phi = 0$. This means that a direct test for the unit root hypothesis is infeasible since ϕ is not identified when testing $H_0 : \gamma = 0$ and vice versa. A popular approach to avoid the presence of nuisance parameters under the null hypothesis is to use a Taylor approximation of the smooth transition function $G(y_{t-1}; \gamma, c) = 1 - \exp\{-\gamma(y_{t-1} - c)^2\})$ around $\gamma = 0$, see Luukkonen et al. (1988). This approach was adopted by Kapetanios et al. (2003) and we construct the new test on the same basis. More specifically, Kapetanios et al. (2003) make the restriction $c = 0$ and consider the model

$$\triangle y_t = \phi y_{t-1}(1 - \exp\{-\gamma y_{t-1}^2\}) + \varepsilon_t.$$

An application of a first-order Taylor approximation leads to the auxiliary regression

$$\triangle y_t = \beta_1 y_{t-1}^3 + u_t, \tag{1}$$

with u_t being a noise term depending on ε_t, ϕ and the remainder of the Taylor expansion. Obviously, it looks very much like the famous Dickey–Fuller test regression without deterministic terms. The cubic term y_{t-1}^3 approximates the ESTAR nonlinearity.

The authors suggest a t test for the unit root hypothesis against globally stationary ESTAR which corresponds to $H_0 : \beta_1 = 0$ versus $H_1 : \beta_1 < 0$. Hence, the unit root test is carried out by estimating the auxiliary regression (1) and computing a Dickey–Fuller-type t test, labeled as KSS,

$$\text{KSS} \equiv \frac{\hat{\beta}_1}{\sqrt{\hat{\text{var}}(\hat{\beta}_1)}} = \frac{\sum_{t=1}^{T} y_{t-1}^3 \triangle y_t}{\sqrt{\hat{\sigma}^2 \sum_{t=1}^{T} y_{t-1}^6}}, \tag{2}$$

where $\hat{\sigma}^2 = \frac{1}{T} \sum_{t=1}^{T} (\triangle y_t - \hat{\beta}_1 y_{t-1}^3)^2$ is the usual estimator of the error variance. Let $W(r)$ be the Brownian motion defined on $r \in [0, 1]$ and let $\Rightarrow$ denote convergence in

distribution. The limiting distribution of the KSS statistic is then given by

$$\text{KSS} \Rightarrow \frac{\frac{1}{4}W(1)^4 - \frac{3}{2}\int_0^1 W(r)^2 dr}{\left(\int_0^1 W(r)^6 dr\right)^{1/2}},$$

see Theorem 1 in Kapetanios et al. (2003). Regarding deterministic terms, they suggest to demean or detrend the data in a first step, i.e.

$$y_t = \omega' d_t + \xi_t$$

with $d_t = 1$ or $d_t = [1 \quad t]'$ and ω is a parameter vector of suitable dimension. In a second step, the unit root test is applied to $\widehat{\xi}_t$. As a consequence, the asymptotic distribution of the KSS statistic depends on functionals of a demeaned or detrended Brownian motion, respectively. The demeaned and detrended Brownian motion are given by

$$W(r) - \int_0^1 W(s) ds,$$

$$W(r) + (6r - 4) \int_0^1 W(s) ds + (12r - 6) \int_0^1 s W(s) ds,$$

respectively. For details concerning this test such as proofs and critical values see Kapetanios et al. (2003).

3 Modified Wald-type test

In order to allow for a nonzero location parameter c in the exponential transition function we consider the nonlinear time series model

$$\triangle y_t = \phi y_{t-1}(1 - \exp\{-\gamma(y_{t-1} - c)^2\}) + \varepsilon_t. \tag{3}$$

Following Kapetanios et al. (2003), we apply a first-order Taylor approximation to $G(y_{t-1}; \gamma, c) = (1 - \exp\{-\gamma(y_{t-1} - c)^2\})$ around $\gamma = 0$ and proceed with the test regression

$$\triangle y_t = \beta_1 y_{t-1}^3 + \beta_2 y_{t-1}^2 + \beta_3 y_{t-1} + u_t.$$

Following Kapetanios et al. (2003) we impose $\beta_3 = 0$ to improve the power of the test, see Kapetanios et al. (2003, footnote 5). Henceforth, we proceed with

$$\triangle y_t = \beta_1 y_{t-1}^3 + \beta_2 y_{t-1}^2 + u_t. \tag{4}$$

where $\beta_1 = \gamma\phi$ and $\beta_2 = -2c\gamma\phi$. We are interested in the pair of hypotheses given by $H_0 : \gamma = 0$ against $H_1 : \gamma > 0$. In the test regression (4), this pair of hypothesis is equivalent to $H_0 : \beta_1 = \beta_2 = 0$ against $H_1 : \beta_1 < 0, \beta_2 \neq 0$. Note that the two-sidedness of β_2 under H_1 stems from the fact that c is allowed to take real values. This testing problem is nonstandard in the sense that one parameter is one-sided under H_1 while the other one is two-sided. A standard Wald test would be inappropriate and we therefore apply the methods of Abadir and Distaso (2007) to derive a suitable test. In a nutshell, the one-sided parameter is orthogonalized with respect to the two-sided one. The modified Wald test builds upon the one-sided parameter (β_1) and the transformed two-sided parameter, say $\beta_2^\perp$, that are stochastically independent by definition.

