ELSEVIER

Journal of Econometrics 112 (2003) 359–379

JOURNAL OF Econometrics

www.elsevier.com/locate/econbase

Testing for a unit root in the nonlinear STAR framework

George Kapetanios[a], Yongcheol Shin[b,*], Andy Snell[b]

[a]*Department of Economics, Queen Mary University of London, London, UK*
[b]*Department of Economics, University of Edinburgh, 50 George Square, Edinburgh, EH8 9JY UK*

Abstract

In this paper, we propose a simple testing procedure to detect the presence of nonstationarity against nonlinear but globally stationary exponential smooth transition autoregressive processes. We provide an advance over the existing literature in three senses. *First*, we derive the limiting nonstandard distribution of the proposed tests. *Second*, we find via Monte Carlo simulation exercises that under the alternative of a globally stationary ESTAR process, our proposed test has better power than the standard Dickey–Fuller test, in the region of the null, where the processes are highly persistent. *Third*, we provide an application to ex post real interest rates and bilateral real exchange rates with the US Dollar from the 11 major OECD countries, and find our test is able to reject a unit root in many cases, whereas the linear DF tests fail, providing some evidence of nonlinear mean-reversion in both real interest and exchange rates.

JEL classification: C12; C32; E43; F31

Keywords: Exponential smooth transition autoregressive model; Unit roots; Globally stationary nonlinear processes; Monte Carlo simulations; Real interest and exchange rates

1. Introduction

There is a growing dissatisfaction with the standard linear ARMA framework which investigators use to test for unit roots. Much of this arises from the fact that in several areas of economics a theoretical prediction of stationarity is confounded in practice by the persistent failure of the standard Dickey–Fuller (DF) test to reject the null of a unit root. For example, in international monetary economics the regular finding of a unit root in real exchange rates causes discomfort to economists who wish to build

* Corresponding author.
E-mail address: y.shin@ed.ac.uk (Y. Shin).

0304-4076/02/$ - see front matter
PII: S0304-4076(02)00202-6

models around a long run purchasing power parity (PPP) relationship (e.g., Taylor et al., 2001). Another notable area of discomfort occurs in macro-finance where apparent unit root behaviour in real interest rates violates mean reversion in the aggregate marginal product of capital and makes the adherence to simple constant returns to scale production functions used in modern macroeconomic growth and RBC theory hard to maintain (see for example Rose, 1988).

Some authors have accepted the results of the tests and have reformulated the economic theory. For example, Edison and Kloveland (1987) point out that whilst the homogeneity postulate behind the standard view of PPP is only likely to hold in the long run, long runs of data may encounter regime changes in tastes and technology which in turn imply permanent movements in the terms of trade or in the relative price of traded to nontraded goods. They find that adjusting for "general equilibrium" shocks, enables them to reject the unit root in real exchange rates and provide support for the PPP hypothesis. Increasingly though, investigators are looking to alternative frameworks within which to test for unit roots. The literature here has two branches. The first focuses on the use of panel data and its role in increasing the power of standard unit root tests. Abuaf and Jorion (1990) used a panel data test and rejected the joint hypothesis of unit roots in each of a group of real exchange rates against an alternative that they are stationary. See also Frankel and Rose (1996) and Wu (1996). For a general econometric discussion see Im et al. (2002). The second branch uses alternative forms of stationarity to AR or ARMA models such as fractional integration (e.g., Mills, 1993) and nonlinear transition dynamics (e.g., Pesaran and Potter, 1997).

Recently, Balke and Fomby (1997) have popularised a joint analysis of nonstationarity and nonlinearity in the context of threshold cointegration. In particular, using Monte Carlo experiments based on the threshold autoregressive model with three regimes they have shown that the power of the DF test falls dramatically with threshold parameters. See also Pippenger and Goering (1993). Many other authors have attempted to address similar issues in the context of a threshold autoregressive (TAR) model. For example, papers by Enders and Granger (1998), Berben and van Dijk (1999), Caner and Hansen (2001), Lo and Zivot (2001) and Kapetanios and Shin (2001) all form part of the growing literature that examines the interplay between nonstationarity, cointegration and nonlinearity.

In this paper, we analyse the implications of the existence of a particular kind of nonlinear dynamics for unit root testing procedures, and provide an alternative framework for a test of the null of a unit root process against an alternative of a nonlinear exponential smooth transition autoregressive (ESTAR) process, which is globally stationary. For a survey on the recent developments of STAR modelling see van Dijk et al. (2002). There has been, however, no serious attempt to directly distinguish nonstationary linear systems from stationary nonlinear STAR ones. Sercu et al. (1995) show that equilibrium models of real exchange rate determination in the presence of transactions costs imply a nonlinear adjustment process toward PPP. Furthermore, Michael et al. (1997) argue that conventional cointegration or unit root tests which ignore the effect of the STAR nonlinearity, may be biased against the long run PPP hypothesis. To this end, we develop a test procedure that is specifically designed to have power against the globally stationary ESTAR process.

The current paper provides an advance over the existing literature in three senses. *First*, we derive the nonstandard limiting distribution of the suggested tests. *Second*, we conduct Monte Carlo simulation exercises and examine small sample size and power performance of our proposed tests. We find *inter alia* that under the alternative of a globally stationary ESTAR process, our proposed test has a power gain over the DF tests that is highest in the region of the null i.e., in cases where the processes are highly persistent. *Third*, we provide an application to ex post real interest rates and bilateral real exchange rates from the 11 major OECD countries, which demonstrates the empirical worth of our approach. In particular, our proposed test is able to reject a unit root in several cases where the DF tests fails to do so, providing some evidence of nonlinear mean-reversion in both real interest and exchange rates.

The plan of the paper is as follows: Section 2 defines a globally stationary ESTAR process, develops the proposed test statistics, derives their asymptotic distributions and provides asymptotic critical values. Section 3 addresses the issue of the small sample performance of the proposed tests that take account of the specific nonlinear nature of the alternative. Section 4 presents empirical applications. Section 5 contains some concluding remarks. Mathematical proofs are proved in the appendix.

2. Testing the null of a unit root against the alternative of a globally stationary ESTAR process

Consider a univariate smooth transition autoregressive of order 1, STAR(1) model,

$$y_t = \beta y_{t-1} + \gamma y_{t-1}\Theta(\theta; y_{t-d}) + \varepsilon_t, \quad t = 1,\dots,T, \tag{1}$$

where $\varepsilon_t \sim iid(0,\sigma^2)$, and β and γ are unknown parameters. To begin with we assume that y_t is a mean zero stochastic process. We discuss processes with nonzero mean and/or with a linear time trend after Theorem 2.1. Following the literature on STAR models, the transition function adopted here is of the exponential form, i.e.,

$$\Theta(\theta; y_{t-d}) = 1 - \exp(-\theta y_{t-d}^2), \tag{2}$$

where we assume that $\theta \geqslant 0$, and $d \geqslant 1$ is the delay parameter. The exponential transition function is bounded between zero and 1, i.e. $\Theta : \mathbb{R} \to [0,1]$ has the properties:

$$\Theta(0) = 0; \qquad \lim_{x\to\pm\infty} \Theta(x) = 1$$

and is symmetrically U-shaped around zero.[1]

[1] An alternative nonlinear adjustment scheme to the one favoured in this paper is given either by the first-order logistic function, $\Theta(\theta; y_{t-d}) = 2/(1+\exp(-\theta y_{t-d})) - 1$, which is bounded between -1 and 1, or by the second-order logistic function, $\Theta(\theta; y_{t-d}) = 2/(1+\exp(-\theta y_{t-d}^2)) - 1$, which is bounded between 0 and 1. In other applications this scheme might be useful, but here we focus on the exponential case.

Using (2) in (1) we obtain an exponential STAR (ESTAR) model,

$$y_t = \beta y_{t-1} + \gamma y_{t-1}[1 - \exp(-\theta y_{t-d}^2)] + \varepsilon_t, \tag{3}$$

which can be conveniently reparameterised as

$$\Delta y_t = \phi y_{t-1} + \gamma y_{t-1}[1 - \exp(-\theta y_{t-d}^2)] + \varepsilon_t, \tag{4}$$

where $\phi = \beta - 1$. If θ is positive, then it effectively determines the speed of mean reversion. Representation (3) makes economic sense in that many economic models predict that the underlying system tends to display a dampened behaviour towards an attractor when it is (sufficiently far) away from it, but that it shows some instability within the locality of that attractor. In assets markets too there are applications. If the differential between the risk adjusted returns on two assets is wide, the profitability of "arbitrage" is higher than when this differential is low due to the existence of fixed transactions costs. As a result the speed of reversion to equilibrium, i.e., the speed with which returns are equalised varies inversely with the size of the differential itself.

