ECONOMETRIC INSTITUTE RESEARCH REPORT EI-9902/A

Unit root tests and asymmetric adjustment
A reassessment*

Robert-Paul Berben†
Department of Applied Economics, University of Nijmegen

Dick van Dijk‡
Tinbergen Institute, Erasmus University Rotterdam

January 20, 1999

Abstract

Standard unit root tests are misspecified in case the variable of interest is stationary but displays asymmetric adjustment towards its long-run equilibrium and, consequently, may suffer from a lack of power against such alternatives. This observation recently has aroused interest in developing test statistics which can be used to test the null hypothesis of a unit root against the alternative of stationarity with asymmetric adjustment. In this paper we reconsider the test statistics put forward by Enders and Granger (1998). We point out an important deficiency of their tests and develop an alternative one which is based on more solid statistical grounds. Monte Carlo experiments demonstrate that our new test outperforms standard unit roots and the tests of Enders and Granger (1998) in terms of power against the alternative of interest. An empirical illustration involving the forward premium is provided to demonstrate the practical usefulness of our test statistic.

Keywords: threshold autoregressive process, unidentified nuisance parameter, asymptotic distribution, forward premium

---

*Part of this research was done while the second author was visiting the University of Chicago Graduate School of Business. The hospitality and stimulating research environment which were provided during this visit are gratefully acknowledged. The first author gratefully acknowledges support from the Economics Research Foundation (ESR), which is part of the Netherlands Organization for Scientific Research (NWO).

†Department of Applied Economics, University of Nijmegen, P.O. Box 9108, NL-6500 HK, Nijmegen, The Netherlands, email: r.berben@bw.kun.nl

‡Tinbergen Institute, Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR, Rotterdam, The Netherlands, email: djvandijk@few.eur.nl (corresponding author)

1 Introduction

Classical tests for a unit root, such as the widely used Dickey-Fuller statistic, are based upon the 'linearity paradigm' which has long dominated the time series literature. By this we mean that these tests assume that under the alternative of stationarity the adjustment process towards the equilibrium (which can be either a constant or deterministic trend) is linear. Evidently, such tests then are misspecified in case the true adjustment process is nonlinear and, consequently, they may have diminished ability to detect such stationary alternatives. Pippenger and Goering (1993) and Balke and Fomby (1997) provide Monte Carlo evidence which clearly demonstrates this.

Given the dominant position which the unit root testing problem has occupied in time series analysis over the last two decades and the recent increase in interest in nonlinear time series models, it comes as no surprise that the problem of testing the composite hypothesis of linearity and nonstationarity of a time series against specific nonlinear and stationary (ergodic) alternatives has received some attention. In particular, testing the null hypothesis of a linear autoregressive process containing a unit root against the alternative of asymmetric adjustment as implied by threshold autoregressive [TAR] processes has been addressed, see Caner and Hansen (1997), González and Gonzalo (1997), and Enders and Granger (1998).

In this paper we reconsider the approach of Enders and Granger (1998). We argue that the test statistic they propose is suboptimal or *ad hoc*, in the sense that it makes use of a biased estimate of the threshold (nuisance) parameter in the model under the alternative hypothesis. Hence it might be expected that this test has inferior power compared to statistics which are based on a consistent estimate of the unidentified threshold. In fact, the simulation evidence presented in Enders and Granger (1998) suggests that their test is even less powerful than the standard Dickey-Fuller statistic. We demonstrate that it is fairly straightforward to obtain an alternative 'consistent' test statistic, and show that this test indeed has better power than the statistic of Enders and Granger (1998) as well as the Dickey-Fuller statistic and some other commonly applied unit root tests.

The rest of this paper is organized as follows. In Section 2, we briefly discuss TAR models in general and the continuous TAR [CTAR] model on which we focus here. This section also serves to develop some basic notation and to discuss the hypothesis of interest. In Section 3 we address the question of testing the unit root hypothesis against the alternative of a stationary CTAR model, both in case the threshold is known and in case it is unknown. We discuss the tests suggested by Enders and Granger (1998), point out their main deficiency, and propose an alternative statistic. We demonstrate that our test statistic is readily obtained as a by-product of the estimation of the CTAR model - this topic therefore is briefly discussed in Section 3.2. We derive the asymptotic distribution of the test statistic, using a novel way to characterize the set of allowable threshold values in the estimation procedure. In Section 4 we examine the power properties of our test, and compare it with the tests of Enders and Granger (1998) and some other unit root tests which are popular among practitioners. In Section 5 we illustrate the use of our test by examining several forward premium series. We find that in the three cases considered, the forward premium is not only stationary, but that adjustment is clearly asymmetric as well. Finally, Section 6 contains some concluding remarks. The proofs of Theorems 1 and 2 are deferred to an appendix.

2 The continuous TAR model

An observed time series $y_t,\ t = 1, \ldots, n$, follows a two-regime threshold autoregressive [TAR] model with threshold variable s_t if it satisfies

$$y_t = \begin{cases} \phi_{10} + \phi_{11} y_{t-1} + \ldots + \phi_{1p} y_{t-p} + \varepsilon_t & \text{if } s_t \leq \gamma, \\ \phi_{20} + \phi_{21} y_{t-1} + \ldots + \phi_{2p} y_{t-p} + \varepsilon_t & \text{if } s_t > \gamma, \end{cases} \tag{1}$$

where γ is the threshold and ε_t a martingale difference sequence with respect to the history of y_t. For the formal statistical theory to be developed in the next section we will assume that ε_t is i.i.d. $(0, \sigma^2)$. Throughout we assume that the pre-sample values $y_0, y_{-1}, \ldots, y_{1-p}$ are given and fixed. The TAR model implies a piece-wise linear autoregressive process for the variable y_t, where the linear process that governs the dynamics at any given point in time depends on the value taken by the threshold variable s_t relative to the threshold γ. Particular interest has been given to the so-called Self-Exciting TAR [SETAR] model, in which the threshold variable is a lagged endogenous value, i.e., $s_t = y_{t-d}$ for certain integer $d > 0$. One of the reasons for the interest in this SETAR model is that it can exhibit nonlinear characteristics such as limit cycles and jump phenomena, see Tong and Lim (1980). For a general treatment of (SE)TAR models we refer to Tong (1990).

Following Chan and Tsay (1998), the SETAR model (1) with $s_t = y_{t-d}$ is said to have a continuous autoregressive function if for all points $X_* = (1, x_{p-1}, \ldots, x_0)$ with $x_{p-d} = \gamma$, it is true that $\Phi_1' X_* = \Phi_2' X_*$, where Φ_1 and Φ_2 denote the vectors of autoregressive parameters in the two regimes, i.e., $\Phi_i = (\phi_{i0}, \phi_{i1}, \ldots, \phi_{ip})$, $i = 1, 2$. It is easily seen that this is equivalent to the requirement that $\phi_{1j} = \phi_{2j} \equiv \phi_j$ for $1 \leq j \neq d \leq p$ and $\phi_{10} + \gamma\phi_{1d} = \phi_{20} + \gamma\phi_{2d}$. The resulting continuous TAR [CTAR] model can be written as

$$y_t = \phi_0 + \sum_{j=1, \neq d}^{p} \phi_j y_{t-j} + \begin{cases} \phi_{1d}(y_{t-d} - \gamma) + \varepsilon_t & \text{if } y_{t-d} \leq \gamma, \\ \phi_{2d}(y_{t-d} - \gamma) + \varepsilon_t & \text{if } y_{t-d} > \gamma, \end{cases} \tag{2}$$

where $\phi_0 = \phi_{10} + \gamma\phi_{1d}$. The model which is considered by Enders and Granger (1998) is in turn a special case of the CTAR model and is obtained from (2) by requiring in addition that the series y_t has a unique equilibrium point, or attractor, equal to the threshold or, equivalently, that the means of the autoregressive processes in both regimes are equal to γ. This is enforced by imposing the constraint $\phi_{i0}/(1 - \sum_{j=1,\neq d}^{p} \phi_j - \phi_{id}) \equiv \gamma$, $i = 1, 2$. In case $p = d = 1$ the resulting model is

$$y_t = \gamma + \begin{cases} \phi_{11}(y_{t-1} - \gamma) + \varepsilon_t & \text{if } y_{t-1} \leq \gamma, \\ \phi_{21}(y_{t-1} - \gamma) + \varepsilon_t & \text{if } y_{t-1} > \gamma. \end{cases} \tag{3}$$

To analyze stationarity and ergodicity of this CTAR model it is necessary to consider the stability properties of the 'skeleton' or deterministic part of the model - which simply is the model (3) with the noise ε_t turned off. As shown by Chan and Tong (1985), asymptotic stability of the skeleton of (3) is guaranteed if $\max(|\phi_{11}|, |\phi_{21}|) < 1$[1]. Together with the assumption that $\{y_t\}$ is an irreducible Markov chain this implies that $\{y_t\}$ is geometrically ergodic.

[1] Obviously, this is equivalent to the requirement that the AR processes in the two regimes are stationary. Chan, Petruccelli, Tong and Woolford (1985) show that stability of the simple first order model considered here actually holds under less restrictive conditions. In particular, the skeleton of the SETAR model (1) with $p = 1$ and $s_t = y_{t-1}$ is asymptotically stable if and only if one of the following conditions holds,

(i) $\phi_{11} < 1, \phi_{21} < 1, \phi_{11}\phi_{21} < 1$;

(ii) $\phi_{11} = 1, \phi_{21} < 1, \phi_{10} > 0$;

(iii) $\phi_{11} < 1, \phi_{21} = 1, \phi_{20} < 0$;

(iv) $\phi_{11} = 1, \phi_{21} = 1, \phi_{20} < 0 < \phi_{10}$;

(v) $\phi_{11}\phi_{21} = 1, \phi_{11} < 0, \phi_{20} + \phi_{21}\phi_{10} > 0$;

Conditions (ii)-(iv) are self-contradicting if applied to the CTAR model (2). As far as condition (v) is concerned, we feel that negative autoregressive parameters are not relevant given our interest in contrasting the CTAR model with a random walk (which of course implies $\phi_{11} = \phi_{21} = 1$). For the same reason we only consider the restrictions on ϕ_{11} and ϕ_{21} individually from (i) and disregard the restriction on their product. See also Chen and Tsay (1991).