Let the parameter vector of the regression model (4) be $\theta = [\beta_1 \ \beta_2]'$. Following the notation of Abadir and Distaso (2007), the null hypothesis of a unit root is rewritten as

$$H_0 : h(\theta) \equiv [h_1(\theta) \ h_2(\theta)]' = [\beta_1 \ \beta_2]' = [0 \ 0]'.$$

The alternative hypothesis of a globally stationary ESTAR model is given by

$$H_1 : h_1(\theta) < 0 \text{ or } h_2(\theta) \neq 0,$$

which includes the subset hypothesis $H_1^\cap : h_1(\theta) < 0$ and $h_2(\theta) \neq 0$. Theorem 6 in Abadir and Distaso (2007) states that the modified Wald test is consistent against H_1 as well as $H_1^\cap$. The standard Wald test statistic based on the Hessian matrix $\mathcal{H}$ is

$$W_{\mathcal{H}} = h(\hat{\theta})' \hat{V}^{-1} h(\hat{\theta}),$$

where $\hat{V} \equiv \left[\frac{\partial h(\theta)}{\partial \theta'} \big|_{\theta=\hat{\theta}} (-\hat{\mathcal{H}})^{-1} \frac{\partial h(\theta)'}{\partial \theta} \big|_{\theta=\hat{\theta}} \right]$ with elements v_{ij}. In general, the modified Wald test statistic of Abadir and Distaso (2007) is given by

$$\tau = \left(\frac{\partial h_{2\cdot 1}(\theta)}{\partial \theta'} \Big|_{\theta=\hat{\theta}} (-\hat{\mathcal{H}})^{-1} \frac{\partial h_{2\cdot 1}(\theta)'}{\partial \theta} \Big|_{\theta=\hat{\theta}} \right)^{-1} \hat{h}_{2\cdot 1}(\hat{\theta})^2$$
$$+ 1(h_1(\hat{\theta}) < 0) \left(\frac{\partial h_1(\theta)}{\partial \theta'} \Big|_{\theta=\hat{\theta}} (-\hat{\mathcal{H}})^{-1} \frac{\partial h_1(\theta)'}{\partial \theta} \Big|_{\theta=\hat{\theta}} \right)^{-1} \hat{h}_1(\hat{\theta})^2,$$

with $h_{2\cdot 1}(\theta)$ being

$$h_{2\cdot 1}(\theta) = h_2(\theta) - \frac{h_1(\theta) v_{21}}{v_{11}}.$$

The estimator of $h_{2\cdot 1}(\theta)$ is simply given by $\hat{h}_{2\cdot 1}(\hat{\theta}) = h_2(\hat{\theta}) - \frac{h_1(\hat{\theta})\hat{v}_{21}}{\hat{v}_{11}}$. Based on these results, straightforward calculations lead us to

$$\tau = \left(\hat{v}_{22} - \frac{\hat{v}_{21}^2}{\hat{v}_{11}} \right)^{-1} \left(\hat{\beta}_2 - \hat{\beta}_1 \frac{\hat{v}_{21}}{\hat{v}_{11}} \right)^2 + 1(\hat{\beta}_1 < 0) \frac{\hat{\beta}_1^2}{\hat{v}_{11}},$$

which is the new test statistic for the unit root hypothesis against globally stationary ESTAR. A simpler and more intuitive way to formulate this statistic is

$$\tau = t^2_{\beta_2^\perp = 0} + 1(\hat{\beta}_1 < 0)t^2_{\beta_1 = 0}.$$

The two summands appearing in the test statistic τ can be interpreted as follows: the first term is a squared t-statistic for the hypothesis $\beta_2^\perp \equiv \beta_2 - \beta_1 v_{21}/v_{11} = 0$ with $\beta_2^\perp$ being orthogonal to β_1. Additionally, the second term is a squared t-statistic for the hypothesis $\beta_1 = 0$, the one-sidedness under H_1 is achieved by the multiplied indicator function. In the next step, the limiting distribution of τ is derived.

Assumption 1 y_t is a random walk, i.e. $y_t = y_{t-1} + \varepsilon_t$ with $y_0 = 0$ and u_t being iid$(0, \sigma^2)$ and $E|u_t|^\delta < \infty$ for some $\delta \geq 6$.

In Theorem 1 we derive the asymptotic distribution of τ under the null hypothesis $H_0 : \gamma = 0$.

Theorem 1 *Under Assumption 1 the τ statistic has the following asymptotic distribution which is free of nuisance parameters*:

$$\tau \Rightarrow \mathcal{A}(W(r)) + \mathcal{B}(W(r)),$$

where $\mathcal{A}$ and $\mathcal{B}$ are functions of the Brownian motion $W(r)$ that are given in the proof. Under the alternative hypothesis $H_1 : \gamma > 0$ the τ statistic diverges with rate T.

Proof See Appendix.

We follow the approach of Kapetanios et al. (2003) and demean or detrend the data in a first step when allowing for deterministic terms. This means that the Brownian motion $W(r)$ appearing in the limiting distribution of the τ statistic has to be replaced by a demeaned or detrended Brownian motion, respectively.