The application that motivates our model is that of Sercu et al. (1995) and of Michael et al. (1997). These authors analyse nonlinearities in the PPP relationship. They adopt a null of a unit root for real exchange rates and have an alternative hypothesis of stationarity, namely the long run PPP. Their theory suggests that the larger the deviation from PPP, the stronger the tendency to move back to equilibrium. In the context of our model, this would imply that while $\phi \geqslant 0$ is possible, we must have $\gamma < 0$ and $\phi + \gamma < 0$ for the process to be globally stationary. Under these conditions, the process might display unit root or explosive behaviour in the middle regime for small y_{t-d}^2, but for large y_{t-d}^2, it has stable dynamics and as a result is geometrically ergodic. They claim that the *ADF* test may lack power against such stationary alternatives and one of the contributions of this paper is to provide an alternative test designed to have power against such an ESTAR processes.

More formally, geometric ergodicity and the associated asymptotic stationarity can be established by the drift condition of Tweedie (1975). A variant of the condition states that an irreducible aperiodic Markov chain y_t is geometrically ergodic if there exists constants $\delta < 1$, $B, L < \infty$ and a small set C such that

$$\mathrm{E}[\|y_t\| \mid y_{t-1} = y] < \delta\|y\| + L \quad \forall y \notin C, \tag{5}$$

$$\mathrm{E}[\|y_t\| \mid y_{t-1} = y] \leqslant B \quad \forall y \in C. \tag{6}$$

The concept of the small set is the equivalent of a discrete Markov chain state in a continuous context. For more details see Tweedie (1975), Balke and Fomby (1997) and Kapetanios (1999). Now using this we show under $\theta > 0$ that the condition we need for geometric ergodicity of model (3) is in fact $|\beta + \gamma| < 1$ or $|\phi + \gamma| < 0$. First, if $|\beta + \gamma| < 1$, then there exists some finite $y^* > 0$ such that for all $y < -y^*$ and $y > y^*$, $\beta + \gamma[1 - \exp(-\theta y^2)] < 1$, where $0 < 1 - \exp(-\theta y^2) < 1$ and $\theta > 0$. Define the small set $C = [-y^*, y^*]$. Then, by the finiteness of $\mathrm{E}(|\varepsilon_t|)$, condition (6) is easily seen to be satisfied. We then need to prove that condition (5) holds. But, since

$\beta + \gamma[1 - \exp(-\theta y^2)] < 1$, it follows that

$$E[\|y_t\| \mid y_{t-1} = y] \leqslant \|\beta + \gamma[1 - \exp(-\theta y^2)]\| \|y\| + L$$

for all $y \notin C$ and for some finite L. Therefore, geometric ergodicity is proved for the ESTAR process given by (3).[2]

Following the practice in the literature (e.g. Balke and Fomby, 1997, in the context of TAR models and Michael et al., 1997 in the context of ESTAR models), we impose $\phi=0$ in (4), implying that y_t follows a unit root process in the middle regime. We now consider a null hypothesis that is a special case of a linear unit root which in terms of the above model implies that $\phi = 0$ and $\theta = 0$. Under the alternative hypothesis ($\phi = 0$ but $\theta > 0$), then y_t follows a nonlinear but globally stationary process provided that $-2 < \gamma < 0$, which we assume holds. In practice, there is likely to be little theoretical or prior guidance as to the value of the delay parameter d. We would suggest that d be chosen to maximise goodness of fit over $d = \{1, 2, \ldots, d_{\max}\}$. In what follows, to clarify ideas and in keeping with empirical practice to date (as in for example Michael et al.), we set $d = 1$.

Imposing $\phi = 0$ and $d = 1$ gives our specific ESTAR model (4) as

$$\Delta y_t = \gamma y_{t-1}\{1 - \exp(-\theta y_{t-1}^2)\} + \varepsilon_t. \tag{7}$$

We might expect that the standard linear ADF test may not be very powerful when the true process is stationary but nonlinear, so we will develop a testing framework for this context. Our test directly focuses on a specific parameter, θ, which is zero under the null and positive under the alternative. Hence we test

$$H_0: \theta = 0 \tag{8}$$

against the alternative

$$H_1: \theta > 0. \tag{9}$$

Obviously, testing the null hypothesis (8) directly is not feasible, since γ is not identified under the null. See for example Davies (1987). To overcome this problem, we follow Luukkonen et al. (1988), and derive a t-type test statistic. If we compute a first-order Talyor series approximation to the ESTAR model under the null, we get the auxiliary regression

$$\Delta y_t = \delta y_{t-1}^3 + \text{error}. \tag{10}$$

This suggests that we could obtain the t-statistic for $\delta = 0$ against $\delta < 0$ as[3]

$$t_{\text{NL}} = \hat{\delta}/\text{s.e.}(\hat{\delta}), \tag{11}$$

[2] The general case for lags of $p \geqslant 1$ and $d \geqslant 1$ may be similarly proved by defining a Markov chain as $\mathbf{y}_t = (y_{t-1}, \ldots, y_{t-\max(p,d)})$ and carrying out similar steps.

[3] An LM-type test statistic may be obtained via a similar route. See Granger and Teräsvirta (1993) and Teräsvirta (1994) for more details. The advantage of the t-test over the LM-test is that the t-test deals with one sided alternatives of stationarity explicitly, and thus is expected to be more powerful.

where $\hat{\delta}$ is the OLS estimate of δ and s.e.$(\hat{\delta})$ is the standard error of $\hat{\delta}$. Our test is motivated by the fact that the auxiliary regression is testing the significance of the score vector from the quasi-likelihood function of the ESTAR model, evaluated at $\theta=0$. Unlike the case of testing linearity against nonlinearity for the stationary process, the t_{NL} test does not have an asymptotic standard normal distribution.

Theorem 2.1. *Under the null of a unit root* (8) *the* t_{NL} *statistic defined by* (11) *has the following asymptotic distribution*:

$$t_{\mathrm{NL}} \Rightarrow \frac{\{\frac{1}{4}W(1)^4 - \frac{3}{2}\int_0^1 W(r)^2\,\mathrm{d}r\}}{\sqrt{\int W(r)^6\,\mathrm{d}r}}, \tag{12}$$

where $W(r)$ *is the standard Brownian motion defined on* $r\in[0,1]$. *Under the alternative hypothesis* (9) *with the ESTAR model* (7), *the* t_{NL} *statistic is consistent.*

Proof. See the appendix.

To accommodate stochastic processes with nonzero means and/or linear deterministic trends, we need the following modifications. In the case where the data has nonzero mean, i.e., where $x_t = \mu + y_t$, we use the de-meaned data $y_t = x_t - \bar{x}$, where $\bar{x}$ is the sample mean. In this case the asymptotic distribution of the t_{NL} statistic is basically the same as (12), except that $W(r)$ is replaced by the de-meaned standard Brownian motion $\tilde{W}(r)$ defined on $r\in[0,1]$. Similarly, for the case with nonzero mean and nonzero linear trend, i.e., where $x_t=\mu+\delta t+y_t$, we use the de-meaned and de-trended data $y_t = x_t - \hat{\mu} - \hat{\delta}t$, where $\hat{\mu}$ and $\hat{\delta}$ are the OLS estimators of μ and δ. Now the associated asymptotic distributions are such that $W(r)$ is replaced by the de-meaned and de-trended standard Brownian motion $\hat{W}(r)$.

In nonlinear models, the modelling of intercepts and trends is not straightforward. In particular, our use of de-meaned and/or de-trended data implies a specific view of the way that the intercept and/or trend enter the model under the alternative. We should stress that although finite sample power may be affected, our suggested testing procedure is asymptotically similar with respect to intercepts or time trends.

Asymptotic critical values of the t_{NL} statistics for the above three cases, denoted Case 1, Case 2 and Case 3, respectively, have been tabulated via stochastic simulations with $T = 1,000$ and 50,000 replications, and presented in Table 1.

Table 1
Asymptotic critical values of t_{NL} statistic

Fractile (%)	Case 1	Case 2	Case 3
1	−2.82	−3.48	−3.93
5	−2.22	−2.93	−3.40
10	−1.92	−2.66	−3.13

Note: Case 1, Case 2 and Case 3 refer to the underlying model with the raw data, the de-meaned data and the de-trended data, respectively.