Notice that the CTAR model (3) reduces to a linear model if $\phi_{11} = \phi_{21}$ and that the threshold γ is equal to the mean of the process in this case. Furthermore, in case both autoregressive parameters are equal to 1, the model collapses to a driftless random walk and the parameter γ is no longer identified. Enders and Granger (1998) discuss testing the composite null hypothesis of linearity and nonstationarity, which thus is given by $H_0 : \phi_{11} = \phi_{21} = 1$, against the alternative of a stationary and ergodic CTAR model as given in (3). In the next section we discuss the test statistics which they propose to test this hypothesis, which we denote as the unit root hypothesis in the following for sake of brevity, in more detail.

At this point it is good to note that Caner and Hansen (1997) and González and Gonzalo (1997) analyze testing for a unit root against a stationary TAR alternative in the more general two-regime CTAR model (2), with one important difference, in that they only allow for *stationary* threshold variables s_t. Hence, the possibility that $s_t = y_{t-d}$ (which is nonstationary under the null hypothesis of interest) is excluded explicitly. The reasons for them doing so will be pointed out in the next section as well.

3 Testing for a unit root against the CTAR alternative

In this section, we address the question how the unit root hypothesis can be tested against the alternative of a stationary CTAR model, both in case the threshold γ is known and unknown. We discuss the test statistics proposed by Enders and Granger (1998) and develop an alternative test for the case of unknown threshold which has more desirable properties. It turns out that our test statistic can easily be obtained from the estimation results of the alternative model. Hence we briefly discuss the issue of estimating the parameters in the CTAR model as well.

3.1 Testing for a unit root in case of known threshold

In this section we discuss how the null hypothesis of linearity and nonstationarity can be tested against the alternative of a stationary CTAR model for fixed value of the threshold γ. It is useful to rewrite (2) in 'first-difference' form as

$$\Delta y_t = \begin{cases} (\phi_1 - 1)(y_{t-1} - \gamma) + \varepsilon_t & \text{if } y_{t-1} \leq \gamma, \\ (\phi_2 - 1)(y_{t-1} - \gamma) + \varepsilon_t & \text{if } y_{t-1} > \gamma, \end{cases} \tag{4}$$

where we have defined $\phi_i = \phi_{i1}$, $i = 1, 2$, to simplify notation. Formulated more compactly,

$$\Delta y_t = \rho_1(y_{t-1} - \gamma)1(y_{t-1} \leq \gamma) + \rho_2(y_{t-1} - \gamma)1(y_{t-1} > \gamma) + \varepsilon_t, \tag{5}$$

where Δ denotes the first difference operator, i.e., $\Delta x_t = x_t - x_{t-1}$, $1(A)$ denotes the indicator function for the event A, i.e., $1(A) = 1$ if A is true and $1(A) = 0$ otherwise, and $\rho_i = \phi_i - 1, i = 1, 2$.

Using the compact formulation of the model as given in (5), the null hypothesis can be rewritten as $H_0 : \rho_1 = \rho_2 = 0$, while the alternative hypothesis is given by $H_a : \rho_1 < 0, \rho_2 < 0$. Note that for known γ, the model is linear in the remaining parameters ρ_1 and ρ_2. Hence, least squares estimates of $\rho = (\rho_1, \rho_2)$ are easily obtained as

$$\hat{\rho}(\gamma) = \left(\sum_{t=1}^{n} y_{t-1}(\gamma)y_{t-1}(\gamma)'\right)^{-1} \left(\sum_{t=1}^{n} y_{t-1}(\gamma)\Delta y_t\right). \tag{6}$$

where $y_{t-1}(\gamma) = ((y_{t-1} - \gamma)1(y_{t-1} \leq \gamma), (y_{t-1} - \gamma)1(y_{t-1} > \gamma))'$. The corresponding residuals are denoted $\hat{\varepsilon}_t(\gamma) = \Delta y_t - y_{t-1}(\gamma)'\hat{\rho}(\gamma)$ with variance $\hat{\sigma}_n^2(\gamma) = \frac{1}{n}\sum_{t=1}^{n} \hat{\varepsilon}_t(\gamma)^2$.

The unit root hypothesis now might be tested using a conventional F-test, which is computed as

$$F_n(\gamma) = \frac{1}{2} \frac{\sum_{t=1}^{n} \Delta y_t y_{t-1}(\gamma)'(\sum_{t=1}^{n} y_{t-1}(\gamma)y_{t-1}(\gamma)')^{-1} \sum_{t=1}^{n} y_{t-1}(\gamma)\Delta y_t}{\hat{\sigma}_n^2(\gamma)}. \tag{7}$$

It turns out that for *fixed* γ it is impossible to obtain a proper asymptotic distribution for $F_n(\gamma)$. The reason for this is the erratic limiting behavior of the terms $1(y_{t-1} \leq \gamma)$ and $1(y_{t-1} > \gamma)$ in $y_{t-1}(\gamma)$ when y_t is nonstationary. Note that in case of *stationary* y_t, $\frac{1}{n}\sum_{t=1}^{n} 1(y_{t-1} \leq \gamma)$ converges to $F(\gamma)$ in case $n \to \infty$, where $F(\cdot)$ is the CDF of the stationary distribution of y_t. By contrast, in case of *nonstationary* y_t, it follows that,

$$\begin{aligned}
\frac{1}{n}\sum_{t=1}^{n} 1(y_{t-1} \leq \gamma) &= \int_0^1 1(y_{[r(n-1)]} \leq \gamma)dr \\
&= \int_0^1 1(\frac{y_{[r(n-1)]}}{\sigma\sqrt{n}} \leq \frac{\gamma}{\sigma\sqrt{n}})dr \\
&\Rightarrow \int_0^1 1(W(r) \leq 0),
\end{aligned} \tag{8}$$

where $[\cdot]$ denotes integer part and $W(r)$ denotes a standard Wiener process on $[0, 1]$. Note that (8) is independent of the value of γ and, consequently, the asymptotic distribution of the F-statistic would be invariant to the location of the threshold. This in fact is the reason why both Caner and Hansen (1997) and González and Gonzalo (1997) only allow for stationary threshold variables s_t in (2).

The problem noted above can be circumvented however if we allow for a 'drifting' threshold by redefining the threshold in terms of the maximum and minimum of the sample. In particular, define the threshold $\gamma(\tau)$ as

$$\gamma(\tau) = (1-\tau) \min_{1 \leq t \leq n} y_{t-1} + \tau \max_{1 \leq t \leq n} y_{t-1} = (1-\tau) y_{(0)} + \tau y_{(n-1)}, \tag{9}$$

for certain $\tau \in (0,1)$, where $y_{(0)}, \ldots, y_{(n-1)}$ denote the order statistics of the threshold variable y_{t-1}, $y_{(0)} \leq \ldots \leq y_{(n-1)}$. Using standard arguments we then arrive at the asymptotic distribution of the test statistic $F_n(\gamma(\tau))$ as stated in the following theorem, the proof of which is given in the appendix.

Theorem 1 (Asymptotic distribution of $F_n(\gamma(\tau))$ in case of known threshold) *Under the null hypothesis of no threshold and a unit root, the $F_n(\gamma(\tau))$ statistic in (7), with $\gamma(\tau)$ given by (9) for fixed $\tau \in (0,1)$, has the following asymptotic distribution*

$$F_n(\gamma(\tau)) \Rightarrow \frac{1}{2} \int_0^1 dW(r) K_\tau(r)' \left(\int_0^1 K_\tau(r) K_\tau(r)' dr \right)^{-1} \int_0^1 K_\tau(r) dW(r),$$

where

$$K_\tau(r) = \begin{pmatrix} V_\tau(r) 1(V_\tau(r) < 0) \\ V_\tau(r) 1(V_\tau(r) > 0) \end{pmatrix},$$

with

$$V_\tau(r) = W(r) - (1-\tau) \min_{0 \leq s \leq 1} W(s) - \tau \max_{0 \leq s \leq 1} W(s),$$

where $W(r)$ denotes a standard Wiener process on the interval $[0,1]$.

To conclude this section, it is worth noting that the notion of a 'drifting threshold' as defined in (9) bears some similarity to the analysis in Leybourne, Mills and Newbold (1998), who discuss the behavior of the Dickey-Fuller tests in the presence of a break under the null hypothesis. In that case the null distribution is invariant to the size of the break for *fixed* break size. If the size of the break is allowed to grow with the sample size, a limiting distribution for the test statistic is obtained which does depend on the break size[2].

[2] Another analogy with our approach can be found in the literature on the effects of outliers on estimators and test statistics, where it is assumed that outliers occur with a certain probability and, hence, the number of outliers increases as the sample size increases. See Franses and Haldrup (1994) for an application of this idea in the context of outliers and unit root tests.

3.2 Estimation of the CTAR model

In practice, the value of the threshold γ will of course be unknown in most instances. In this section we briefly recapitulate estimation of the parameters in the CTAR model, including the threshold parameter. The main reason for this digression is that the statistic to test the unit root hypothesis in case of unknown threshold (which is discussed in the next section) can be obtained directly from the estimation results. More general discussions on estimating (SE)TAR models can be found in Tong (1990) and Hansen (1996b,1997a).

Starting from the compact formulation of the model as given in (5), it is easily understood that the parameters of interest, i.e., ρ_1, ρ_2, γ and σ^2 can conveniently be estimated by sequential conditional least squares. As already noted above, for known and fixed γ, the model is linear in the remaining parameters and least squares estimates of $\rho = (\rho_1, \rho_2)$ are easily obtained from (6), with corresponding residual variance $\hat{\sigma}_n^2(\gamma)$. The least squares estimate of γ then is obtained by minimizing this residual variance, i.e.,

$$\hat{\gamma} = \underset{\gamma \in \Gamma}{\operatorname{argmin}}\, \hat{\sigma}_n^2(\gamma), \tag{10}$$

where Γ denotes the set of all allowable threshold values. The final estimates of the autoregressive parameters then are given by $\hat{\rho} = \hat{\rho}(\hat{\gamma})$, while the residual variance is estimated as $\hat{\sigma}_n^2 = \hat{\sigma}_n^2(\hat{\gamma})$.

For estimation purposes, the exact definition of the set Γ is not essential. The parameters of the model can be estimated consistently as long as the true value of γ is in the interior of Γ. In addition, Γ should be such that each regime contains enough data for the estimator defined above to produce reliable estimates of the autoregressive parameters. A popular choice for Γ is to require that each regime contains at least a (pre-specified) fraction π_0 of the observations[3], i.e.,

$$\Gamma = \{\gamma | y_{([\pi_0(n-1)])} \leq \gamma \leq y_{([(1-\pi_0)(n-1)])}\}, \tag{11}$$

where $y_{(0)}, \ldots, y_{(n-1)}$ again denote the order statistics of the threshold variable y_{t-1}, $y_{(0)} \leq \cdots \leq y_{(n-1)}$.