Next, we consider the case of serially correlated errors. We allow for stationary linear innovations that are generated by a short-memory process v_t.

Assumption 2 $v_t = \psi(L)u_t = \sum_{j=0}^\infty \psi_j u_{t-j}$, where $\sum_{j=0}^\infty j|\psi_j| < \infty$ and $u_t \sim$ iid$(0, \sigma^2)$.

In Theorem 2 we show that the asymptotic distribution of the τ statistic does not change when adding a sum of lagged differences on the right hand side of the test regression, i.e.

$$\triangle y_t = \beta_1 y_{t-1}^3 + \beta_2 y_{t-1}^2 + \sum_{i=1}^p \rho_i \triangle y_{t-i} + u_t. \tag{5}$$

Theorem 2 *The asymptotic distribution of the τ statistic does not change when the test regression in (5) is used instead of (4).*

Proof See Appendix.

Table 1 Critical values of the τ statistic

	$d_t = 0$	$d_t = 1$	$d_t = [1 \quad t]'$
1%	13.15	13.75	17.10
5%	9.53	10.17	12.82
10%	7.85	8.60	11.10

Alternatively, one could consider more general error processes and derive a Phillips-Perron-type test for the unit root hypothesis, see Rothe and Sibbertsen (2006) and Sandberg (2009). However, we focus on the augmented Dickey–Fuller version in this paper and leave the other for future research.

4 Monte Carlo study

This section covers the Monte Carlo study that compares the small sample performance of the new unit root test and the existing test by Kapetanios et al. (2003). Throughout this section we set the number of observations T equal to 150 and 300 which are reasonable sample sizes for many macroeconomic and financial time series like unemployment rates and interest rates. Furthermore, 500 initial observations are deleted to reduce the effect of initial conditions. Asymptotic critical values for the modified Wald-type test τ are provided in Table 1. They are based on 20,000 replications and $T = 1,000$. We report critical values for raw ($d_t = 0$), demeaned ($d_t = 1$) and detrended data ($d_t = [1 \quad t]'$) for nominal significance levels of 1, 5 and 10%, respectively.

We investigate the size of both tests under the following data generating processes

$$y_t = y_{t-1} + \varepsilon_t \quad \text{with } \varepsilon_t = u_t \tag{6}$$

$$y_t = y_{t-1} + \varepsilon_t \quad \text{with } \varepsilon_t = \rho \varepsilon_{t-1} + u_t \tag{7}$$

$$y_t = y_{t-1} + \varepsilon_t \quad \text{with } \varepsilon_t = u_t - \theta u_{t-1}, \tag{8}$$

where u_t is drawn from the standard normal distribution. The errors ε_t follow an iid process, an AR(1) or MA(1) process. We adopt the approach of Phillips and Sul (2003) and sample parameters from uniform distributions in order to cover a wide range of values in a relatively small number of experiments. Hence, the autoregressive and moving average parameters ρ and θ are drawn from uniform distributions:

$$\rho \sim U\left[\underline{\rho}, \bar{\rho}\right] \quad \text{and} \quad \theta \sim U\left[\underline{\theta}, \bar{\theta}\right].$$

We specify $\underline{\rho} = \underline{\theta} = 0$ and $\bar{\rho} = \bar{\theta} = 0.4$. The size experiments are based on 5,000 replications, results can be found in Table 2. Most rejection rates under the null hypothesis are quite close to the nominal ones which suggests that both tests are correctly sized. We observe that both tests are a little bit oversized in the presence of errors that follow a first-order moving average process.

Table 2 Size experiments

d_t	Test	iid			AR(1)			MA(1)		
		1.0	5.0	10.0	1.0	5.0	10.0	1.0	5.0	10.0
$T = 150$										
0	KSS	0.9	5.0	9.7	1.1	4.7	9.6	0.9	4.5	9.8
	τ	0.9	4.9	9.5	1.4	4.8	9.2	1.1	5.4	10.2
1	KSS	1.2	5.1	9.4	1.1	5.5	9.8	1.4	6.5	12.1
	τ	1.0	4.8	9.2	1.0	4.9	9.7	1.3	6.1	11.9
$[1 \quad t]'$	KSS	1.0	4.7	9.6	1.2	4.8	9.8	1.8	6.7	12.1
	τ	0.9	5.3	9.8	1.0	5.4	9.7	1.6	6.8	12.0
$T = 300$										
0	KSS	0.9	5.1	9.9	1.0	5.0	10.0	1.2	4.7	10.1
	τ	1.1	5.2	10.9	1.0	5.0	10.0	1.7	5.2	11.9
1	KSS	1.1	5.1	9.9	1.1	4.8	9.7	1.3	6.0	11.8
	τ	1.1	4.9	9.8	0.9	4.6	9.5	1.2	5.6	11.8
$[1 \quad t]'$	KSS	1.1	4.9	10.2	0.8	5.2	9.6	1.4	6.5	10.9
	τ	1.0	5.2	10.1	0.9	5.4	10.3	1.5	6.6	11.2