We now consider the more general case where the errors in (7) are serially correlated. Assuming that these serially correlated errors enter in a linear fashion, then following the well-established Dickey and Fuller (1979) and Said and Dickey (1984) corrections, we may extend model (7) to

$$\Delta y_t = \sum_{j=1}^{p} \rho_j \Delta y_{t-j} + \gamma y_{t-1}\{1 - \exp(-\theta y_{t-1}^2)\} + \varepsilon_t, \tag{13}$$

where $\varepsilon_t \sim iid(0, \sigma^2)$. The t_{NL} statistic for testing $\theta = 0$ in this set up is given by the same expression as in (11), where $\hat{\delta}$ is the OLS estimate of δ and s.e.$(\hat{\delta})$ is the standard error of $\hat{\delta}$ obtained from the following auxiliary regression with the p augmentations:

$$\Delta y_t = \sum_{j=1}^{p} \rho_j \Delta y_{t-j} + \delta y_{t-1}^3 + \text{error}. \tag{14}$$

Theorem 2.2. *Consider the nonlinear ADF regression* (13). *Under the null* (8) *the* t_{NL} *statistic obtained from* (14) *has the same asymptotic distribution as obtained under nonserially correlated errors. Under the alternative hypothesis, the* t_{NL} *statistic is consistent.*

Proof. See the appendix.

In practice, the number of augmentations p must be selected prior to the test. We would propose that standard model selection criteria or significance testing procedure be used for this purpose because under the null of a linear model, the properties of these criteria are well understood.[4] See Ng and Perron (1995) for a thorough discussion on the lag selection in the linear univariate models. In fact this is the approach taken in most STAR models, see, for example, van Dijk et al. (2002).

3. Small sample properties

In this section, we undertake a small-scale Monte Carlo investigation of the small sample size and power performance of our proposed t_{NL} test, and compare it with that of the Dickey–Fuller test.[5] We also consider the F-test recently proposed by Enders and Granger (1998) that is designed to have power against the two-regime stationary

[4] The augmentations may enter in a nonlinear way. In such cases, we would view the use of linear augmentations as a first-order approximation to the underlying dynamics rather than a strict view about the exact nature of the dynamic process itself. Of course the criteria and resultant lag selection will have implications for power but this problem plagues all unit root tests that use parametric corrections for auto-correlation. Alternatively, we would follow the semi-parametric correction method advanced by Phillips and Perron (1988).

[5] In a further case, we relax our maintained assumption, $\phi = 0$ in (7), and consider an F-type test for $\phi = \delta = 0$ in the following auxiliary regression: $\Delta y_t = \phi y_{t-1} + \delta y_{t-1}^3 + \varepsilon_t$. This F-type test tends to under-reject somewhat. More importantly, as expected, its power is relatively poor as compared to the t_{NL} test in most cases considered. These simulation results are available upon request.

Table 2
The size of alternative tests

	Case 1			Case 2			Case 3		
	t_{NL}	EG	DF	t_{NL}	EG	DF	t_{NL}	EG	DF
$\rho=0$									
$T=50$	0.042	0.029	0.051	0.044	0.019	0.055	0.047	0.013	0.059
$T=100$	0.045	0.037	0.049	0.046	0.028	0.050	0.048	0.028	0.058
$T=200$	0.050	0.040	0.045	0.046	0.032	0.047	0.048	0.035	0.051
$\rho=0.5$									
$T=50$	0.046	0.036	0.053	0.051	0.018	0.054	0.065	0.013	0.066
$T=100$	0.046	0.043	0.050	0.052	0.029	0.050	0.057	0.028	0.060
$T=200$	0.049	0.045	0.047	0.047	0.032	0.047	0.053	0.039	0.056

Note: To compute the rejection probabilities, the data under the null is generated by (16).

TAR processes given by

$$\Delta y_t = \begin{Bmatrix} \phi_1 y_{t-1} + u_t & \text{if } y_{t-1} \leqslant 0 \\ \phi_2 y_{t-1} + u_t & \text{if } y_{t-1} > 0 \end{Bmatrix}, \quad t = 1, 2, \ldots, T. \tag{15}$$

The Enders and Granger F-statistic (hereafter EG) tests for $\phi_1 = \phi_2 = 0$ in (15).[6]

Since the tests are similar with respect to intercepts and/or time trends and for the sake of simplicity, we set trend and intercept parameters to zero. In the first set of experiments we focus on the size of the tests and thus construct the null model with possibly serially correlated errors by

$$y_t = y_{t-1} + \varepsilon_t \quad \text{with } \varepsilon_t = \rho\varepsilon_{t-1} + u_t, \tag{16}$$

where u_t is drawn from the standard normal distribution, and we consider $\rho = \{0, 0.5\}$.

Secondly, in order to evaluate the power of alternative tests against globally stationary ESTAR processes, we generate the DGP as follows:

$$\Delta y_t = \gamma y_{t-1}[1 - \exp(-\theta y_{t-1}^2)] + \varepsilon_t, \tag{17}$$

where $\varepsilon_t \sim \mathrm{N}(0, 1)$. In particular, we choose a broad range of parameter values for $\gamma = \{-1.5, -1, -0.5, -0.1\}$ and $\theta = \{0.01, 0.05, 0.1, 1\}$ for a general power comparison.

For each experiment, we have computed the rejection probability of the null hypothesis. The nominal size of the tests is set at 0.05, the number of replications at 20,000 and the sample size is considered for $T = 50, 100, 200$.

Table 2 presents results on the size of the various tests. All sizes for the t_{NL} test are close to the nominal level of 5% even in the presence of serially correlated errors

[6] Enders and Granger (1998) implicitly assume that a threshold value is known. Like our testing procedure, the de-meaned data and the de-trended data are then used, respectively, to accommodate stochastic processes with nonzero means and/or linear deterministic trends. The distribution of the F-statistic for the three cases are also nonstandard, and their asymptotic critical values are tabulated via simulation. The 95% critical values are 3.75, 4.56 and 6.08 for Cases I, II and III, respectively.

(in this case we use the correct specification with one augmentation for all the tests), whereas the *EG* test tends to under-reject when the sample size is relatively small.

We next turn to the power performance of the tests, which is summarised in Table 3(a)–(c). A general finding is that our suggested t_{NL} test is relatively more powerful when θ is relatively small regardless of the values of γ. For example, when looking at Table 3(b) with $\theta = 0.01$ and $\gamma = -1$, the powers of the t_{NL} test are 0.183 and 0.488 for $T = 50$ and 100, whereas those of the *DF* test are 0.160 and 0.341. But, this power gain decreases as θ increases. In fact, when θ is sufficiently large (e.g., $\theta = 1$), the power of the Dickey–Fuller test dominates. This is not a surprising finding because as θ grows large, the model becomes approximately linear. Notice also in this case that the power of all tests are close to 1 unless γ is very small. Interestingly, in our application below we find that the estimates of θ (which we obtain under the constraint that $\gamma = -1$) are indeed quite small, ranging as they do between 0.01 and 0.1. Our simulation result clearly shows that over this range of θ our suggested tests are more powerful than the *DF* tests. In general, it is hard to get an exact definition of "small" and "large" θ because it is not a scale-free parameter. But, it is easily seen that given the values of σ^2 and γ, as θ grows, $E(e^{-\theta y_{t-1}^2})$ decreases and the series becomes less persistent, where we note that the term $e^{-\theta y_{t-1}^2}$ here measures the size of the largest root of the series at time t. For example, we find via simulation that for $\sigma^2 = 1$ and $\gamma = -1$ with $T = 1,000$, the average sample values of $e^{-\theta y_{t-1}^2}$ are 0.95, 0.89, 0.83 and 0.34 for $\theta = 0.01$, 0.05, 0.1 and 1, respectively. The power performance of the tests aforementioned then implies that the t_{NL} test performs best relative to the *ADF* test in the region of the null, where the series is relatively more persistent (in this regard, the relatively small value of θ is associated with the relatively more persistent series). Considering that most economic time series are likely to be highly persistent or stay near unit root, this might be a useful finding at least empirically.