The minimization problem (10) can be solved by means of direct search, although there is a notable difference with estimation of the parameters in the corresponding 'unconstrained' SETAR model, i.e., (1) with $p = 1$ and $s_t = y_{t-1}$. In the latter case, it suffices

[3] Since Andrews (1993), the default value for π_0 appears to be 0.15.

to compute the residual variance $\hat{\sigma}_n^2(\gamma)$ only for threshold values equal to the order statistics of y_{t-1}, i.e., with $\gamma = y_{(i)}$ for each i such that $y_{(i)} \in \Gamma$. This is easily understood since no observations are transferred from one regime to the other when γ is varied between two consecutive order statistics. Consequently, the value of $\hat{\sigma}_n^2(\gamma)$ does not change in that case. By contrast, in the CTAR model, changing the threshold γ changes the regressor $y_{t-1}(\gamma)$ and, hence, the value of $\hat{\sigma}_n^2(\gamma)$, even if no observations move from one regime to the other. Since a detailed grid search over Γ would be rather time-consuming, a pragmatic approach is to proceed as in the unconstrained SETAR case and restrict the grid search to values of γ equal to the order statistics of y_{t-1}. In practice the resulting parameter estimates will be reasonably close to the estimates which would be obtained from a full search. If desired, a second grid search around the initial estimate of γ can be done to refine the estimates.

Under the stationarity assumption $|\phi_1| < 1$ and $|\phi_2| < 1$, Chan and Tsay (1998) demonstrate that the conditional least squares estimators of all parameters, including the threshold γ, are $\sqrt{n}$ consistent and asymptotically normal (where in addition $\hat{\gamma}$ is correlated with the other parameters). This also is rather different from the the general (discontinuous) TAR model (1), for which Chan (1993) demonstrates that the estimator of the threshold is consistent at rate n and asymptotically independent of the other parameter estimators. Chan (1993) also shows that the asymptotic distribution of $\hat{\gamma}$ depends upon a host of nuisance parameters, including the true regression parameters ρ. Using an alternative approach, Hansen (1997a) derives a limiting distribution for $\hat{\gamma}$ which is free of nuisance parameters (up to scale).

3.3 Testing for a unit root in case of unknown threshold

In case the threshold is unknown, the testing problem suffers from what has become known as the 'Davies-problem', in the sense that γ is a nuisance parameter which is not identified under the null hypothesis. For general expositions of testing under such circumstances we refer to Davies (1977,1987), Andrews (1994) and Hansen (1996a). Enders and Granger (1998) suggest to estimate γ by the sample mean of y_{t-1}, i.e., $\overline{\gamma} = \frac{1}{n}\sum_{t=1}^{n} y_{t-1}$, and use $F_n(\overline{\gamma})$, computed from (7), to test the unit root hypothesis. However, if the strength of adjustment towards the attractor γ is really different in the two regimes, the sample mean is a biased estimator of γ, as shown already by Tong (1983). Given that the objective of the analysis in Enders and Granger (1998) is to obtain statistics which have better (power) properties than standard unit root tests against alternatives with asymmetric adjustment, it appears rather strange or 'suboptimal' to use such a biased estimate of the crucial parameter in the model under the alternative. In fact, the simulation evidence presented in Enders and Granger (1998) suggests that their test has less power than the standard Dickey-Fuller statistic against the alternative model (3).

Instead of the sample mean of y_{t-1}, it might make more sense to use the consistent estimate $\hat{\gamma}$ as obtained from (10) and test the unit root hypothesis using the statistic $F_n(\hat{\gamma})$. To enable obtaining an analytic asymptotic distribution of $F_n(\hat{\gamma})$, we reformulate the estimation of the CTAR model in terms of the drifting threshold $\gamma(\tau)$ as follows

$$\hat{\tau} = \underset{\tau \in \mathcal{T}}{\operatorname{argmin}}\, \hat{\sigma}_n^2(\gamma(\tau)), \tag{12}$$

with

$$\mathcal{T} = \{\tau | \tau_0 \leq \tau \leq 1 - \tau_1\}, \tag{13}$$

for certain $\tau_0, \tau_1 \in (0,1)$ and $\gamma(\tau)$ as defined in (9). The threshold estimate then is given by $\hat{\gamma}(\hat{\tau}) = (1-\hat{\tau})y_{(0)} + \hat{\tau} y_{(n-1)}$. It should be intuitively clear that the parameters τ_0 and τ_1 in (13) can be chosen such that the set of allowable threshold values implied by (9) and (13) is exactly the same as the set Γ defined in (11). In particular, if τ_0 and τ_1 are set such that

$$\tau_0 = \frac{y_{([\pi_0(n-1)])} - y_{(0)}}{y_{(n-1)} - y_{(0)}}, \tag{14}$$

$$\tau_1 = \frac{y_{(n-1)} - y_{([(1-\pi_0)(n-1)])}}{y_{(n-1)} - y_{(0)}}, \tag{15}$$

both regimes always contain at least a fraction π_0 of the observations. Hence, at the cost of introducing an additional nuisance parameter it is possible to redefine the conventional set Γ in terms of the minimum and maximum of the sample[4].

[4] The number of nuisance parameters can be restricted to 1 if (13) is replaced by

$$\mathcal{T} = \{\tau | \tau_0 \leq \tau \leq 1 - \tau_0\}, \tag{16}$$

for certain $\tau_0 \in (0,1)$. Obviously, the set of allowable thresholds implied by (9) and (16) will differ from the set Γ as given in (11) since in general it will be impossible to choose the parameter τ_0 such that both regimes contain at least a fraction π_0 of the observations. In fact, if the distribution of y_t is highly asymmetric, τ_0 might have to be chosen quite large in order to guarantee that the regime where the data are sparse contains a minimum fraction of the observations, such that the set of allowable thresholds (unnecessarily) becomes quite small.

Notice that the F-statistic (7) can be rewritten in terms of the residual variance $\hat{\sigma}_n^2(\gamma(\tau))$ as

$$F_n(\gamma(\tau)) = \frac{n}{2}\left(\frac{\tilde{\sigma}_n^2 - \hat{\sigma}_n^2(\gamma(\tau))}{\hat{\sigma}_n^2(\gamma(\tau))}\right), \tag{17}$$

where $\tilde{\sigma}_n^2 = \sum_{t=1}^n \Delta y_t$ estimates the residual variance under the null hypothesis. Since this is a monotonic transformation of $\hat{\sigma}_n^2(\gamma(\tau))$ it follows that $F_n(\hat{\gamma}(\hat{\tau}))$ is equivalent to the supremum[5] of the pointwise test statistic $F_n(\gamma(\tau))$ over the set $\mathcal{T}$ of allowable values for $\gamma(\tau)$,

$$F_n \equiv F_n(\hat{\gamma}(\hat{\tau})) = \sup_{\tau \in \mathcal{T}} F_n(\gamma(\tau)). \tag{18}$$

Note that a reverse interpretation of our proposed procedure thus would be that the estimator of the threshold is defined to be the value of $\gamma(\tau)$ for which the evidence against the null-hypothesis (in terms of the F-test (17)) is maximized.

At this point it should be remarked that Enders and Granger (1998) also discuss the use of the t-statistics of ρ_1 and ρ_2 for discriminating between the random walk model and the CTAR model. In particular, they define the $t_{\max}(\overline{\gamma})$ and $t_{\min}(\overline{\gamma})$ statistics as the most and least significant of $t_{\rho_1(\overline{\gamma})=0}$ and $t_{\rho_2(\overline{\gamma})=0}$, respectively, where the notation is such that it is emphasized that the threshold is estimated by the sample mean of y_{t-1}. The $t_{\max}(\overline{\gamma})$ test is dubbed 'max' because it is equal to the t-statistic that yields maximal evidence against the null hypothesis of nonstationarity. The reverse holds for the $t_{\min}(\overline{\gamma})$ test[6]. A possible advantage of the t-tests over the F-test is that the first two statistics might be used to exploit the one-sided nature of the alternative hypothesis ($\rho_1 < 0$ and $\rho_2 < 0$), whereas the F-test can not. On the other hand, the t-tests do not properly use all information. The $t_{\min}$ test neglects information contained in the smaller t-statistic, whereas the $t_{\max}$ test does not use the exact magnitude of the larger t-statistic.

Translating these t-statistics to our framework would involve defining the $t_{n,\max}$ and $t_{n,\min}$ statistics as the minima of $t_{\max}(\gamma(\tau))$ and $t_{\min}(\gamma(\tau))$, respectively, over the region

[5] Alternative test statistics, such as the Ave or Exp might also be considered, cf. Andrews and Ploberger (1994).

[6] However, an alternative interpretation would be that finding a significant $t_{\min}$ statistic provides maximal evidence against the null, since this automatically implies that both ρ_1 and ρ_2 are significantly different from zero. In addition, given the one-sided nature of the alternative hypothesis, it might be more sensible to consider just the largest and smallest t-statistics as test statistics. Hence, one might define the $t_{\max}(\overline{\gamma})$ statistic simply as the smaller of $t_{\rho_1(\overline{\gamma})=0}$ and $t_{\rho_2(\overline{\gamma})=0}$, and the $t_{\min}(\overline{\gamma})$ statistic as the larger of the two.

$\mathcal{T}$. The main objection against the use of the t-statistics is that the threshold values at which the $t_{\max}(\gamma(\tau))$ and $t_{\min}(\gamma(\tau))$ are minimized, need not be the same as the threshold estimate $\hat{\gamma}(\hat{\tau})$ obtained from (12). Moreover, it does not appear feasible to prove that the value at which they are minimized provides a consistent estimate for the threshold γ (which is necessary to derive the limiting distribution of the t-tests). An alternative definition of the $t_{n,\max}$ and $t_{n,\min}$ statistics as the most and least significant (or smallest and largest) t-statistic at the threshold estimate $\hat{\gamma}(\hat{\tau})$ suffers from the drawback that the t-tests no longer are the infimum of the pointwise t-stats over the region $\mathcal{T}$, which again prevents the derivation of the limiting distribution. Hence we do not consider the t-tests any further.