Reported values are rejection rates of KSS and τ test under the validity of H_0

Next, we study the power of both tests by considering different settings in the globally stationary nonlinear ESTAR process for the parameters c and γ. The data generating process we consider under the nonlinear ESTAR alternative is

$$\triangle y_t = \phi y_{t-1}(1 - \exp\{-\gamma(y_{t-1} - c)^2\}) + \varepsilon_t \qquad (9)$$

with $\phi = -1$. This restriction is often imposed in empirical studies, see for example Taylor et al (2001). The location parameter c is set either equal to zero or it is drawn from a uniform distribution with lower and upper bound, $\underline{c}$ and $\bar{c}$, respectively. Analogously, the smoothness parameter γ is drawn from a uniform distribution with lower and upper bound $\underline{\gamma}$ and $\bar{\gamma}$, respectively:

$$c \sim U\left[\underline{c}, \bar{c}\right] \quad \text{and} \quad \gamma \sim U\left[\underline{\gamma}, \bar{\gamma}\right].$$

Results of these power experiments are reported in Table 3. In the first experiment we specify a zero location parameter ($c_0 \equiv c = 0$) and slow transition between regimes ($\gamma_l \equiv \gamma \sim U[0.001, 0.01]$). In the second and third experiments a nonzero location parameter is allowed by drawing it from a uniform distribution with lower and upper bound of -5 (-10) and 5 (10), respectively. The fourth and fifth settings restrict the upper bound $\bar{c}$ to zero in order to have a nonzero mean of c. Please note that we do not report results for experiments where to the lower bound $\underline{c}$ is restricted to zero because there is no qualitative difference because of symmetry. The last two experiments are like the two previous ones but with fast transition between regimes, i.e. $\gamma_h \equiv \gamma \sim U[0.01, 0.1]$.

Table 3 Power experiments for exponential STAR

d_t	Test	c_0, γ_l	$c_{\pm 5}, \gamma_l$	$c_{\pm 10}, \gamma_l$	c_{-5}, γ_l	c_{-10}, γ_l	c_{-5}, γ_h	c_{-10}, γ_h
$T = 150$								
0	KSS	88.9	60.8	55.2	62.2	56.0	88.2	90.7
	τ	58.4	49.3	51.7	50.3	49.3	90.4	92.4
1	KSS	51.9	47.8	50.4	46.6	49.0	92.8	93.1
	τ	54.0	51.9	60.8	52.3	60.1	99.5	99.4
$[1 \quad t]'$	KSS	30.7	26.5	37.3	28.9	37.6	87.1	92.1
	τ	33.6	30.3	42.3	30.4	44.3	95.4	96.6
$T = 300$								
0	KSS	99.6	92.4	73.0	92.0	74.4	97.0	97.2
	τ	95.3	88.1	72.6	88.5	74.5	98.2	97.9
1	KSS	91.1	87.6	77.8	88.4	79.7	98.1	97.5
	τ	92.9	92.3	93.3	91.8	93.9	100.0	100.0
$[1 \quad t]'$	KSS	77.5	74.2	64.1	73.4	65.9	97.7	96.3
	τ	81.6	78.9	78.4	79.6	78.7	100.0	100.0

Reported values are rejection rates of KSS test (upper entries) and τ test (lower entries). Nominal significance level is 5%

When interpreting the simulated rejection probabilities against ESTAR we observe that the new test is generally superior to the Kapetanios et al. (2003) test in terms of power. Only in some cases where the unit root tests are applied to raw data ($d_t = 0$), the KSS test performs better than the modified Wald test. Please note, that most applications of unit root tests in economics involve deterministic terms. When data is demeaned or detrended, power gains up to 11% (for $T = 150$) and 15% (for $T = 300$) can be achieved by applying the new test. Furthermore, even for $T = 150$ the new test dominates the one proposed by Kapetanios et al. (2003) for nonzero deterministic terms ($d_t \neq 0$), although the differences between the tests are smaller than for $T = 300$. Given the larger sample size of $T = 300$, the empirical power of both tests is quite high and we observe that it drops significantly for both tests, especially for low values of γ (γ_l), if the smaller sample size of $T = 150$ is considered.

In addition, we study the power of both tests against globally stationary logistic STAR (LSTAR) models. As noted by Kapetanios et al. (2003), an alternative nonlinear adjustment scheme to the exponential one is a logistic smooth transition function. We use the second-order logistic function

$$G(y_{t-1}; \gamma, c_1, c_2) = 2/[1 + \exp(-\gamma(y_{t-1} - c_1)(y_{t-1} - c_2))] - 1$$

that has two location parameters, namely c_1 and c_2. Like the exponential smooth transition function it becomes constant if $\gamma \to 0$, which means that the nonlinear logistic smooth transition model becomes linear. Without loss of generality, we set $c_1 = 0$ and draw c_2 from the uniform distribution as done before in the case of an exponential smooth transition. The data generating process is now given by