By contrast, the power of the *EG* test is quite poor in most cases, unless both the sample size and θ are sufficiently large. This implies that the unit root test derived in the nonlinear TAR framework does not seem to be powerful against the nonlinear STAR processes. [7]

To further investigate and thus highlight the relative power performance of alternative tests, we consider a third set of experiments in which the DGP is given by

$$\Delta y_t = \phi y_{t-1} + \gamma y_{t-1}[1 - \exp(-\theta y_{t-1}^2)] + \varepsilon_t, \tag{18}$$

where $\phi = 0.1$, $\varepsilon_t \sim N(0,1)$, $\gamma = \{-1.5, -1, -0.5\}$ and $\theta = \{0.01, 0.05, 0.1, 1\}$. As discussed earlier, under these parameter values, the process is locally explosive but still globally geometrically ergodic. Though this case is not explicitly covered in our theoretical discussion (i.e., under the maintained assumption $\phi = 0$), it is also of interest for empirical work. Table 4 presents the simulation results for this experiment. The picture is similar to before, but now the power gain of our t_{NL} test over the *DF* test is

[7] This is also partially due to the *EG* test being under-sized. Since its power is much lower than the t_{NT} test in most cases we do not attempt to compute the size-adjusted power.

Table 3

	$\theta=0.01$			$\theta=0.05$			$\theta=0.1$			$\theta=1$		
	t_{NL}	EG	DF	t_{NL}	EG	DF	t_{NL}	EG	DF	t_{NL}	EG	DF
(a) The power of alternative tests: Case 1												
$\gamma=-1.5$												
$T=50$	0.629	0.067	0.470	0.984	0.668	0.995	0.999	0.968	1.0	1.0	1.0	1.0
$T=100$	0.980	0.461	0.984	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
$T=200$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
$\gamma=-1.0$												
$T=50$	0.458	0.047	0.298	0.940	0.338	0.953	0.989	0.769	0.998	1.0	1.0	1.0
$T=100$	0.930	0.249	0.895	1.0	0.992	1.0	1.0	1.0	1.0	1.0	1.0	1.0
$T=200$	1.0	0.973	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
$\gamma=-0.5$												
$T=50$	0.232	0.033	0.154	0.706	0.292	0.631	0.869	0.241	0.899	0.968	0.940	1.0
$T=100$	0.695	0.105	0.518	0.990	0.781	0.998	0.999	0.966	1.0	1.0	1.0	1.0
$T=200$	0.992	0.584	0.995	1.0	0.997	1.0	1.0	1.0	1.0	1.0	1.0	1.0
$\gamma=-0.1$												
$T=50$	0.071	0.027	0.073	0.153	0.032	0.130	0.205	0.039	0.181	0.269	0.079	0.347
$T=100$	0.147	0.040	0.112	0.426	0.082	0.345	0.529	0.127	0.521	0.564	0.317	0.780
$T=200$	0.480	0.087	0.329	0.872	0.349	0.920	0.905	0.590	0.984	0.883	0.891	0.998
(b) The power of alternative tests: Case 2												
$\gamma=-1.5$												
$T=50$	0.250	0.085	0.195	0.826	0.521	0.797	0.968	0.911	0.984	1.0	1.0	1.0
$T=100$	0.692	0.377	0.543	0.998	0.999	1.0	1.0	1.0	1.0	1.0	1.0	1.0
$T=200$	0.992	0.995	0.999	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
$\gamma=-1.0$												
$T=50$	0.183	0.060	0.160	0.626	0.264	0.519	0.855	0.617	0.868	0.997	1.0	1.0
$T=100$	0.488	0.225	0.341	0.980	0.969	0.993	0.999	1.0	1.0	1.0	1.0	1.0
$T=200$	0.955	0.910	0.963	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
$\gamma=-0.5$												
$T=50$	0.108	0.035	0.121	0.296	0.077	0.242	0.469	0.192	0.408	0.794	0.845	0.968
$T=100$	0.244	0.118	0.188	0.781	0.535	0.731	0.936	0.900	0.970	0.993	1.0	1.0
$T=200$	0.725	0.474	0.590	0.997	0.975	1.0	1.0	1.0	1.0	1.0	1.0	1.0
$\gamma=-0.1$												
$T=50$	0.060	0.020	0.080	0.080	0.027	0.100	0.090	0.035	0.111	0.106	0.052	0.142
$T=100$	0.086	0.040	0.093	0.146	0.081	0.146	0.191	0.110	0.187	0.255	0.212	0.331
$T=200$	0.156	0.100	0.142	0.433	0.270	0.370	0.556	0.457	0.577	0.620	0.767	0.854
(c) The power of alternative tests: Case 3												
$\gamma=-1.5$												
$T=50$	0.164	0.048	0.171	0.655	0.267	0.605	0.900	0.676	0.926	1.0	1.0	1.0
$T=100$	0.441	0.220	0.361	0.985	0.983	0.997	1.0	1.0	1.0	1.0	1.0	1.0
$T=200$	0.948	0.917	0.965	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
$\gamma=-1.0$												
$T=50$	0.120	0.036	0.141	0.429	0.135	0.374	0.691	0.327	0.688	0.985	0.997	1.0
$T=100$	0.277	0.144	0.243	0.910	0.811	0.934	0.992	0.974	0.999	1.0	1.0	1.0
$T=200$	0.815	0.666	0.776	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
$\gamma=-0.5$												
$T=50$	0.082	0.023	0.109	0.181	0.059	0.197	0.287	0.099	0.297	0.611	0.554	0.875
$T=100$	0.141	0.084	0.151	0.535	0.319	0.498	0.784	0.649	0.830	0.969	1.0	1.0
$T=200$	0.440	0.291	0.375	0.975	0.988	0.996	0.996	1.0	1.0	1.0	1.0	1.0

Table 3 (*continued*)

	$\theta = 0.01$			$\theta = 0.05$			$\theta = 0.1$			$\theta = 1$		
	t_{NL}	*EG*	*DF*	t_{NL}	*EG*	*DF*	t_{NL}	*EG*	*DF*	t_{NL}	*EG*	*DF*
$\gamma = -0.1$												
$T = 50$	0.053	0.017	0.078	0.062	0.020	0.092	0.066	0.024	0.099	0.081	0.030	0.119
$T = 100$	0.062	0.035	0.082	0.091	0.061	0.116	0.110	0.076	0.138	0.145	0.123	0.216
$T = 200$	0.098	0.077	0.108	0.230	0.179	0.241	0.320	0.280	0.359	0.410	0.541	0.644

Note: To compute the rejection probabilities, the data under the alternative is generated by (17).

for some cases even greater. For example, when looking at Table 4(b) with $\theta = 0.01$ and $\gamma = -1.5$, the power of the t_{NL} test is 0.220, 0.493 and 0.919 for $T = 50, 100, 200$, respectively, whereas the corresponding numbers for the *DF* test are 0.211, 0.247, 0.782.

4. Empirical application: real interest rates and real exchange rates

The apparent unit root behaviour in real interest rates and real exchange rates has become an awkward puzzle for economists. We argued above that transactions costs in financial assets markets are likely to lead to nonlinear speeds of convergence to equilibrium of rates of return. In the context of real interest rates, the Fisher hypothesis predicts that the long run equilibrium value stays around common constant. The apparent nonstationarity of real interest rates clearly violates this hypothesis creating the interest rate paradox referred to above, see Rose (1988) and Barro and Sala-i-Martin (1990) for a further discussion. The difficulty of rejecting a unit root in real exchange rates also implies similar problems because nonstationarity here implies potentially unbounded gains from arbitrage in traded goods.

Owing to transaction costs and other frictions, it is quite plausible that the more these variables deviate from their equilibrium values, the larger will be the investment/arbitrage adjustment flows that drive them back again. If so, the results in this paper suggest that the failure to reject the unit root may be due to the lack of power of the standard *ADF* test traditionally used in this context, and thus we suggest to use the t_{NL} procedure as well. Here we apply the t_{NL} test to ex post real interest rates and bilateral real exchange rates from the eleven major OECD economies and we also estimate the alternative ESTAR models.

Quarterly data on ex post short term real interest rates,[8] and real bilateral exchange rates with the US dollar for the major economies of the EU (France, Germany, Italy, the Netherlands, Spain and the UK), N.America (Canada and the US), Australasia (Australia and New Zealand) and Japan were collected from the International Financial

[8] Because the difference between ex ante and ex post rates is a forecast error, which in most economic worlds is a stationary nonpersistent process (white noise under rational expectations), this measurement error is unlikely to have a profound impact on our inferences.