The following theorem establishes the limiting behaviour of the F_n test statistic as defined in (18).

Theorem 2 (Asymptotic distribution of F_n in case of unknown threshold) *Under the null hypothesis of no threshold and a unit root, the F_n statistic in* (18), *with $\mathcal{T}$ defined in* (13), *has the following asymptotic distribution*

$$F_n \Rightarrow \sup_{\tau \in \mathcal{T}} \frac{1}{2} \int_0^1 dW(r) K_\tau(r)' \left(\int_0^1 K_\tau(r) K_\tau(r)' dr \right)^{-1} \int_0^1 K_\tau(r) dW(r). \tag{19}$$

The limiting distribution of our test statistic has both 'Dickey-Fuller' and 'sup' characteristics and, hence, critical values have to be calculated by simulation. We simulate the finite sample distributions for $n = 50, 75, 100, 150, 200, 250, 500$ and 1000 for a number of pairs (τ_0, τ_1) by varying τ_0, τ_1 among $\tau_0, \tau_1 \in \{0.10, 0.15, \ldots, 0.55, 0.6 | \tau_0 + \tau_1 < 1\}$. We use 10000 replications and approximate the continuous time Wiener processes in (19) with discrete time random walks consisting of n observations. Estimates of quantiles of the asymptotic distribution in (19) are obtained from response surfaces. That is, we estimate equations like

$$q^p_{\tau_0,\tau_1}(n) = q^p_{\tau_0,\tau_1}(\infty) + b_1 n^{-1} + b_2 n^{-2} + \varepsilon^p_{\tau_0,\tau_1}. \tag{20}$$

In this equation, $q^p_{\tau_0,\tau_1}(n)$ denotes the pth quantile of the distribution of F_n for a particular combination (τ_0, τ_1) and a sample size of n observations. The intercept $q^p_{\tau_0,\tau_1}(\infty)$ provides an estimate of the pth quantile of the corresponding asymptotic distribution, cf. MacKinnon (1994)[7].

[7]In order to get some insight into the reliability of these estimates, we assess the accuracy of the $q^p_{\tau_0,\tau_1}(n)$

It will be clear that it is nor feasible nor desirable to present critical values for all possible combinations of (τ_0, τ_1). In order to get an impression of the magnitude of the respective critical values, Table 1 presents a sample of the results for $n = 100, 250, 500$ and 1000 as well as estimates of quantiles of the asymptotic distribution $q^p_{\tau_0,\tau_1}(\infty)$ for $p = 0.9, 0.95$ and 0.99.

- insert Table 1 about here -

In addition, Figure 1 shows the complete set of 90, 95 and 99% percentiles of the distributions of the F_n-statistic for $n = 100, 250$ and 500.

- insert Figure 1 about here -

From the entries in Table 1 and the surfaces in Figure 1 it is seen that the critical values of the F-statistic are decreasing in both τ_0 and τ_1. This makes sense intuitively, since lowering either τ_0 or τ_1 enlarges the set of allowable threshold values over which the supremum of the pointwise statistic is to be taken. Furthermore, the difference between the quantiles of the asymptotic distribution and of the finite sample distribution for $n \geq 500$ varies from very small to quite substantial, depending on the values of τ_0 and τ_1. It appears that even for these sample sizes the asymptotic distribution does not always offer a good approximation to the finite sample distribution. This being the case, the main practical use of the asymptotic distribution is to get a first indication of the potential significance of the F_n statistic. Of course the finite sample critical values reported in Table 1 might be used as well, although one should realize that these depend on the particular model which is used[8], and are available only fir a limited number of sample sizes and combinations

using the formula for the variance of a quantile estimate (Kendall and Stuart, 1977 p. 252)

$$\mathrm{VAR}(q^p_{\tau_0,\tau_1}(n)) = p(1-p)/(mf^2).$$

In this formula, m denotes the number of replications and f denotes the height of the density function of the population distribution at the estimated quantile. In our case, the population distribution is unknown, hence we approximate the distribution with a Gamma distribution. A first choice would have been to approximate the population distribution with an F-distribution, as the F-test has an F-distribution for fixed $\gamma(\tau)$ in case the process is assumed stationary under the null hypothesis. However, we tried to match an F-distribution to the simulated distribution by the method of moments, the number of degrees of freedom that resulted from matching the sample mean and the population expectation, indicated that the population variance did not exist. We felt this to be unrealistic, and hence discarded the F-distribution. The average standard error of the various quantile estimates equals 0.033. Hence, we conclude that the estimates of the quantiles of the asymptotic distribution are fairly accurate.

[8]For example, if additional lagged first differences are incorporated in the CTAR model, the resulting F_n-statistic is only asymptotically similar to the F_n-statistic based upon the basic model (3). See Cheung and Lai (1995a) for a discussion of this point for the standard ADF test.

(τ_0, τ_1). Therefore it seems advisable to use bootstrap techniques to assess the significance of the F_n statistic more precisely.

Also note that critical values as reported here are limited in the sense that they only allow to test the null hypothesis at certain specific significance levels, whereas in practice p-values may be more informative or desirable, cf. Hansen (1997b). It turns out that computationally convenient approximations to the limiting distribution of the F_n-statistic can be obtained by fitting a Gamma distribution to the estimated asymptotic quantiles $q^{p}_{\tau_0,\tau_1}(\infty)$. Similar to Hansen (1997b), we use a weighted loss function over the p-value space. In our case, we use p equal to $(0.50, 0.60, 0.70, 0.80, 0.82, \ldots, 0.88, 0.90, 0.91, \ldots, 0.99)$. This implies that we implicitly give no weight to the left tail of the distribution and increasing weights to the right tail of the distribution. We use a quadratic loss function - higher order loss functions appeared less satisfactory. A comprehensive list of parameter estimates is provided in Table 2. The approximation appears to be sufficiently accurate for empirical applications. For instance, the maximum absolute error over all combinations (τ_0, τ_1) is below 0.0055 for p equal to $(0.90, 0.91, \ldots, 0.99)$.

- insert Table 2 about here -

Again, parameter estimates are only available for a limited number of combinations of (τ_0, τ_1). p-values for arbitrary values of (τ_0, τ_1) can be obtained by interpolation, as discussed in the next Section.

4 Power in finite samples

In this section we use Monte Carlo experiments to investigate the finite sample performance of the F_n test statistic developed in Section 3. We also compare its properties with the $F_n(\overline{\gamma})$ test of Enders and Granger (1998) and with some commonly applied unit root tests, i.e., the standard Augmented Dickey-Fuller [ADF] test, the efficient ADF-GLS test of Elliot, Rothenberg and Stock (1996), and the test statistic of Kwiatkowski, Phillips, Schmidt and Shin (1992) [KPSS].

4.1 Monte Carlo design

In the simulations, we generate 5000 series of length $n + 100$ from the CTAR model as given in (3). The necessary starting value y_0 is fixed at 0, while the first 100 observations of each series are discarded, such that the effective sample size is equal to n, which we set to $n = 100$ and 250. The autoregressive parameters ϕ_1 and ϕ_2 are varied among $\phi_i = \{0.95, 0.9, 0.8, 0.7, 0.6, 0.5\}$, $i = 1, 2$. Since the CTAR model with $\phi_1 = \phi_l$ and $\phi_2 = \phi_u$ is in a sense equivalent to the model with $\phi_1 = \phi_u$ and $\phi_2 = \phi_l$ for all ϕ_l and ϕ_u, we only report results for parameterizations for which $\phi_2 \leq \phi_1$. Because the threshold γ is only a location parameter in the CTAR model we set it equal to zero without loss of generality. Finally, the errors ε_t are taken to be i.i.d standard normal.

Even though the Monte Carlo experiments are not all that extensive, using bootstrap techniques for each and every replication would amount to an excessive computational burden. Therefore we use a more pragmatic way to determine the appropriate critical values as follows.

We simulate the distribution of the F_n statistic under the null hypothesis for the relevant sample sizes $n = 100$ and 250, for all combinations (τ_0, τ_1) such that $\tau_0, \tau_1 \in \{0.10, 0.125, 0.15, \ldots, 0.575, 0.6 | \tau_0 + \tau_1 < 1\}$ using 10000 replications, and retain the critical values at 1, 5 and 10% significance levels. Recall that the usual practice when estimating TAR models is to define the set of allowable threshold values such that both regimes of the model contain at least a fraction π_0 of the observations. We follow this convention by setting τ_0 and τ_1 according to (14) and (15) given above for each replication in the power experiments, where we set $\pi_0 = 0.15$. The resulting 'estimates' of τ_0 and τ_1 are denoted $\hat{\tau}_0$ and $\hat{\tau}_1$, respectively. The appropriate critical value at a given significance level for any $(\hat{\tau}_0, \hat{\tau}_1)$ then is determined by interpolating the corresponding critical values from the four grid points surrounding[9] $(\hat{\tau}_0, \hat{\tau}_1)$, using the Euclidean metric to determine the weights which are to be assigned to each of the four grid points[10].

Critical values for the Enders-Granger tests are obtained by simulation as well. The response surfaces from Cheung and Lai (1995a,b) are used to obtain the appropriate critical values for the ADF and ADF-GLS statistics, while critical values for the KPSS statistic are taken from Kwiatkowski *et al.* (1992). The ADF test, the ADF-GLS test, and the

[9]It follows easily from (14) and (15) that in practice both τ_0 and τ_1 will be larger than π_0. Hence the grid of points (τ_0, τ_1) for which critical values have been simulated is such that $(\hat{\tau}_0, \hat{\tau}_1)$ is in the 'interior' of the grid, given our choice of $\pi_0 = 0.15$.

[10]In particular, let $\tau_{0,i}$ and $\tau_{1,j}$ denote the grid values of τ_0 and τ_1 such that $\tau_{0,i} \leq \hat{\tau}_0 \leq \tau_{0,i+1}$ and $\tau_{1,j} \leq \hat{\tau}_1 \leq \tau_{1,j+1}$. The weight assigned to the critical value at, for example, $(\tau_{0,i}, \tau_{1,j})$ then is equal to $d(i,j)/(d(i,j) + d(i,j+1) + d(i+1,j) + d(i+1,j+1))$, where $d(i,j)$ is the Euclidean distance from $(\hat{\tau}_0, \hat{\tau}_1)$ to $(\tau_{0,i}, \tau_{1,j})$.

KPSS test statistic are computed including a constant in the test equation.