Table 4 Power experiments for logistic STAR

d_t	Test	c_0, γ_l	$c_{\pm 5}, \gamma_l$	$c_{\pm 10}, \gamma_l$	c_{-5}, γ_l	c_{-10}, γ_l	c_{-5}, γ_h	c_{-10}, γ_h
$T = 150$								
0	KSS	74.5	63.2	42.6	62.3	44.9	98.9	98.1
	τ	35.0	34.1	44.2	33.9	44.2	96.8	98.0
1	KSS	30.6	27.6	32.2	28.8	32.9	89.2	88.5
	τ	31.3	29.7	38.9	30.4	39.7	95.6	97.9
$[1 \quad t]'$	KSS	24.3	15.7	22.3	16.5	23.1	81.2	83.2
	τ	28.1	17.2	26.4	18.0	27.2	88.4	92.4
$T = 300$								
0	KSS	97.8	95.4	87.8	94.9	89.2	99.9	99.9
	τ	83.2	82.5	82.7	80.3	83.3	99.9	100.0
1	KSS	75.6	72.4	63.7	71.4	64.6	99.0	94.4
	τ	79.0	77.5	78.3	77.9	80.0	100.0	100.0
$[1 \quad t]'$	KSS	53.3	50.4	46.5	50.8	48.5	97.5	92.3
	τ	58.6	54.9	57.7	55.4	59.4	99.5	99.8

Reported values are rejection rates of KSS test (upper entries) and τ test (lower entries). Nominal significance level is 5%

$$\triangle y_t = y_{t-1}(1 - 2/[1 + \exp(-\gamma(y_{t-1} - c_1)(y_{t-1} - c_2))]) + \varepsilon_t. \qquad (10)$$

Empirical rejection frequencies are reported in Table 4. Both tests have higher power against ESTAR than LSTAR models which is not surprising since both have the former one as specific alternative. Nonetheless, one might expect that both tests have substantial power against logistic STAR models because the Taylor approximation of a logistic STAR model is quite similar. Thus, a rejection of the null hypothesis does not necessarily contain information about the specific *type* of nonlinear adjustment. When comparing both unit root tests, we come to the same conclusions as before. In addition we observe that the power is lower for detrended data than for demeaned data which is due to an additional parameter that has to be estimated when detrending the data.

In sum, the new test shows good overall performance and is quite often more powerful than the existing test by Kapetanios et al. (2003). In the cases of demeaned and detrended data, which are the most important cases in practice, the new test is indeed uniformly better in terms of power.

5 Empirical application

Unit root tests have become a very popular tool in the literature that is concerned with testing validity of the PPP which counts to one of the most important parities in international macroeconomics. One can state that PPP holds if and only if the real exchange rate is stationary. Henceforth, testing the unit root hypothesis means testing the nonvalidity of the PPP theory. Since linear unit root tests like the ones of
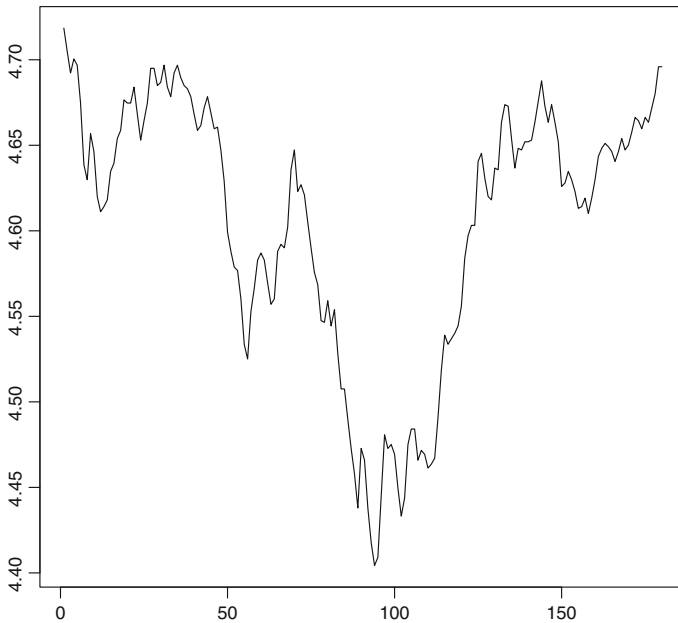
Fig. 1 Logarithm of EU real effective exchange rate (January 1993 to December 2007)

Dickey and Fuller (1979) and Phillips and Perron (1988) often fail to reject the null hypothesis of nonstationarity when being applied to real exchange rate data, researchers tend to use nonlinear unit root tests where the specific model that is true under the alternative is congruent with economic models of financial markets. For example, STAR models for the real exchange rate can be interpreted in the context of transaction costs and arbitrage, see Dumas (1992), Sercu et al. (1995) and Michael et al. (1997). However, rejecting the null hypothesis in favor of a nonlinear alternative while a linear Dickey–Fuller test does not reject in favor of a linear alternative might hint at non-linearities. Nonetheless, one should be careful with the conclusion that nonlinearity is of ESTAR-type because the test regression approximates a lot of nonlinear models and such tests can have substantial power against other nonlinear mean-reverting processes, see Sect. 4.

We apply both unit root tests against nonlinear alternatives and two famous tests against linear alternatives to the monthly real effective exchange rate time series for the European Union. Our data is taken from Datastream (code: EMXTW..RF) and spans from January 1993 to December 2007 implying 180 observations. The logged time series is depicted in Fig. 1. No linear trend can be seen in the data but the mean appears to be highly significant. Hence, we demean the data in a first step. In a second step we estimate the test regressions with a lag length chosen accordingly to the Schwarz information criterion ($\widehat{p} = 1$). We obtain KSS $= -2.21$ which is not significant at the 10% level suggesting that PPP does not hold. On the contrary, using the new test one has to reject the null hypothesis since $\tau = 9.19$ is significant at the 10% level which indicates that PPP holds true. Furthermore, the unit root tests against

linear alternatives by Dickey and Fuller (1979) (DF) and Phillips and Perron (1988) (PP) do not provide any evidence against the null hypothesis. The test statistics are DF $= -1.76$ and PP $= -1.60$, respectively.