Table 4

	$\theta=0.01$			$\theta=0.05$			$\theta=0.1$			$\theta=1$		
	t_{NL}	*EG*	*DF*	t_{NL}	*EG*	*DF*	t_{NL}	*EG*	*DF*	t_{NL}	*EG*	*DF*
(a) The power of alternative tests: Case 1												
$\gamma=-1.5$												
$T=50$	0.545	0.019	0.120	0.979	0.393	0.971	0.998	0.898	0.999	1.0	1.0	1.0
$T=100$	0.974	0.068	0.661	1.0	0.997	1.0	1.0	1.0	1.0	1.0	1.0	1.0
$T=200$	1.0	0.768	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
$\gamma=-1$												
$T=50$	0.321	0.014	0.047	0.917	0.119	0.798	0.983	0.506	0.987	1.0	1.0	1.0
$T=100$	0.898	0.028	0.237	1.0	0.806	1.0	1.0	0.999	0.999	1.0	1.0	1.0
$T=200$	1.0	0.191	0.974	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
$\gamma=-0.5$												
$T=50$	0.066	0.012	0.015	0.582	0.027	0.223	0.786	0.070	0.597	0.916	0.763	0.997
$T=100$	0.458	0.013	0.017	0.977	0.157	0.877	0.996	0.636	0.997	0.998	1.0	1.0
$T=200$	0.989	0.019	0.139	1.0	0.965	1.0	1.0	1.0	1.0	1.0	1.0	1.0
(b) The power of alternative tests: Case 2												
$\gamma=-1.5$												
$T=50$	0.220	0.074	0.211	0.763	0.333	0.594	0.951	0.794	0.948	1.0	1.0	1.0
$T=100$	0.493	0.215	0.247	0.994	0.986	0.997	1.0	1.0	1.0	1.0	1.0	1.0
$T=200$	0.919	0.678	0.782	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
$\gamma=-1$												
$T=50$	0.180	0.042	0.217	0.519	0.168	0.316	0.787	0.396	0.683	0.992	0.999	1.0
$T=100$	0.307	0.194	0.244	0.943	0.777	0.909	0.996	0.996	0.999	1.0	1.0	1.0
$T=200$	0.734	0.358	0.351	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
$\gamma=-0.5$												
$T=50$	0.155	0.022	0.221	0.205	0.080	0.191	0.323	0.117	0.236	0.646	0.589	0.849
$T=100$	0.254	0.127	0.351	0.554	0.234	0.307	0.819	0.529	0.703	0.970	1.0	1.0
$T=200$	0.345	0.308	0.316	0.961	0.894	0.955	0.998	1.0	1.0	1.0	1.0	1.0
(c) The power of alternative tests: Case 3												
$\gamma=-1.5$												
$T=50$	0.167	0.043	0.190	0.584	0.181	0.436	0.860	0.503	0.832	1.0	1.0	1.0
$T=100$	0.309	0.168	0.247	0.969	0.886	0.968	0.999	0.999	1.0	1.0	1.0	1.0
$T=200$	0.777	0.458	0.520	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
$\gamma=-1$												
$T=50$	0.137	0.026	0.184	0.344	0.101	0.268	0.601	0.205	0.498	0.969	0.989	1.0
$T=100$	0.219	0.142	0.234	0.818	0.505	0.694	0.977	0.942	0.987	1.0	1.0	1.0
$T=200$	0.491	0.299	0.326	0.998	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
$\gamma=-0.5$												
$T=50$	0.113	0.015	0.173	0.145	0.046	0.176	0.197	0.070	0.209	0.442	0.306	0.666
$T=100$	0.192	0.072	0.280	0.322	0.169	0.259	0.580	0.314	0.476	0.899	0.988	0.998
$T=200$	0.265	0.257	0.318	0.854	0.640	0.734	0.981	0.986	0.996	0.998	1.0	1.0

Note: To compute the rejection probabilities, the data under the alternative is generated by (18).

Statistics CD-Rom (2001 release date), covering the period 1957(1)–2000(3) for interest rates and 1957(1)–1998(4) for exchange rates. The price deflators used throughout were consumer price indices and the nominal interest rate variables used were treasury bill rates for Canada, the UK and US, call money rates for France, Germany, Japan and Spain, and discount rates for Italy, the Netherlands and New Zealand.[9]

Figs. 1 and 2 plot the real interest rates and real exchange rates. The theories predict that the series should be untrended and this appears generally to be the case with two notable exceptions, namely, the Canadian and Japanese real exchange rates.[10] Both sets of series are highly volatile and persistent. In the case of real exchange rates, for example, volatility increases after the break up of the Bretton Woods system in 1972.

Because neither theory nor empirics support the idea of a trend in either real interest rates or real exchange rates, the relevant t_{NL} statistic is that based on de-meaned data. An $AR(8)$ regression model for Δy_t was estimated for each series, and insignificant augmentation terms were excluded. Then, regression (13) with selected augmentations was estimated to compute the test statistics. The test and estimation results are summarised in Table 5.

To examine the issue of time varying volatility more generally, we also computed Breusch Pagan χ^2 tests for ARCH effects and these are presented in columns 2 and 7 of Table 5. This confirms that as is often the case with financial variables, ARCH effects are prominent in several series. This may raise two problems. First, heteroscedasticity will interfere with inference on the appropriate number of augmentations to be included for the test. Second, the power of both the t_{NT} and ADF tests may be affected, though the asymptotic size of both tests is not affected by the existence of heteroscedasticity. We addressed the first issue by using heteroscedastic consistent standard errors in regressions determining the augmentation terms, but we can do little about the second issue. However, there is no prior reason to suppose that the power of our test is less or more affected by heteroscedasticity than that of the ADF test, so in this sense comparison with the ADF test remains "fair".

For real interest rates, the ADF test rejects in 2 out of 11 cases at the 5% significance level (Germany and Japan), and rejects another (Spain) at the 10% significance level. By contrast the t_{NL} test yields 5 significant statistics at the 5% level and 2 additional cases at the 10% level, giving much stronger overall support to the long run Fisher

[9] Interest rate data for Spain prior to 1974 was not available so the sample for this country is restricted to 1974(1)–2000(3). Similarly, the Euro-zone exchange rates were only available until 1998(4). The last 24 quarterly data points for the Netherlands had to be extrapolated using Treasury Bill Rate movements, and the last 4 quarters for the UK and US had to be extrapolated using call money rates. For Australia the 2 year government bond yield was used because it was the only short to medium term rate for which a sufficiently long time series was available. In general the decision favouring one rate over another was entirely driven by availability over the relevant quarterly time span but casual experimentation with the data showed that all of the possible short term rates were very coherent within each respective country.

[10] The Canadian dollar and Japanese Yen real exchange rates exhibit secular depreciation and appreciation respectively against the US dollar during the sample period and any trend in real exchange rates would constitute a rejection of long run PPP in its simplest form. However, as has often been pointed out in the literature, the existence of secular differential rates of productivity growth between two countries induces a trend in PPP relationships—the so-called Balassa–Samuelson effect. To investigate the existence and nature of this trend for Japan and Canada would require data on relative productivity growth and this is beyond the scope of the current paper. In the meantime, we should note that it is unlikely that the unit root hypothesis in either of these two cases will be rejected by either the ADF or t_{NL} tests considered below.