Finally it should be remarked that the ADF and ADF-GLS statistics are misspecified under the CTAR DGP since they assume symmetric adjustment under the alternative. By contrast, the KPSS statistic has stationarity as null hypothesis. As shown by Corradi, Swanson and White (1998), the CTAR model is allowed under this null, in the sense that the asymptotic distribution as derived by Kwiatkowski *et al.* (1992) remains valid. Hence, one might perhaps expect the KPSS test to perform better relative to the ADF and ADF-GLS statistics.

4.2 Monte Carlo results

The percentage of times the unit root null hypothesis is correctly rejected at the 1, 5 and 10% significance levels is reported in Tables 3 and 4 for sample sizes $n = 100$ and 250, respectively. To facilitate comparison of different test statistics, the column headed KPSS reports the percentage of non-rejections of its null hypothesis of stationarity.

- insert Tables 3 and 4 about here -

Several conclusions emerge from these Tables. First of all, the F_n test statistic clearly outperforms the $F_n(\overline{\gamma})$ statistic, often by quite a wide margin. This demonstrates that the use of a consistent estimate of the threshold parameter instead of an *ad-hoc* estimate pays off in terms of power of the resulting test statistic. Second, the F_n test has better power than the standard Dickey-Fuller statistic, while the $F_n(\overline{\gamma})$ has lower power than the ADF statistic. The latter confirms the findings of Enders and Granger (1998), while the former shows that it is worthwhile to properly specify the alternative against which the unit root hypothesis is tested. A somewhat puzzling result is that the F_n statistic even has higher power than the ADF test in case of symmetric mean-reversion. One would expect that the unnecessary estimation of an additional parameter would result in a loss in power. Thirdly, for the smaller sample size $n = 100$, the efficient ADF-GLS test has better power than the F_n test for those combinations of ϕ_1 and ϕ_2 which imply slow mean-reversion in both regimes, e.g., $(\phi_1, \phi_2) = (0.95, 0.8)$. If mean-reversion in one or both of the regimes is fast however, the F_n statistic has higher power. Notice that this is also the case for highly asymmetric DGP's, such as $(\phi_1, \phi_2) = (0.95, 0.6)$. For $n = 250$ the ADF-GLS statistic then has even lower power than the standard Dickey-Fuller statistic in most cases[11]. Similar observations can be made for the KPSS test. If both regimes contain a near unit-root, the non-rejection of the null hypothesis of stationarity by the KPSS is (much) higher than the power of our F_n statistic, while the opposite holds if mean reversion in one or both of the regimes is fast.

5 Empirical illustration

The aim of this section is to demonstrate the practical usefulness of the F_n statistic. To this end, we use several forward premium series, previously analyzed by Hai, Mark and Wu (1997). More specifically, we analyze the one month ahead forward premium for the US dollar value of the UK pound, the French franc, and the Japanese yen. The data are sampled at the monthly frequency over the period January 1976 to August 1992 (200 observations). We refer to Hai *et al.* (1997) for a more extensive description of the data.

The forward premium y_t is defined as $y_t = f_t - s_t$, where f_t and s_t denote the logarithms of the forward and spot exchange rates, respectively. We address the question whether the forward premium is nonstationary or can best be described as a stationary process with asymmetric adjustment. The issue of stationarity of the forward premium is linked to the question whether foreign exchange markets are efficient or not, cf. Engel (1996). If the risk premium is stationary, then stationarity of the forward premium is a necessary condition for efficiency of the foreign exchange market.

Hai *et al.* (1997) investigate the stationarity of the forward premium using a number of well known univariate unit root test. They conclude that for the UK pound and the French franc, nonstationarity can be rejected at the 5% significance level, while for the Japanese yen this is possible only at the 10% significance level. Using the ADF test and the lag length selection method advocated by Ng and Perron (1995) we are able to replicate their results, at least qualitatively - see Table 5. The initial lag length of the AR model is set equal to six. Lagged first difference variables are deleted when they are insignificant at the 5% level[12].

---

[11] This is probably due to the setup of the power simulations in which the first 100 observations are discarded. As a result, the initial observation used in the calculation of the test statistics is not fixed but distributed according to the stationary distribution of the process. Under these circumstances, the power of the ADF-GLS test quickly deteriorates, cf. Elliot *et al.* (1996).

[12] Unfortunately, Hai *et al.* (1997) do not supply details on the initial number of lags and the significance

Next, we apply the $F_n(\overline{\gamma})$ and F_n statistics to the three forward premium series. Similar to the ADF test, we add lagged first differences to equation (5) to handle remaining autocorrelation in the error terms. The number of lags is chosen equal to the number of lags used in the linear model. The values of the test statistics as well as p-values are presented in Table 5. The approximate asymptotic p-value for the F_n statistic, $p_{F_n,\infty}$, is obtained using the Gamma approximation outlined in Section 3.3. This p-value gives an initial hint regarding the significance level of the F_n statistic. The second p-value, $p_{F_n,b}$, is derived using a nonparametric bootstrap method. In constructing the bootstrap samples, we use the 'estimated' lag length in order to approximate the finite sample distribution of the F_n and $F_n(\overline{\gamma})$ statistics as close as possible.

- insert Table 5 about here -

Finally, the Table also contains estimates of the boundaries of the set $\mathcal{T}$ of allowed threshold values, τ_0, and τ_1, the estimate of the threshold value γ, and estimates of the parameters for the lagged level term y_{t-1} in the two regimes, ϕ_1, ϕ_2. The last row in Table 5 shows a Wald test of the hypothesis of symmetric adjustment versus asymmetric adjustment given that the forward premium follows a stationary two regime CTAR process[13].

From Table 5 we draw the following conclusion. First, both the $F_n(\overline{\gamma})$ and F_n statistics suggest that nonstationarity of the forward premium can be strongly rejected in all cases, including the case of the Japanese yen. Second, adjustment to the long-run equilibrium is clearly asymmetric. Formal support for this assertion is given by the Wald test. Third, the forward premium of the Japanese yen appears to behave differently from the two European series. In the Japanese case, the lower regime is most persistent, whereas in the cases of the UK pound and the French franc, it is the upper regime that features most persistence.

Generalizations to the Vector Error-Correction framework will be needed in order to address the question whether the two regime CTAR process is able to shed some light on the 'forward premium puzzle' (see Lewis (1995)).

level in their Table III. This makes replicating their results exactly less straightforward.

[13]The Wald test takes into account the fact in the CTAR model the regression parameters and the threshold variable are jointly asymptotically normally distributed, cf. Chan and Tsay (1998).

6 Concluding remarks

The literature on test statistics which can be used to test the null hypothesis of a unit root against an alternative of stationarity with asymmetric adjustment is growing. In the present paper, we have reconsidered the approach put forward by Enders and Granger (1998). We have shown that their test is suboptimal in the sense that a biased estimate of the threshold is used. An alternative test with a more solid statistical foundation has been developed. This test statistic can easily be obtained by sequential least squares estimation. However, in order to derive a meaningful description of the limiting behaviour of this test statistic, the threshold value is allowed to depend on a specific functional of the sample. Several ways of obtaining appropriate critical values have been discussed. The Monte Carlo experiments show that the new test clearly outperforms standard unit root tests and the test of Enders and Granger (1998). The application to the forward premium series demonstrates the usefulness of our test statistic, and suggests that it might be worthwhile to consider the possibility of asymmetric adjustment for these series.

A topic for further research is to extend the methods presented in this paper towards testing for cointegration in the presence of asymmetric error-correction.

Appendix: Mathematical proofs

Proof of Theorem 1:
Throughout the proof we assume that $\hat{\sigma}_n^2(\gamma(\tau)) \xrightarrow{P} \sigma^2$. Define $y_t(\gamma(\tau))$ as

$$y_t(\gamma(\tau)) = \begin{pmatrix} (y_{t-1} - \gamma(\tau))1(y_{t-1} \leq \gamma(\tau)) \\ (y_{t-1} - \gamma(\tau))1(y_{t-1} > \gamma(\tau)) \end{pmatrix}, \tag{21}$$

with

$$\gamma(\tau) = (1-\tau) \min_{1 \leq t \leq n} y_{t-1} + \tau \max_{1 \leq t \leq n} y_{t-1},$$

for certain $\tau \in (0,1)$. Furthermore, define the process $V_\tau(r)$ as

$$V_\tau(r) = W(r) - (1-\tau) \min_{0 \leq s \leq 1} W(s) - \tau \max_{0 \leq s \leq 1} W(s),$$

for all $r \in (0,1)$ with $W(s)$ a standard Wiener process. Then, using standard arguments it follows that, for any $r \in (0,1)$,

$$\frac{1}{\sqrt{\sigma^2 n}} y_{[rn]}(\gamma(\tau)) \Rightarrow \begin{pmatrix} V_\tau(r)1(V_\tau(r) < 0) \\ V_\tau(r)1(V_\tau(r) > 0) \end{pmatrix} \equiv K_\tau(r),$$

where $[\cdot]$ denotes integer part. Now, introduce the following functional

$$f(X(\cdot)) = \frac{1}{2} \int_0^1 dW(r)X(r)' \left(\int_0^1 X(r)X(r)'dr \right)^{-1} \int_0^1 X(r)dW(r).$$

Then it immediately follows that for fixed τ

$$F_n(\gamma(\tau)) \Rightarrow f(K_\tau(\cdot)).$$

$\square$

Proof of Theorem 2:
We want to show that

$$F_n = \sup_{\tau \in \mathcal{T}} F_n(\gamma_n(\tau)) \Rightarrow \sup_{\tau \in \mathcal{T}} f(K_\tau(\cdot)).$$