We conclude that nonlinearities, potentially of exponential STAR-type with non-zero location, are present in the data but that they are not detected by applying existing tests. The modified Wald unit root test yields new evidence on the stationarity of the EU real effective exchange rate which suggests the validity of PPP.

6 Conclusions

This paper contributes to the literature on nonlinear unit root tests by generalizing the existing test by Kapetanios et al. (2003) with respect to a nonzero location parameter. The resulting nonstandard testing problem is tackled by deriving a modified Wald-type test that builds upon the inference techniques by Abadir and Distaso (2007). The nonstandard limiting distribution of the test statistic has been derived under standard assumptions and consistency has been proven. The Monte Carlo study shows that the new test is in most situations superior to the extant test. An empirical application to the EU real effective exchange rate underpins its usefulness.

Appendix

Proof of Theorem 1 In order to simplify the notation, we write $\int$ instead of $\int_0^1$ in the following. The proof makes use of the following convergence results, see Hansen (1992) and Hamilton (1994). We have

$$\frac{1}{T^{(i+2)/2}} \sum_{t=1}^{T} y_{t-1}^i \Rightarrow \sigma^i \int W(r)^i dr \qquad \text{for } i = 4, 5, 6$$

$$\frac{1}{T^{3/2}} \sum_{t=1}^{T} y_{t-1}^2 \triangle y_t \Rightarrow \sigma^3 \left(\frac{1}{3} W(1)^3 - \int W(r) dr \right)$$

$$\frac{1}{T^2} \sum_{t=1}^{T} y_{t-1}^3 \triangle y_t \Rightarrow \sigma^4 \left(\frac{1}{4} W(1)^4 - \frac{3}{2} \int W(r)^2 dr \right).$$

We first note that the second summand of τ is given by $1(\hat{\beta}_1 < 0) t_{\hat{\beta}_1=0}^2$ and that the OLS estimator for β_1 in (4) is given by

$$\hat{\beta}_1 = \frac{\sum_{t=1}^{T} y_{t-1}^4 \sum_{t=1}^{T} y_{t-1}^3 \triangle y_t - \sum_{t=1}^{T} y_{t-1}^2 \triangle y_t \sum_{t=1}^{T} y_{t-1}^5}{\sum_{t=1}^{T} y_{t-1}^4 \sum_{t=1}^{T} y_{t-1}^6 - \left(\sum_{t=1}^{T} y_{t-1}^5 \right)^2}.$$

Under the assumptions of Theorem 1 and by using the convergence results from above we obtain $\hat{\beta}_1 \xrightarrow{P} \beta_1$ and $\hat{\beta}_1 = O_P(T^{-2})$. Furthermore,

$$t_{\beta_1=0} = \frac{\sum_{t=1}^{T} y_{t-1}^4 \sum_{t=1}^{T} y_{t-1}^3 \triangle y_t - \sum_{t=1}^{T} y_{t-1}^2 \triangle y_t \sum_{t=1}^{T} y_{t-1}^5}{\sqrt{\hat{\sigma}^2 \left(\left(\sum_{t=1}^{T} y_{t-1}^4 \right)^2 \sum_{t=1}^{T} y_{t-1}^6 - \sum_{t=1}^{T} y_{t-1}^4 \left(\sum_{t=1}^{T} y_{t-1}^5 \right)^2 \right)}}.$$

Again, by using the convergence results it follows that

$$t_{\beta_1=0} \Rightarrow$$
$$\frac{\left(\int W(r)^4 dr \right) \left(\frac{1}{4} W(1)^4 - \frac{3}{2} \int W(r)^2 dr \right) - \left(\frac{1}{3} W(1)^3 - \int W(r) dr \right) \left(\int W(r)^5 dr \right)}{\sqrt{\left(\int W(r)^4 dr \right)^2 \left(\int W(r)^6 dr \right) - \left(\int W(r)^4 dr \right) \left(\int W(r)^5 dr \right)^2}},$$

and by applying the CMT it follows directly that $t_{\beta_1=0}^2$ converges to the square of the previous function which gives an expression for $\mathcal{B}(W(r))$. Regarding the first summand of τ, we have for the nominator of $t_{\beta_2^\perp=0}$

$$\hat{\beta}_2 - \hat{\beta}_1 \frac{\hat{v}_{21}}{\hat{v}_{11}} = \frac{\sum_{t=1}^{T} y_{t-1}^4 \sum_{t=1}^{T} y_{t-1}^6 \sum_{t=1}^{T} y_{t-1}^2 \triangle y_t - \sum_{t=1}^{T} y_{t-1}^2 \triangle y_t \left(\sum_{t=1}^{T} y_{t-1}^5 \right)^2}{\left(\sum_{t=1}^{T} y_{t-1}^4 \right)^2 \sum_{t=1}^{T} y_{t-1}^6 - \sum_{t=1}^{T} y_{t-1}^4 \left(\sum_{t=1}^{T} y_{t-1}^5 \right)^2}.$$