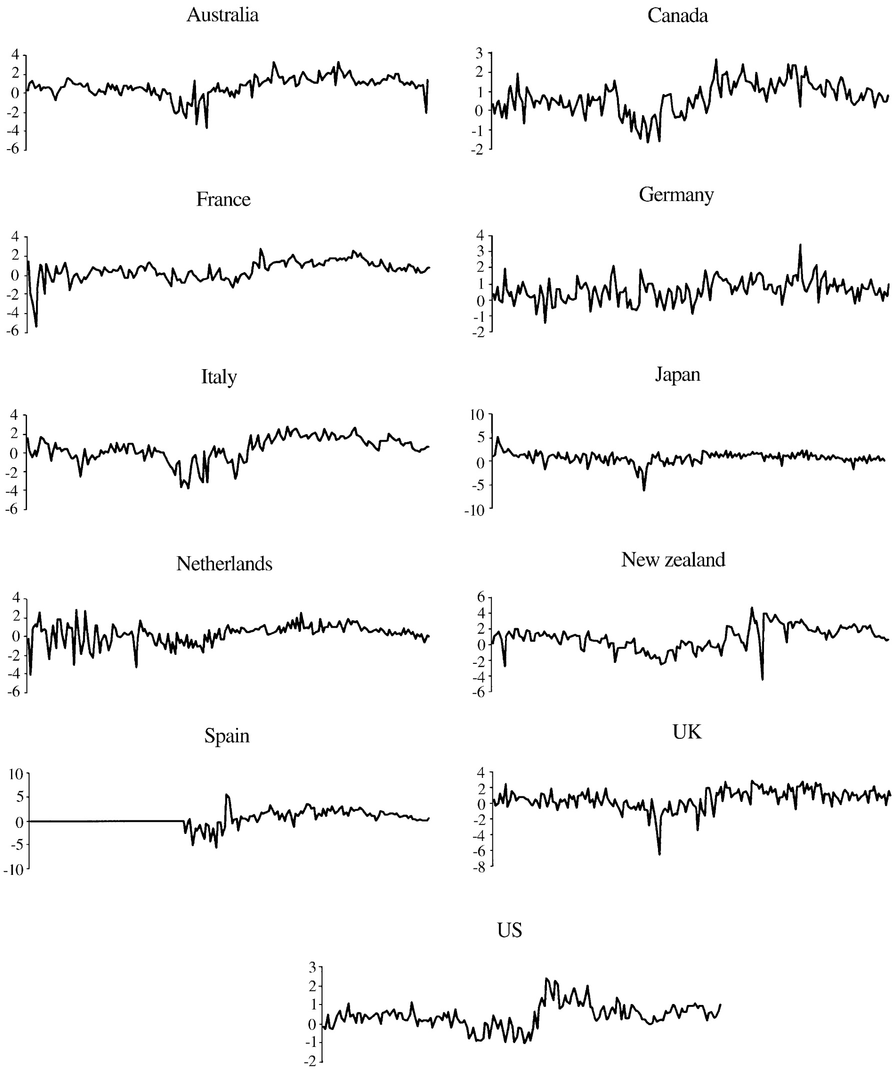

Fig. 1. Real interest rates.

hypothesis. For real exchange rates, the *ADF* test is unable to reject a unit root for *any* of the countries at the 5% significance level, although in the case of New Zealand and Italy it can reject at the 10% significance level. Again the t_{NL} test improves the situation rejecting the null in 5 cases at the 5% significance level and another at the 10% significance level, giving stronger support to simple long run PPP.

Table 5 also displays the estimation results of ESTAR models. Initial estimation found γ to be very poorly identified, a result that has been found elsewhere (e.g., Taylor et al., 2001). As a consequence, we follow the procedure of the numerical

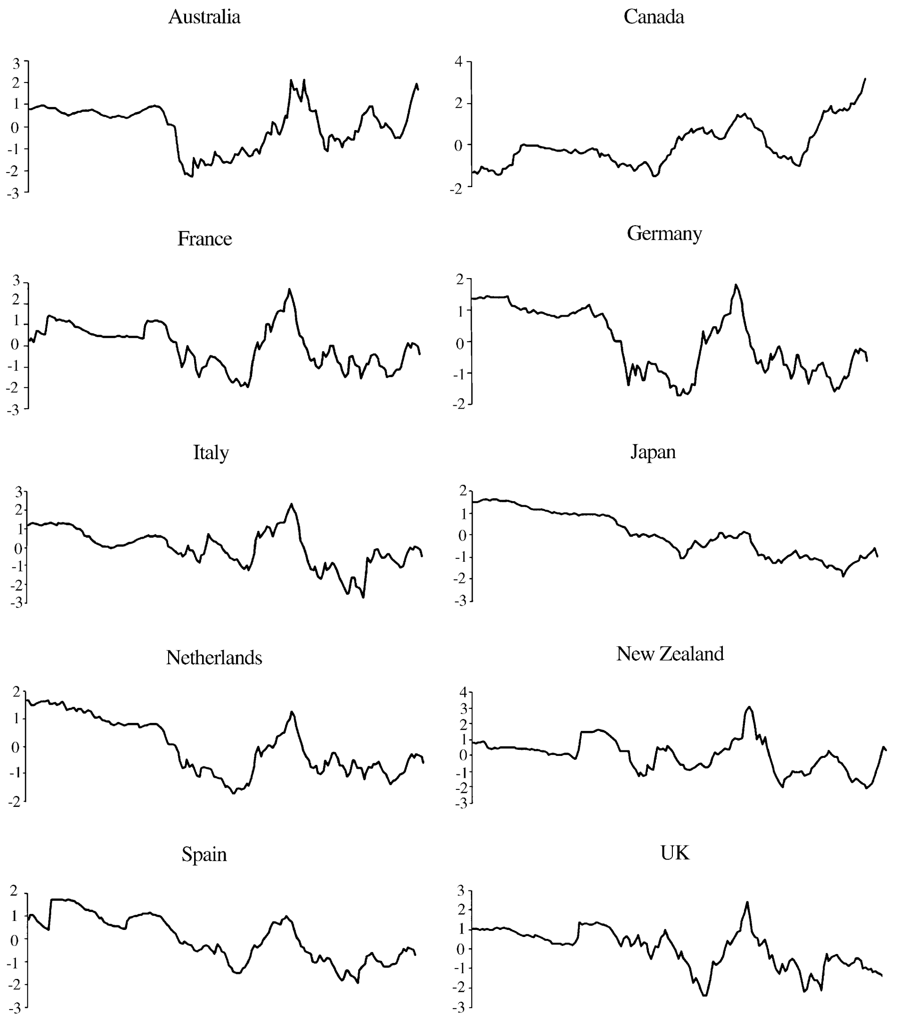

Fig. 2. Real bilateral exchange rates with the US dollar.

section above and set γ to minus unity.[11] All estimates of θ except for the Canadian real exchange rate are correctly signed.[12] Although the t-ratio does not provide a valid

[11] We experimented with different values for γ, but found that it made no qualitative difference to the results in terms of the significance of θ. For example, whilst halving γ to -0.5 resulted broadly speaking in a trebling of the value for θ, the significance or more correctly, the numerical value of its t-ratio, was not greatly affected.

[12] The perverse coefficient for this case that implies explosive behaviour under the alternative was perhaps to be expected for reasons outlined in the footnote 10. In fact, we have also reestimated the regression model with de-trended data for both the Canadian dollar and Japanese Yen, and found that both t_{NL} and *ADF* tests cannot reject the null of unit root for the former, but can reject it for the latter. In the case of Canada, detrending also gets rid of the perverse sign of $\hat{\theta}$.

Table 5
Unit root test results for 11 real interest rates and 10 bilateral real exchange rates

	Real interest rates					Real exchange rates				
Country	χ^2_8	*ADF*	t_{NL}	$\hat{\theta}$	s.e.($\hat{\theta}$)	χ^2_8	*ADF*	t_{NL}	$\hat{\theta}$	s.e.($\hat{\theta}$)
AU	26.0*	−2.07	−2.06	0.020	0.011	22.7*	−2.55	−3.95*	0.032	0.009
CA	6.22	−2.45	−2.72**	0.045	0.019	4.76	−0.61	0.54	−0.002	0.004
FR	2.90	−2.34	−4.23*	0.100	0.030	13.3	−2.54	−2.93*	0.022	0.008
GE	13.7**	−3.44*	−2.96*	0.079	0.045	22.8*	−1.95	−2.70**	0.024	0.009
IT	30.6*	−2.47	−3.51*	0.037	0.013	5.20	−2.59**	−4.07*	0.024	0.007
JA	26.3*	−4.27*	−2.26	2.33	1.58	7.53	−1.18	−1.49	0.006	0.004
NE	22.2*	−1.58	−2.71**	0.077	0.041	16.9*	−1.83	−1.89	0.013	0.007
NZ	48.6*	−1.75	−6.14*	0.080	0.025	10.2	−2.87**	−3.37*	0.020	0.006
SP	39.6*	−2.70**	−3.40*	0.019	0.009	7.75	−1.59	−2.22	0.015	0.007
UK	40.9*	−2.24	−0.55	0.003	0.005	2.50	−1.78	−3.61*	0.029	0.009
US	23.9*	−2.56	−1.93	0.024	0.013					