The limiting distribution of $F_n(\gamma(\tau))$ for fixed τ is given by Theorem 1. If we can derive the limiting distribution of the estimator of τ, then the continuous mapping theorem [CMT] can be invoked in order to obtain the limiting distribution of F_n. Define

$$\hat{\tau} = \arg\max_{\tau \in \mathcal{T}} F_n(\gamma(\tau)).$$

We want to use Theorem 2.7 of Kim and Pollard (1990). In order to do so, the following requirements have to be fulfilled: (1) $\hat{\tau}$ has be bounded in probability, (2) $\hat{\tau}$ must be near-maximizing, and (3) $f(K_\tau(\cdot))$ has to be continuous in τ and attain an unique maximum on $\mathcal{T}$. These assumptions are satisfied except for the continuity of $f(K_\tau(\cdot))$. To tackle this problem, we modify the estimation problem in the following way. We replace the indicator functions, which obviously are discontinuous, by continuous functions which can approximate the indicator function arbitrarily closely. In particular, $1(y_{t-1} > \gamma(\tau))$ is replaced by

$$h_u^\delta(y_{t-1} - \gamma(\tau)) = \begin{cases} 1 & \text{if } y_{t-1} - \gamma(\tau) \geq 0, \\ 1 + \frac{1}{\delta}(y_{t-1} - \gamma(\tau)) & \text{if } -\delta < y_{t-1} - \gamma(\tau) < 0, \\ 0 & \text{if } y_{t-1} - \gamma(\tau) \leq -\delta, \end{cases}$$

and likewise $1(y_{t-1} < \gamma(\tau))$ is replaced by h_l^δ which is similarly defined. If we modify all relevant entities defined above in the appropriate way by inserting h_u^δ and h_u^δ at the proper places, then according to Theorem 2.7 in Kim and Pollard (1990)

$$\hat{\tau} \to \arg\max_{\tau \in \mathcal{T}} f(K_\tau^\delta(\cdot)) = \tau^*,$$

where $K_\tau^\delta(\cdot)$ is obviously defined. Hence, using the CMT,

$$\sup_{\tau \in \mathcal{T}} F_n^\delta(\gamma(\tau)) = F_n^\delta(\gamma(\hat{\tau})) \Rightarrow f(K_\tau^\delta(\cdot, \tau^*)) = \sup_{\tau \in \mathcal{T}} f(K_\tau^\delta(\cdot, \tau)).$$

Thus

$$\sup_{\tau \in \mathcal{T}} F_n^\delta(\gamma(\tau)) \Rightarrow \sup_{\tau \in \mathcal{T}} f(K_\tau^\delta(\cdot, \tau)).$$

As h_l^δ and h_u^δ converge pointwise to the indicator functions $1(y_{t-1} < \gamma(\tau))$ and $1(y_{t-1} > \gamma(\tau))$, respectively, as $\delta \downarrow 0$, it follows that

$$\sup_{\tau \in \mathcal{T}} F_n^\delta(\gamma(\tau)) \xrightarrow{d} \sup_{\tau \in \mathcal{T}} F_n(\gamma(\tau)), \quad \text{as } \delta \downarrow 0,$$

for fixed n. Moreover, because

$$\sup_{\tau \in \mathcal{T}} f(K_\tau^\delta(\cdot)) \xrightarrow{d} \sup_{\tau \in \mathcal{T}} f(K_\tau(\cdot)), \quad \text{as } \delta \downarrow 0,$$

it follows that

$$F_n \Rightarrow \sup_{\tau \in \mathcal{T}} f(K_\tau(\cdot)).$$

$\square$

References

Andrews, D.W.K., 1993, Tests for parameter instability and structural change with unknown change point, *Econometrica* **61**, 821-856.

Andrews, D.W.K., 1994, Empirical process methods in econometrics, in R.F. Engle and D.L. McFadden (editors), *Handbook of econometrics, vol. IV*, Amsterdam: Elsevier Science, p. 2248-2296.

Andrews, D.W.K. and W. Ploberger, 1994, Optimal tests when a nuisance parameter is present only under the alternative, *Econometrica* **62**, 1383-1414.

Balke, N.S. and T.B. Fomby, 1997, Threshold cointegration, *International Economic Review* **38**, 627-646.

Caner, M. and B.E. Hansen, 1997, Threshold autoregressions with a unit root, Working Papers in Economics No. 381, Boston College.

Chan, K.S., 1993, Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model, *Annals of Statistics* **21**, 520-533.

Chan, K.S. and H. Tong, 1985, On the use of the deterministic Lyapunov function for the ergodicity of stochastic difference equations, *Advances in Applied Probability* **17**, 666-678.

Chan, K.S., J.D. Petrucelli, H. Tong and S.W. Woolford, 1985, A multiple threshold AR(1) model, *Journal of Applied Probability* **22**, 267-279.

Chan, K.S. and R.S. Tsay, 1998, Limiting properties of the least squares estimator of a continuous threshold autoregressive model, *Biometrika* **85**, 413-426.

Chen, R. and R.S. Tsay, 1991, On the ergodicity of TAR(1) processes, *Annals of Applied Probability* **1**, 613-634.

Cheung, Y.-W. and K.S. Lai, 1995a, Lag order and critical values of a modified Dickey-Fuller test, *Oxford Bulletin of Economics and Statistics* **57**, 411-419.

Cheung, Y.-W. and K.S. Lai, 1995b, Lag order and critical values of the augmented Dickey-Fuller test, *Journal of Business & Economic Statistics* **13**, 277-280.

Corradi, V., N.R. Swanson and H. White, 1998, Testing for stationarity-ergodicity and for co-movements between nonlinear discrete time Markov processes, working paper, University of Pennsylvania.

Davies, R.B., 1977, Hypothesis testing when a nuisance parameter is present only under the alternative, *Biometrika* **64**, 247-254.

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

Elliot, G., T.J. Rothenberg and J.H. Stock, 1996, Efficient tests for an autoregressive unit root, *Econometrica* **64**, 813-836.

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

Engel, C., 1996, A note on cointegration and capital market efficiency, *Journal of International Money and Finance* **15**, 657-660.

Franses, P.H. and N. Haldrup, 1994, The effects of additive outliers on tests for unit roots and cointegration, *Journal of Business & Economic Statistics* **12**, 471-478.

González, M. and J. Gonzalo, 1997, Threshold unit root models, Working paper 97-50, Universidad Carlos III de Madrid.

Hai, W., N.C. Mark and Y. Wu, 1997, Understanding spot and forward exchange rate regressions, *Journal of Applied Econometrics* **12**, 715-734.

Hansen, B.E., 1996a, Inference when a nuisance parameter is not identified under the null hypothesis, *Econometrica* **64**, 413-430.

Hansen, B.E., 1996b, Sample splitting and threshold estimation, Working paper, Boston College.

Hansen, B.E., 1997a, Inference in TAR models, *Studies in Nonlinear Dynamics and Econometrics* **2**, 1-14.

Hansen, B.E., 1997b, Approximate asymptotic p-values for structural-change tests, *Journal of Business & Economics Statistics* **15**, 60-67.

Kim, J. and D. Pollard, 1990, Cube root asymptotics, *Annals of Statistics* **18**, 191-219.

Kwiatkowski, D., P.C.B. Phillips, P. Schmidt and Y. Shin, 1992, Testing the null hypothesis of stationarity against the alternative of a unit root, *Journal of Econometrics* **54**, 159-178.

Lewis, K.K., 1995, Puzzles in international financial markets, in G. Grossman and K. Rogoff (editors), *The Handbook of International Economics, Vol. 3*, Amsterdam: Elsevier Science, p. 1913-1971.

Leybourne, S.J., T.C. Mills and P.Newbold, 1998, Spurious rejections by Dickey-Fuller tests in the presence of a break under the null, *Journal of Econometrics* **87**, 191-203.

MacKinnon, J.G., 1994, Approximate asymptotic distribution functions for unit-root and cointegration tests, *Journal of Business & Economics Statistics* **12**, 167-176.

Ng, S., en P. Perron, 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.

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

Tong, H., 1983, *Threshold Models in Non-Linear Time Series Analysis*, New York: Springer-Verlag.

Tong, H., 1990, *Non-Linear Time Series: a Dynamical Systems Approach*, Oxford: Oxford University Press.

Tong, H. and K.S. Lim, 1980, Threshold autoregressions, limit cycles, and data (with discussion, *Journal of the Royal Statistical Society* **B 42**, 245-292.

Table 1: Critical values for test of unit root null hypothesis against the CTAR alternative[1]

		$n=100$			$n=250$			$n=500$			$n=1000$			$n=\infty$		
τ_0	τ_1	90%	95%	99%	90%	95%	99%	90%	95%	99%	90%	95%	99%	90%	95%	99%
0.10	0.10	4.41	5.13	6.58	4.59	5.25	6.71	4.72	5.42	6.76	4.79	5.46	7.05	4.89	5.57	7.03
0.10	0.20	4.14	4.83	6.37	4.20	4.88	6.30	4.32	4.99	6.32	4.36	5.00	6.63	4.37	5.04	6.49
0.10	0.30	3.95	4.67	6.24	4.05	4.80	6.16	4.20	4.87	6.22	4.25	4.89	6.55	4.25	4.91	6.43
0.10	0.40	3.87	4.62	6.15	3.99	4.72	6.11	4.12	4.82	6.22	4.18	4.85	6.49	4.21	4.89	6.39
0.10	0.50	3.80	4.53	6.12	3.94	4.68	6.08	4.10	4.76	6.20	4.16	4.83	6.47	4.20	4.89	6.42
0.10	0.60	3.73	4.49	6.06	3.92	4.65	6.05	4.06	4.74	6.20	4.13	4.81	6.47	4.19	4.90	6.43
0.20	0.20	3.80	4.40	5.69	3.64	4.25	5.61	3.71	4.35	5.57	3.72	4.34	5.78	3.54	4.17	5.58
0.20	0.30	3.56	4.22	5.60	3.39	4.02	5.35	3.48	4.12	5.32	3.50	4.14	5.53	3.27	3.85	5.28
0.20	0.40	3.41	4.10	5.46	3.27	3.89	5.22	3.31	3.96	5.23	3.39	4.00	5.37	3.16	3.72	5.19
0.20	0.50	3.31	3.96	5.35	3.18	3.81	5.10	3.23	3.89	5.12	3.29	3.92	5.25	3.10	3.65	5.11
0.20	0.60	3.23	3.86	5.26	3.13	3.75	5.04	3.17	3.82	5.04	3.22	3.86	5.19	3.06	3.62	5.09
0.30	0.30	3.29	3.91	5.22	3.10	3.68	5.04	3.15	3.80	4.98	3.18	3.82	5.08	2.95	3.52	5.00
0.30	0.40	3.10	3.72	4.97	2.92	3.50	4.86	2.98	3.57	4.82	3.03	3.63	4.89	2.76	3.33	4.74
0.30	0.50	2.96	3.56	4.79	2.80	3.37	4.70	2.82	3.41	4.67	2.90	3.49	4.75	2.64	3.20	4.61
0.30	0.60	2.85	3.42	4.71	2.70	3.27	4.48	2.73	3.33	4.51	2.79	3.40	4.65	2.56	3.12	4.50
0.40	0.40	2.90	3.47	4.77	2.70	3.23	4.63	2.74	3.34	4.59	2.81	3.39	4.74	2.55	3.09	4.48
0.40	0.50	2.73	3.24	4.49	2.55	3.07	4.40	2.57	3.13	4.40	2.66	3.23	4.55	2.38	2.93	4.36

[1] Selection of finite sample and asymptotic critical values for the F_n test statistic (18). Finite sample critical values are based on 10000 replications. Asymptotic critical values are obtained from response surfaces $q^p_{\tau_0,\tau_1}(n) = q^p_{\tau_0,\tau_1}(\infty) + b_1 n^{-1} + b_2 n^{-2} + \varepsilon_{\tau_0,\tau_1,i}$, where $q^p_{\tau_0,\tau_1}(n)$ is the pth quantile of the finite sample distribution with n observations ($n = \{50, 75, 100, 150, 200, 250, 500, 1000\}$). The estimate of $q^p_{\tau_0,\tau_1}(\infty)$ provides an estimate of the asymptotic critical value.