For the denominator of $t_{\beta_2^\perp=0}$ we have after simple calculations

$$\sqrt{\hat{v}_{22} - \frac{\hat{v}_{21}^2}{\hat{v}_{11}}} = \sqrt{\frac{\hat{\sigma}^2}{\sum_{t=1}^{T} y_{t-1}^4}}.$$

Therefore

$$t_{\beta_2^\perp=0} = \frac{\sum_{t=1}^{T} y_{t-1}^2 \triangle y_t \left(\left(\sum_{t=1}^{T} y_{t-1}^4 \right)^{3/2} \sum_{t=1}^{T} y_{t-1}^6 - \left(\sum_{t=1}^{T} y_{t-1}^4 \right)^{1/2} \left(\sum_{t=1}^{T} y_{t-1}^5 \right)^2 \right)}{\hat{\sigma} \left(\left(\sum_{t=1}^{T} y_{t-1}^4 \right)^2 \sum_{t=1}^{T} y_{t-1}^6 - \sum_{t=1}^{T} y_{t-1}^4 \left(\sum_{t=1}^{T} y_{t-1}^5 \right)^2 \right)}.$$

Using the convergence results stated above it follows that

$$t_{\beta_2^\perp=0} \Rightarrow \frac{\left(\frac{1}{3} W(1)^3 - \int W(r) dr \right) \left(\left(\int W(r)^4 dr \right)^{3/2} \left(\int W(r)^6 dr \right) - \left(\int W(r)^4 dr \right)^{1/2} \left(\int W(r)^5 dr \right)^2 \right)}{\left(\int W(r)^4 dr \right)^2 \left(\int W(r)^6 dr \right) - \left(\int W(r)^4 dr \right) \left(\int W(r)^5 dr \right)^2}.$$

Again, by CMT it follows that $t_{\beta_2^\perp=0}^2$ converges in distribution to the square of the previous function which gives an expression for $\mathcal{A}(W(r))$. It is easy to show that $\hat{\sigma}^2 \xrightarrow{p} \sigma^2$, see Kapetanios et al. (2003). Under the alternative hypothesis, $\triangle y_t$, y_{t-1}^2 and y_{t-1}^3 are

$I(0)$ processes and it is easy to show that the terms appearing in the test statistic are $O_P(T)$. Then, $t^2_{\beta_2^\perp=0} = O_P(T)$ and $t^2_{\beta_1=0} = O_P(T)$, therefore $\tau = O_P(T)$. The τ statistic is therefore diverging with rate T.

Proof of Theorem 2 The proof is very similar to the one of Kapetanios et al. (2003) as it uses the same arguments. Let $Z = [\triangle y_{-1}, \triangle y_{-2}, \dots, \triangle y_{-p}]$ with $\triangle y_{-i} = [\triangle y_{-i+1}, \triangle y_{-i+2}, \dots, \triangle y_{T-i}]$ and $M_T = I_T - Z(Z'Z)^{-1}Z$. Note that, $\hat{\sigma}^2 = \frac{1}{T}\varepsilon' M_T \varepsilon = \frac{1}{T}\varepsilon'\varepsilon + o_p(1) \xrightarrow{P} \sigma^2$ with $\varepsilon = [\varepsilon_1, \varepsilon_2, \dots, \varepsilon_T]'$. Moreover, we have

$$\frac{1}{T^{(i+2)/2}} y_{-1}^{i/2'} M_T y_{-1}^{i/2} = \frac{1}{T^{(i+2)/2}} y_{-1}^{i/2'} y_{-1}^{i/2} + o_p(1) \Rightarrow \lambda^i \int W(r)^i dr \quad \text{for } i = 4, 6$$

$$\frac{1}{T^{7/2}} y_{-1}^{5/2'} M_T y_{-1}^{5/2} = \frac{1}{T^{7/2}} y_{-1}^{5/2'} y_{-1}^{5/2} + o_p(1) \Rightarrow \lambda^5 \int W(r)^5 dr,$$

and additionally,

$$\frac{1}{T^{3/2}} y_{-1}^{2'} M_T \varepsilon = \frac{1}{T^{3/2}} y_{-1}^{2'} \varepsilon + o_p(1) \Rightarrow \frac{1}{3}\lambda^3 W(1)^3 - \lambda\sigma^2 \int W(r) dr$$

$$\frac{1}{T^2} y_{-1}^{3'} M_T \varepsilon = \frac{1}{T^2} y_{-1}^{3'} \varepsilon + o_p(1) \Rightarrow \frac{1}{4}\lambda^4 W(1)^4 - \frac{3}{2}\sigma^2\lambda^2 \int W(r)^2 dr,$$

where λ^2 is the long-run variance of $\triangle y_t$ under the null hypothesis. Based on these results we have