Note: χ^2_8 stands for the Breusch–Pagan test for ARCH(8) effects with eight degrees of freedom. The t_{NL} and *DF* statistics are computed using the de-meaned data in a regression model (13) with a maximum of eight augmentations, where the insignificant augmentation terms in a companion $AR(8)$ model for Δy_t were excluded. In all cases * and ** denote significance at 5% and 10% level. $\hat{\theta}$ is estimated imposing $\gamma = -1$, and s.e.($\hat{\theta}$) indicates a standard error.

significance test in the usual way, the 95% asymptotic confidence intervals computed under the alternative for θ include zero in only 6 out of 21 cases; namely, the German, UK and US real interest rates and the Canadian, Japanese and Dutch real exchange rate (with the last of these only marginally including zero). It is interesting to note that, with the exception of the Spanish real exchange rate, whenever t_{NL} is significant at either the 5% or 10% level, then the t-ratio for $\hat{\theta}$ exceeds 1.96 and vice versa. Finally, with the exception of Japanese real interest rates, all estimates of θ lie in the range $[0.01, 0.1]$. This was exactly the range of θ that formed the focus for the numerical experiments on test power in Section 3 above. Thus, it follows that our simulations have indeed focused on an empirically meaningful range of θ.[13]

To get a feel for the influence of the nonlinear terms on persistence of the series, we have plotted one minus the transition function, $e^{-\hat{\theta} y_{t-1}^2}$, in each case. In the absence of augmentation terms, this term measures the single root of the time series at time t conditional on a y_{t-1} value.[14] Note that when each series is at its mean, the root is unity by construction, and the series is locally nonstationary. But, when it is away from its mean, the series exhibits mean reversion. Figs. 3 and 4 present these results.

[13] We noted above that θ is not scale invariant. However, the standard deviation of our data (unity) is roughly the same as that of the data in the numerical experiments (between unity and two).

[14] The plots are drawn for y_{t-1} values in the range $(-2, 2)$ because apart from a very few observations in a few cases, each of the series varied within two standard deviations of its respective mean implying that the normalised data lay (more or less) in the range ± 2.

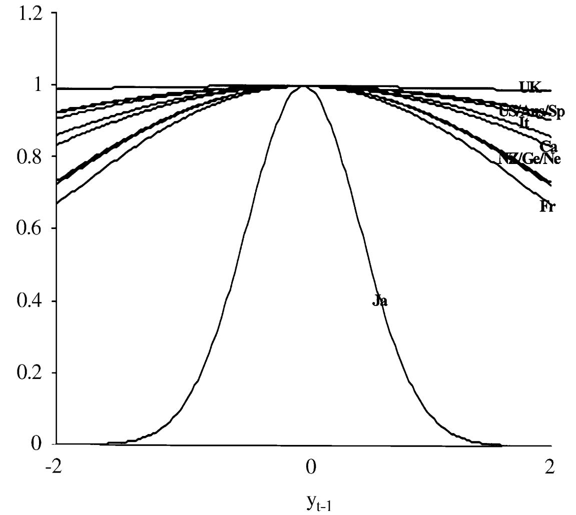

Fig. 3. One minus the estimated transition function ($e^{-\hat{\theta} y_{t-1}^2}$) for real interest rates.

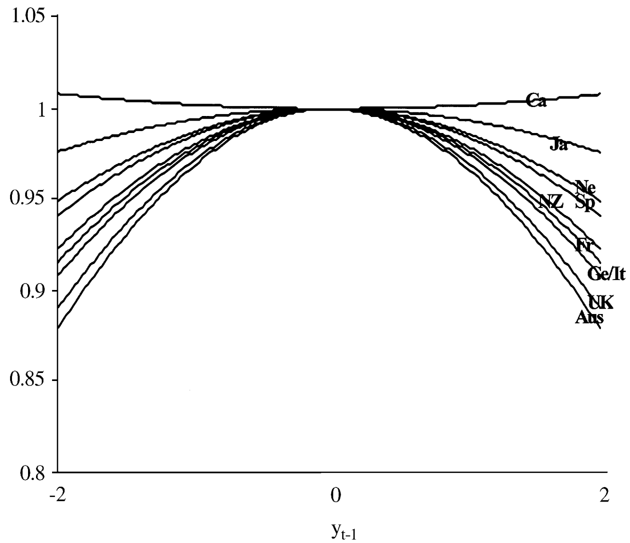

Fig. 4. One minus the estimated transition function ($e^{-\hat{\theta} y_{t-1}^2}$) for real exchange rates.

From Fig. 3 we find that the Japanese real interest rate has the most important nonlinear mean-reversion effects. For example, a two standard deviation shock away from its mean would be followed by a near 100% correction back towards its mean during the next period. We should note, however, that the standard error of $\hat{\theta}$ here is so large that the 95% confidence interval includes zero, and that the t_{NL} test fails to reject a unit root. After Japan, the French real interest rate appears to have the most sizeable nonlinear effect with the corresponding reversion to mean following a 2 standard deviation shock being about 33%. By contrast the Japanese real exchange rate displays little nonlinear adjustment with less than 5% of a two standard deviation shock away from its mean being corrected the following period.

It is also interesting to note that although our t_{NL} test yields p-values that are on the whole less than the *ADF*, the case of Japanese interest rates shows a glaring exception. Here, the p-value for the *ADF* is less than 1% but for t_{NL} it exceeds 10%. However, this case is one where $\hat{\theta}$ is large and persistence is very low whilst the other cases had, on the whole, small values of $\hat{\theta}$ and thus displayed high persistence. This also supports our assertion that t_{NL} performs best relative to the *ADF* in the region of the null and that when we are a long way from the null (high θ and low persistence), the model becomes approximately linear and thus the *ADF* comes back into its own again.

Overall, these results suggest that the t_{NL} test is a useful supplementary statistic to employ in standard unit root testing especially where the series is known to be highly persistent but expected a priori to be stationary. The estimates also suggest that the ESTAR model itself may provide a better alternative to a linear AR model in such cases.

5. Concluding remarks

Empirical univariate analysis of nonstationarity against stationarity has been an integral part of time series econometrics. However, the emphasis of the earlier literature was on the examination of the linear model, implicitly disregarding any possible nonlinearities in the series under consideration. This paper complements other recent studies in trying to fill this vacuum. Its main contribution has been to develop a new unit root test statistic designed to be more powerful against a stationary ESTAR processes than the standard *ADF* test and to establish via Monte Carlo simulations and an empirical application that the suggested testing procedure may be quite useful in practice.

As is always the case when working with nonlinear models there are several generalisations and a number of alternative models that could be analysed in future work. As noted in Section 2, one could consider different types of transition function that allow for asymmetric dynamic adjustment such as the logistic function. Again as noted in Section 2, a more general STAR(p) model could be adopted where all the parameters including coefficients on lagged first differences of the dependent variable are subject to the same nonlinear scheme. For extensions to nonlinear TAR models with unit roots see also Caner and Hansen (2001). Other transition variates could be modelled such as a linear combination of lagged endogenous variables (or first differences), exogenous variables or deterministic time trends. For a survey of extensions in the context of stationary models, see van Dijk et al. (2002). Finally, although our test is univariate, it could be extended to establish the existence of cointegrating equilibrium relationships such as those said to govern real exchange rates. In this regard, a cointegration test based on an error correction model subject to STAR nonlinearity is currently under investigation.

Acknowledgements

We are grateful to the Editor and two anonymous referees for their helpful comments on an earlier version of this paper. Our thanks are also owed to In Choi, Inmoo Kim, Gary Koop and seminar participants at Sungkunkwan University. Partial financial support from the ESRC (grant No. R000223399) is gratefully acknowledged. The usual disclaimer applies.

Appendix A.

A.1. Proof of Theorem 2.1

The asymptotic null distribution of the t_{NL} statistic defined by (11) can be derived as follows. First, under the null (8), $\Delta y_t = \varepsilon_t$, and therefore, we obtain

$$t_{NL} = \frac{\sum_{t=1}^{T} y_{t-1}^3 \varepsilon_t}{\sqrt{\hat{\sigma}^2 \sum_{t=1}^{T} y_{t-1}^6}},$$

where $\hat{\sigma}^2$ is the least-squares estimate of σ^2 from the auxiliary regression. It is easy to show that $\hat{\sigma}^2 \to_p \sigma^2$ under the null, so we only need to find the asymptotics for $\sum_{t=1}^{T} y_{t-1}^3 \varepsilon_t$ and $\sum_{t=1}^{T} y_{t-1}^6$. For the latter, it is easy to show (e.g., Chan and Wei, 1988) that

$$\frac{1}{T^4} \sum_{t=1}^{T} y_{t-1}^6 \Rightarrow \sigma^6 \int W(r)^6 \, \mathrm{d}r,$$

whereas it follows directly from the continuous mapping theorem, weak convergence of stochastic integrals and the semi-martingale property of ε_t (e.g., Hansen, 1992) that

$$\frac{1}{T^2} \sum_{t=1}^{T} y_{t-1}^3 \varepsilon_t \Rightarrow \sigma^4 \int W(r)^3 \, \mathrm{d}W(r) = \sigma^4 \left\{ \frac{1}{4} W(1)^4 - \frac{3}{2} \int_0^1 W(r)^2 \, \mathrm{d}r \right\}.$$

Hence, (12) follows.