Table 2: Parameters of Gamma-distributions that can be used to obtain approximate asymptotic p-values for test of unit root null hypothesis against the CTAR alternative[1]

τ_0	τ_1	r	a	τ_0	τ_1	r	a
0.05	0.05	6.8645	1.7703	0.05	0.10	6.2474	1.7816
0.05	0.15	5.5462	1.6773	0.05	0.20	5.1022	1.5917
0.05	0.25	4.8021	1.5253	0.05	0.30	4.5828	1.4733
0.05	0.35	4.4380	1.4378	0.05	0.40	4.3451	1.4154
0.05	0.45	4.2697	1.3963	0.05	0.50	4.2144	1.3820
0.05	0.55	4.1640	1.3687	0.05	0.60	4.1219	1.3572
0.10	0.05	6.1959	1.7653	0.10	0.10	5.8895	1.9199
0.10	0.15	5.3520	1.8998	0.10	0.20	4.8695	1.8187
0.10	0.25	4.5148	1.7425	0.10	0.30	4.2735	1.6858
0.10	0.35	4.0761	1.6349	0.10	0.40	3.9292	1.5963
0.10	0.45	3.8028	1.5615	0.10	0.50	3.7076	1.5354
0.10	0.55	3.6198	1.5098	0.10	0.60	3.5486	1.4889
0.15	0.05	5.5460	1.6733	0.15	0.10	5.3855	1.9109
0.15	0.15	4.9848	1.9680	0.15	0.20	4.5905	1.9369
0.15	0.25	4.2443	1.8749	0.15	0.30	3.9576	1.8080
0.15	0.35	3.7518	1.7584	0.15	0.40	3.5949	1.7186
0.15	0.45	3.4688	1.6859	0.15	0.50	3.3503	1.6524
0.15	0.55	3.2459	1.6213	0.15	0.60	3.1400	1.5868
0.20	0.05	5.0893	1.5840	0.20	0.10	4.9315	1.8399
0.20	0.15	4.6003	1.9437	0.20	0.20	4.2213	1.9338
0.20	0.25	3.9077	1.8944	0.20	0.30	3.6377	1.8404
0.20	0.35	3.4397	1.7987	0.20	0.40	3.2764	1.7600
0.20	0.45	3.1528	1.7316	0.20	0.50	3.0439	1.7046
0.20	0.55	2.9360	1.6738	0.20	0.60	2.8454	1.6471
0.25	0.05	4.7627	1.5101	0.25	0.10	4.5769	1.7645
0.25	0.15	4.2767	1.8914	0.25	0.20	3.9223	1.9021
0.25	0.25	3.5951	1.8633	0.25	0.30	3.3577	1.8291
0.25	0.35	3.1520	1.7881	0.25	0.40	2.9985	1.7581
0.25	0.45	2.8677	1.7285	0.25	0.50	2.7606	1.7054
0.25	0.55	2.6758	1.6889	0.25	0.60	2.5857	1.6647
0.30	0.05	4.5259	1.4520	0.30	0.10	4.3002	1.6963
0.30	0.15	3.9943	1.8272	0.30	0.20	3.6589	1.8518
0.30	0.25	3.3521	1.8274	0.30	0.30	3.1184	1.7999
0.30	0.35	2.9328	1.7711	0.30	0.40	2.7723	1.7403
0.30	0.45	2.6504	1.7188	0.30	0.50	2.5509	1.7028
0.30	0.55	2.4653	1.6894	0.30	0.60	2.3852	1.6740
0.35	0.05	4.3739	1.4145	0.35	0.10	4.0824	1.6383
0.35	0.15	3.7879	1.7771	0.35	0.20	3.4462	1.8035
0.35	0.25	3.1545	1.7902	0.35	0.30	2.9258	1.7690
0.35	0.35	2.7553	1.7521	0.35	0.40	2.6006	1.7275
0.35	0.45	2.4889	1.7149	0.35	0.50	2.3970	1.7068
0.35	0.55	2.3070	1.6942	0.35	0.60	2.2241	1.6851
0.40	0.05	4.2772	1.3903	0.40	0.10	3.9216	1.5939
0.40	0.15	3.6146	1.7308	0.40	0.20	3.2956	1.7700
0.40	0.25	3.0104	1.7643	0.40	0.30	2.7825	1.7467
0.40	0.35	2.6125	1.7346	0.40	0.40	2.4666	1.7178
0.40	0.45	2.3491	1.7053	0.40	0.50	2.2648	1.7062
0.40	0.55	2.1843	1.7073	0.45	0.05	4.1944	1.3685
0.45	0.10	3.7905	1.5566	0.45	0.15	3.4621	1.6866
0.45	0.20	3.1442	1.7284	0.45	0.25	2.8744	1.7328
0.45	0.30	2.6560	1.7231	0.45	0.35	2.4935	1.7183
0.45	0.40	2.3545	1.7086	0.45	0.45	2.2494	1.7084
0.45	0.50	2.1641	1.7160	0.50	0.05	4.1287	1.3508
0.50	0.10	3.6810	1.5249	0.50	0.15	3.3404	1.6511
0.50	0.20	3.0124	1.6905	0.50	0.25	2.7520	1.7015
0.50	0.30	2.5447	1.7001	0.50	0.35	2.3863	1.7005
0.50	0.40	2.2660	1.7064	0.50	0.45	2.1579	1.7123
0.55	0.05	4.0797	1.3376	0.55	0.10	3.5912	1.4982
0.55	0.15	3.2171	1.6117	0.55	0.20	2.9161	1.6652
0.55	0.25	2.6597	1.6806	0.55	0.30	2.4517	1.6824
0.55	0.35	2.3067	1.6945	0.55	0.40	2.1839	1.7068
0.60	0.05	4.0602	1.3323	0.60	0.10	3.5167	1.4756
0.60	0.15	3.1173	1.5795	0.60	0.20	2.8146	1.6332
0.60	0.25	2.5677	1.6554	0.60	0.30	2.3684	1.6653
0.60	0.35	2.2284	1.6892	0.00	0.00	0.0000	0.0000

[1] This table presents estimates of the paramters r and a of the $\Gamma(r, a)$ distributions that are used to construct approximate p-values for the test of a unit root null hypothesis against the CTAR alternative. Let z denote the value of the F_n-statistic. Then $q^p_{\tau_0,\tau_1}(\infty) \approx \Gamma(z; r, a)$.

Table 3: Power of unit root tests - $n = 100$[1]

		F_n			$F_n(\overline{\gamma})$			ADF			GLS		KPSS[2]		
ϕ_l	ϕ_u	1%	5%	10%	1%	5%	10%	1%	5%	10%	5%	10%	1%	5%	10%
0.95	0.95	6.00	17.98	28.54	2.34	10.58	20.12	2.92	12.72	24.10	20.34	34.74	24.74	37.86	57.80
0.95	0.90	9.32	25.60	38.58	4.04	16.42	28.74	5.02	19.66	33.98	30.18	48.12	30.02	44.28	66.06
0.95	0.80	17.22	38.82	53.00	7.84	26.22	41.82	9.48	30.64	46.76	42.82	61.58	35.50	50.20	72.02
0.95	0.70	25.28	49.86	63.92	11.90	33.96	49.48	14.16	37.46	53.88	49.26	67.38	38.36	53.52	74.16
0.95	0.60	33.66	58.48	72.04	15.86	39.16	54.66	17.46	41.92	58.50	53.80	70.98	40.60	55.20	75.30
0.95	0.50	40.40	65.88	77.18	18.96	43.20	58.58	20.14	45.32	61.64	56.82	73.04	41.60	55.96	76.06
0.90	0.90	15.84	38.98	54.74	8.48	28.86	46.20	10.14	33.82	53.40	48.86	69.40	39.86	54.56	76.16
0.90	0.80	31.66	60.98	76.20	19.00	49.90	69.52	22.64	56.96	75.46	70.34	85.20	49.48	64.04	82.88
0.90	0.70	46.82	75.82	87.22	29.72	64.18	80.36	33.84	69.24	84.82	78.20	90.24	54.10	67.88	85.18
0.90	0.60	60.04	84.82	93.00	39.62	72.96	86.40	42.30	76.46	89.34	82.44	92.44	56.68	70.22	86.44
0.90	0.50	70.34	90.20	95.90	46.98	78.72	90.36	49.10	80.96	92.10	85.78	94.06	58.10	71.38	87.46
0.80	0.80	61.54	88.10	95.22	47.04	83.34	94.14	53.78	88.62	96.82	89.58	96.00	61.66	74.52	90.26
0.80	0.70	81.16	96.72	99.20	69.52	94.94	98.66	75.34	97.08	99.30	94.22	98.12	67.22	78.94	92.78
0.80	0.60	91.50	99.20	99.76	83.06	98.26	99.56	87.18	98.96	99.78	96.36	98.98	69.90	81.18	93.86
0.80	0.50	96.42	99.62	99.98	90.90	99.04	99.86	92.70	99.48	99.90	97.48	99.32	71.96	82.38	94.78
0.70	0.70	95.40	99.62	99.94	90.70	99.32	99.92	94.30	99.78	99.96	97.66	99.48	72.58	83.06	95.18
0.70	0.60	99.04	99.98	99.98	97.58	99.90	100.00	98.70	99.98	99.98	98.72	99.80	75.70	85.74	96.22
0.70	0.50	99.72	99.98	100.00	99.24	100.00	100.00	99.50	99.98	100.00	99.40	99.90	77.36	87.12	96.82
0.60	0.60	99.88	100.00	100.00	99.62	100.00	100.00	99.84	100.00	100.00	99.56	99.90	78.64	88.20	97.24
0.60	0.50	99.98	100.00	100.00	99.94	100.00	100.00	99.98	100.00	100.00	99.80	99.98	80.34	89.40	97.82
0.50	0.50	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.90	99.98	82.38	90.74	98.18

[1] The Table reports the percentage of rejections of the null hypothesis of a unit root at the 1, 5, and 10% significance levels. Artificial time series are generated according to (3), with $\gamma = 0$ and $\varepsilon_t \sim$ i.i.d.$N(0,1)$ for sample size $n = 100$. The Table is based on 5000 replications. F_n denotes the F-statistic (18), $F_n(\overline{\gamma})$ the corresponding statistic of Enders and Granger (1998). ADF en GLS denote the standard Augmented Dickey-Fuller test and the efficient unit root test developed by Elliot *et al.* (1996), respectively. KPSS denotes the test statistic of Kwiatkowski *et al.* (1992). Appropriate critical values are obtained as discussed in Section 4.