$$t_{\beta_1=0} = \frac{\left(y_{-1}^{2'} M_T y_{-1}^2\right)\left(y_{-1}^{3'} M_T \varepsilon\right) - \left(y_{-1}^{2'} M_T \varepsilon\right)\left(y_{-1}^{5/2'} M_T y_{-1}^{5/2}\right)}{\sqrt{\hat{\sigma}^2\left(\left(y_{-1}^{2'} M_T y_{-1}^2\right)^2 \left(y_{-1}^{3'} M_T y_{-1}^3\right) - \left(y_{-1}^{2'} M_T y_{-1}^2\right)\left(y_{-1}^{5/2'} M_T y_{-1}^{5/2}\right)^2\right)}}$$

$$= \frac{\left(y_{-1}^{2'} y_{-1}^2\right)\left(y_{-1}^{3'} \varepsilon\right) - \left(y_{-1}^{2'} \varepsilon\right)\left(y_{-1}^{5/2'} y_{-1}^{5/2}\right)}{\sqrt{\hat{\sigma}^2\left(\left(y_{-1}^{2'} y_{-1}^2\right)^2 \left(y_{-1}^{3'} y_{-1}^3\right) - \left(y_{-1}^{2'} y_{-1}^2\right)\left(y_{-1}^{5/2'} y_{-1}^{5/2}\right)^2\right)}} + o_p(1).$$

Furthermore,

$$t_{\beta_2^\perp=0} = \frac{y_{-1}^{2'} M_T \varepsilon\left(\left(y_{-1}^{2'} M_T y_{-1}^2\right)^{3/2}\left(y_{-1}^{3'} M_T y_{-1}^3\right) - \left(y_{-1}^{2'} M_T y_{-1}^2\right)^{1/2}\left(y_{-1}^{5/2'} M_T y_{-1}^{5/2}\right)^2\right)}{\hat{\sigma}\left(\left(y_{-1}^{2'} M_T y_{-1}^2\right)^2\left(y_{-1}^{3'} M_T y_{-1}^3\right) - \left(y_{-1}^{2'} M_T y_{-1}^2\right)\left(y_{-1}^{5/2'} M_T y_{-1}^{5/2}\right)^2\right)}$$

$$= \frac{y_{-1}^{2'} \varepsilon\left(\left(y_{-1}^{2'} y_{-1}^2\right)^{3/2}\left(y_{-1}^{3'} y_{-1}^3\right) - \left(y_{-1}^{2'} y_{-1}^2\right)^{1/2}\left(y_{-1}^{5/2'} y_{-1}^{5/2}\right)^2\right)}{\hat{\sigma}\left(\left(y_{-1}^{2'} y_{-1}^2\right)^2\left(y_{-1}^{3'} y_{-1}^3\right) - \left(y_{-1}^{2'} y_{-1}^2\right)\left(y_{-1}^{5/2'} y_{-1}^{5/2}\right)^2\right)} + o_p(1),$$

which, as we have shown before, has the asymptotic distribution given in Theorem 1. Finally, among similar lines in the foregoing proof, it is easily seen that the τ test is consistent under the alternative.

References

Abadir KM, Distaso W (2007) Testing joint hypotheses when one of the alternatives is one-sided. J Econom 140:695–718

Dickey DA, Fuller WA (1979) Distribution of the estimators for autoregressive time series with a unit root. J Am Stat Assoc 74:427–431

Dumas B (1992) Dynamic equilibrium and the real exchange rate in a spatially separated World. Rev Financial Stud 5:153–180

Hamilton JD (1994) Time series analysis. Princeton University Press, Princetion

Hansen BE (1992) Convergence to stochastic integrals for dependent heterogeneous processes. Econom Theory 8:489–500

Kapetanios G, Shin Y, Snell A (2003) Testing for a unit root in the nonlinear STAR framework. J Econom 112:359–379

Luukkonen R, Saikkonen P, Teräsvirta T (1988) Testing linearity against smooth transition autoregressive models. Biometrika 75:491–499

Michael P, Nobay AR, Peel DA (1997) Transactions costs and nonlinear adjustment in real exchange rates: an empirical investigation. J Political Econ 105:862–879

Phillips PCB, Perron P (1988) Testing for a unit root in time series regression. Biometrika 75:335–346

Phillips PCB, Sul D (2003) Dynamic panel estimation and homogeneity testing under cross section dependence. Econom J 6:217–259

Rapach DE, Wohar ME (2006) The out-of-sample forecasting performance of nonlinear models of real exchange rate behavior. Int J Forecast 22:341–361

Rothe C, Sibbertsen P (2006) Phillips-Perron-type unit root tests in the nonlinear ESTAR framework. Allgemeines Stat Arch 90:439–456

Sandberg R (2009) Convergence to stochastic power integrals for dependent heterogeneous processes. Econom Theory (forthcoming)

Sarantis N (1999) Modeling non-linearities in real effective exchange rates. J Int Money Finance 18:27–45

Sercu P, Uppal R, Van Hulle C (1995) The exchange rate in the presence of transaction costs: implications for tests of purchasing power parity. J Finance 50:1309–1319

Taylor MP, Peel DA, Sarno L (2001) Nonlinear mean-reversion in real exchange rates: toward a solution to the purchasing power parity puzzles. Int Econ Rev 42:1015–1042

Teräsvirta T (1994) Specification, estimation and evaluation of smooth transition autoregressive models. J Am Stat Assoc 89:208–218