Next, under alternative (9), Δy_t, y_{t-1} and y_{t-1}^3 are $I(0)$ and it is easy to show that

$$\frac{1}{T} \sum_{t=1}^{T} y_{t-1}^3 \Delta y_t = \mathrm{O_p}(1); \qquad \frac{1}{T} \sum_{t=1}^{T} y_{t-1}^6 = \mathrm{O_p}(1).$$

Then, $t_{NL} = \mathrm{O_p}(T^{1/2})$. Hence, the t_{NL} statistic diverges to infinity at rate $T^{1/2}$ under the alternative. □

A.2. Proof of Theorem 2.2

Define the $T \times p$ data matrix $\mathbf{Z} = (\Delta \mathbf{y}_{-1}, \ldots, \Delta \mathbf{y}_{-p})$ with $\Delta \mathbf{y}_{-i} = (\Delta y_{-i+1}, \ldots, \Delta y_{T-i})$, and the $T \times T$ idempotent matrix $\mathbf{M}_T = \mathbf{I}_T - \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'$. Now,

$$\widehat{\sigma^2} = \frac{1}{T} \boldsymbol{\varepsilon}' \mathbf{M}_T \boldsymbol{\varepsilon} = \frac{1}{T} \boldsymbol{\varepsilon}' \boldsymbol{\varepsilon} + \mathrm{o_p}(1) \xrightarrow{p} \sigma^2,$$

where $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_T)'$. Furthermore, under the null it is straightforward to show that

$$\frac{1}{T^2}\mathbf{y}_{-1}^{3'}\mathbf{M}_T\boldsymbol{\varepsilon} = \frac{1}{T^2}\mathbf{y}_{-1}^{3'}\boldsymbol{\varepsilon}_t + o_p(1) \Rightarrow \sigma_{LR}^2\sigma^2\left\{\frac{1}{4}W(1)^4 - \frac{3}{2}\int_0^1 W(r)^2\,dr\right\},$$

$$\frac{1}{T^4}\mathbf{y}_{-1}^{3'}\mathbf{M}_T\mathbf{y}_{-1}^3 = \frac{1}{T^4}\mathbf{y}_{-1}^{3'}\mathbf{y}_{-1}^3 + o_p(1) \Rightarrow \sigma_{LR}^6\int W(r)^6\,dr.$$

where $\mathbf{y}_{-1}^3 = (y_0^3, y_1^3, \dots, y_{T-1}^3)'$ and σ_{LR}^2 is the long-run variance of Δy_t under the null. Using these results, under the null we obtain

$$t_{NL} = \frac{\frac{1}{T^2}\mathbf{y}_{-1}^{3'}\mathbf{M}_T\boldsymbol{\varepsilon}}{\sqrt{\hat{\sigma}^2\frac{1}{T^4}\mathbf{y}_{-1}^{3'}\mathbf{M}_T\mathbf{y}_{-1}^3}} = \frac{\frac{1}{T^2}\mathbf{y}_{-1}^{3'}\boldsymbol{\varepsilon}}{\frac{1}{T^4}\mathbf{y}_{-1}^{3'}\mathbf{y}_{-1}^3} + o_p(1),$$

which as we have shown before has the asymptotic distribution given in (12).

Finally, along similar lines in the proof of Theorem 2.1, it is easily seen that the t_{NL} test is consistent under (9). □

References

Abuaf, N., Jorion, P., 1990. Purchasing power parity in the long run. Journal of Finance 45, 157–174.

Balke, N.S., Fomby, T.B., 1997. Threshold cointegration. International Economic Review 38, 627–645.

Barro, R.J., Sala-i-Martin, X., 1990. World real interest rates. In: Blanchard, O.J., Fisher, S. (Eds.), NBER Macroeconomics Annual. MIT Press, Cambridge, pp. 15–61.

Berben, R., van Dijk, D., 1999. Unit root tests and asymmetric adjustment: a reassessment. Unpublished manuscript, Tinbergen Institute, Erasmus University of Rotterdam.

Caner, M., Hansen, B.E., 2001. Threshold autoregression with a near unit root. Econometrica 69, 1555–1596.

Chan, N.H., Wei, C.Z., 1988. Limiting distributions of least squares estimates of unstable autoregressive processes. Annals of Statistics 16, 367–401.

Davies, R.B., 1987. Hypothesis testing when a nuisance parameter is present under the alternative. Biometrika 74, 33–43.

Dickey, D.A., Fuller, W.A., 1979. Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association 74, 427–431.

Edison, H., Kloveland, J., 1987. A quantitative reassessment of the PPP hypothesis. Journal of Applied Econometrics 2, 309–333.

Enders, W., Granger, C.W.J., 1998. Unit root tests and asymmetric adjustment with an example using the term structure of interest rates. Journal of Business and Economics Statistics 16, 304–311.

Frankel, J.A., Rose, A.K., 1996. A panel project on purchasing power parity: mean reversion within and between countries. Journal of International Economics 40, 209–224.

Hansen, B.E., 1992. Convergence to stochastic integrals for dependent heterogeneous processes. Econometric Theory 8, 489–501.

Granger, C.W.J., Teräsvirta, T., 1993. Modelling Nonlinear Economic Relationships. Oxford University Press, Oxford.

Im, K.S., Pesaran, M.H., Shin, Y., 2002. Testing for unit roots in heterogeneous panels. Journal of Econometrics, forthcoming.

Kapetanios, G., 1999. Essays on the Econometric Analysis of Threshold Models. Unpublished Ph.D. Thesis, University of Cambridge.

Kapetanios, G., Shin, Y., 2001. Unit root tests in three-regime SETAR models. Unpublished manuscript, University of Edinburgh.

Lo, M.C., Zivot, E., 2001. Threshold cointegration and nonlinear adjustment to the law of one price. Macroeconomic Dynamics 5, 533–576.

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

Michael, P., Nobay, R.A., Peel, D.A., 1997. Transactions costs and nonlinear adjustment in real exchange rates: an empirical investigation. Journal of Political Economy 105, 862–879.

Mills, T., 1993. The Econometric Modelling of Financial Time series. Cambridge University Press, Cambridge.

Ng, S., Perron, P., 1995. Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag. Journal of the American Statistical Association 90, 268–281.

Pesaran, M.H., Potter, S., 1997. A floor and ceiling model of US output. Journal of Economic Dynamics and Control 21, 661–695.

Phillips, P.C.B., Perron, P., 1988. Testing for a unit root in time series regression. Biometrika 75, 335–346.

Pippenger, M.K., Goering, G.E., 1993. A note on the empirical power of unit root tests under threshold processes. Oxford Bulletin of Economics and Statistics 55, 473–481.

Rose, A.K., 1988. Is the real interest rate stable? Journal of Finance 43, 1095–1112.

Said, S.E., Dickey, D.A., 1984. Testing for unit roots in autoregressive moving average models of unknown order. Biometrika 71, 599–607.

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

Taylor, M.P., Peel, D.A., Sarno, L., 2001. Nonlinear mean-reversion in real exchange rates: towards a solution to the purchasing power parity puzzles. International Economic Review 42, 1015–1042.

Teräsvirta, T., 1994. Specification, estimation and evaluation of smooth transition autoregressive models. Journal of the American Statistical Association 89, 208–218.

Tweedie, R.L., 1975. Sufficient conditions for ergodicity and recurrence of Markov on a general state space. Stochastic Processes and their Applications 3, 385–403.

van Dijk, D., Teräsvirta, T., Franses, P.H., 2002. Smooth transition autoregressive models—a survey of recent developments. Econometric Reviews 21, 1–47.

Wu, Y., 1996. Are real exchange rates nonstationary? Evidence from panel-data tests. Journal of Money Credit and Banking 28, 54–63.