[2] To facilitate the comparison of different test statistics, the entries in this column are percentage of non-rejections of the null hypothesis of stationarity.

Table 4: Power of unit-root tests - $n = 250$[1]

		F_n			$F_n(\overline{\gamma})$			ADF			GLS		KPSS[2]		
ϕ_l	ϕ_u	1%	5%	10%	1%	5%	10%	1%	5%	10%	5%	10%	1%	5%	10%
0.95	0.95	31.08	60.76	74.28	13.32	41.86	61.14	16.22	48.24	68.04	58.76	76.42	20.56	31.98	53.76
0.95	0.90	56.38	83.96	92.06	30.88	68.72	84.22	36.32	73.52	87.98	76.38	87.24	27.96	40.82	62.84
0.95	0.80	84.02	96.98	99.04	56.70	87.38	95.78	60.34	88.56	96.40	85.00	92.88	34.36	47.22	69.30
0.95	0.70	93.84	99.22	99.78	70.28	92.70	98.14	71.28	93.20	98.16	88.36	94.66	36.26	50.06	72.40
0.95	0.60	96.80	99.68	99.94	77.34	95.52	98.84	76.50	95.08	98.62	90.08	95.66	37.86	52.36	73.66
0.95	0.50	98.38	99.86	99.96	81.48	96.90	99.20	79.70	96.16	98.88	91.40	96.12	39.24	53.38	74.56
0.90	0.90	87.10	98.38	99.66	69.06	95.44	99.02	75.26	97.24	99.52	89.38	95.70	40.42	53.78	75.40
0.90	0.80	99.24	100.00	100.00	95.96	99.86	99.98	96.68	99.88	100.00	95.62	98.76	50.04	64.70	83.32
0.90	0.70	99.98	100.00	100.00	99.24	99.98	100.00	99.32	100.00	100.00	97.44	99.26	54.66	68.80	86.04
0.90	0.60	100.00	100.00	100.00	99.72	100.00	100.00	99.66	100.00	100.00	98.02	99.52	57.32	70.78	87.52
0.90	0.50	100.00	100.00	100.00	99.86	100.00	100.00	99.82	100.00	100.00	98.48	99.70	58.86	72.10	88.42
0.80	0.80	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.14	99.80	63.92	76.42	90.92
0.80	0.70	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.66	99.92	69.36	80.90	93.52
0.80	0.60	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.82	99.96	71.96	82.80	94.82
0.80	0.50	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.90	99.98	73.54	84.06	95.22
0.70	0.70	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.88	99.98	75.10	85.20	95.80
0.70	0.60	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.94	100.00	77.92	87.34	96.64
0.70	0.50	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.98	100.00	79.36	88.56	97.12
0.60	0.60	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	80.56	89.30	97.42
0.60	0.50	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	82.16	90.62	97.80
0.50	0.50	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	84.02	91.66	98.16

[1] The Table reports the percentage of rejections of the null hypothesis of a unit root at the 1, 5, and 10% significance levels. Artificial time series are generated according to (3), with $\gamma = 0$ and $\varepsilon_t \sim$ i.i.d.$N(0,1)$ for sample size $n = 250$. The Table is based on 5000 replications. F_n denotes the F-statistic (18), $F_n(\overline{\gamma})$ the corresponding statistic of Enders and Granger (1998). ADF en GLS denote the standard Augmented Dickey-Fuller test and the efficient unit root test developed by Elliot *et al.* (1996). KPSS denotes the test statistic of Kwiatkowski *et al.* (1992). Appropriate critical values are obtained as discussed in Section 4.

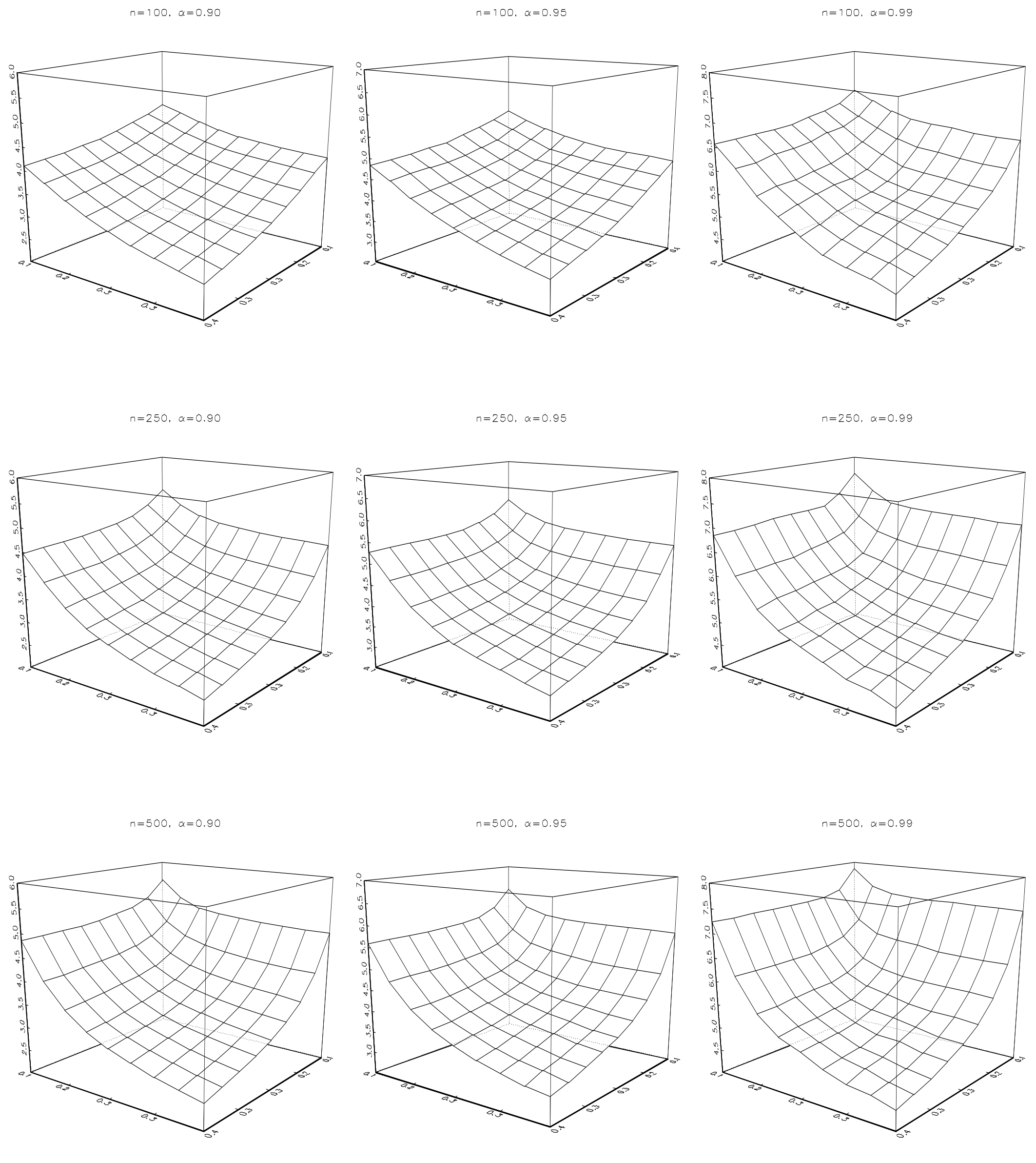
[2] To facilitate the comparison of different test statistics, the entries in this column are percentage of non-rejections of the null hypothesis of stationarity.

Table 5: CTAR model for forward premium[1]

	UK pound	French franc	Japanese yen
ADF	-3.5254	-2.9756	-2.4617
$p_{ADF,b}$	0.0125	0.0370	0.1325
$F_n(\bar{\gamma})$	7.3602	6.7989	4.2418
$p_{F_n(\bar{\gamma}),b}$	0.0130	0.0130	0.0835
F_n	10.5342	15.8499	9.6910
$p_{F_n,\infty}$	0.0000	0.0000	0.0000
$p_{F_n,b}$	0.0245	0.0000	0.0055
τ_0	0.1886	0.3797	0.1690
τ_1	0.5458	0.4014	0.3803
$\hat{\gamma}$	-0.0089	-0.0009	0.0000
$\hat{\phi}_1$	-1.2042	-0.8156	-0.0265
$\hat{\phi}_2$	-0.0748	-0.0712	-0.4273
Wald	8.0901	16.0762	6.9124

[1] $p_{ADF,b}$, $p_{F_n(\bar{\gamma}),b}$, and $p_{F_n,b}$ denote the p-values of the ADF, $F_n(\bar{\gamma})$ and F_n statistics respectively, all based on a nonparametric bootstrap method. $p_{F_n,\infty}$ denotes the approximate asymptotic p-value of the F_n-statistic. The CTAR models are estimated with lagged first differences added to the model. The Ng and Perron (1995) rule indicates that the appropriate number of lagged first differences is 5 in all cases. Significance of lagged first difference variables is assessed at the 5% level. The Wald statistic denotes the Wald test of symmetric versus asymmetric adjustment (see the text for more details), and is asymptotically χ^2_1 distributed.

Figure 1: Finite sample critical values F-test



Note: 90, 95 and 99% percentiles of the small sample distributions of the F_n test (18).