Granger Centre
for time series econometrics

Granger Centre Discussion Paper Series

A powerful test for linearity when the order of integration is unknown

by

David I. Harvey, Stephen J. Leybourne and Bin Xiao

Granger Centre Discussion Paper No. 07/06

The University of Nottingham

A Powerful Test for Linearity when the Order of Integration is Unknown

David I. Harvey, Stephen J. Leybourne and Bin Xiao
School of Economics, University of Nottingham

October 2007

Abstract

In this paper we propose a test of the null hypothesis of time series linearity against a nonlinear alternative, when uncertainty exists as to whether or not the series contains a unit root. We provide a test statistic that has the same limiting null critical values regardless of whether the series under consideration is generated from a linear $I(0)$ or linear $I(1)$ process, and is consistent against nonlinearity of either form, being asymptotically equivalent to the efficient test in each case. Finite sample simulations show that the new procedure has better size control and offers substantial power gains over the recently proposed robust linearity test of Harvey and Leybourne (2007).

Keywords: Nonlinearity testing; Wald tests; Unit root tests; Stationarity tests.

JEL Classifications: C22.

1 Introduction

The recent time series literature has shown that it is often the case that economic variables are better characterised by nonlinear, rather than linear, time series models. It is important therefore, for the purposes of both modelling and forecasting, that reliable tests be available to determine whether a linear or nonlinear model is appropriate for a given series.

The standard tests of the null of linearity against a nonlinear alternative, proposed by Teräsvirta (1994) and Luukkonen *et al.* (1988), rely on an assumption of $I(0)$ behaviour in the underlying series. However, if the series is in fact generated by an $I(1)$ linear process, then spurious rejections of the linearity null hypothesis will occur. In practice, the validity of an $I(0)$ assumption may often be questionable; indeed, it is frequently the case that uncertainty exists as to the order of integration. In an attempt to circumvent this difficulty, Harvey and Leybourne (2007) proposed a test for linearity which does not require an *a priori* assumption as to the order of integration of the process. Their approach involves estimating a hybrid regression allowing for both $I(0)$ and $I(1)$ linear and nonlinear components, and conducting a Wald test for the exclusion of all nonlinear variables. The limiting null distribution of this Wald statistic is different for $I(0)$ series than for $I(1)$ series, but this is overcome by use of a multiplicative scaling factor modification of the kind suggested by Vogelsang (1998). This gives rise to a test statistic that has identical (and standard) critical values for a given significance level under both $I(0)$ and $I(1)$ cases.

In this paper, we continue in the vein of Harvey and Leybourne (2007) to develop tests of linearity that do not require knowledge of the order of integration. Specifically, we propose a test that is comprised of a simple data-dependent weighted average of two Wald test statistics; one that is efficient when the data are generated by an $I(0)$ process and a second that is efficient when the data are $I(1)$. The weights are determined from an auxiliary statistic, which ensures a switch between the two efficient Wald statistics, depending on whether the data are generated by an $I(0)$ or $I(1)$ process. The new weighted statistic has a standard chi-squared limiting null distribution in both the $I(0)$ and $I(1)$ cases, and is shown to have better finite sample size properties and substantially improved power in comparison to the robust test proposed by Harvey and Leybourne (2007). We therefore recommend use of the new test proposed in this paper for practical applications.

The paper is structured as follows. In the next section we present models of linearity and nonlinearity, showing how they approximate specifications such as exponential and logistic smooth transition autoregressive models (ESTAR and LSTAR) in both the $I(0)$ and $I(1)$ contexts. We also describe the efficient Wald statistics that assume knowledge of the order of integration. In Section 3, we develop our robust statistic for testing the hypothesis of testing $I(0)$ or $I(1)$ linearity against nonlinearity, as well as deriving its null asymptotic distribution and establishing consistency under the nonlinear alternative. Section 4 provides finite sample size and power Monte Carlo simulation results, comparing the newly proposed test with the efficient tests (that assume a known order of integration), the Harvey and Leybourne (2007) test, and an alternative approach based on pre-testing for the order of integration. Section 5 presents an application of the tests to short-term interest rate series, while Section 6 concludes.

2 The Model and Standard Linearity Tests

Consider a nonlinear AR(1) model for an $I(0)$ time series y_t, $t = 1, ..., T$

$$\begin{aligned} y_t &= \mu + v_t, \\ v_t &= \rho v_{t-1} + \delta f(v_{t-1}, \theta) v_{t-1} + \varepsilon_t \end{aligned} \tag{1}$$

where ε_t is a zero mean IID white noise process that possesses finite moments up to order 12 (as in Assumption 1 of Harvey and Leybourne, 2007), and where ρ, δ and the function $f(.,\theta)$ are chosen such that v_t is globally stationary. Assuming that the function $f(.,\theta)$ admits a Taylor series expansion around $\theta = 0$, this model can be approximated to the second order by

$$\begin{aligned} y_t &= \mu + v_t, \\ v_t &= \delta_1 v_{t-1} + \delta_2 v_{t-1}^2 + \delta_3 v_{t-1}^3 + \varepsilon_t. \end{aligned} \tag{2}$$

Typical specifications for $f(.,\theta)$ include the well-known ESTAR and LSTAR models

$$\begin{aligned} \text{ESTAR} \quad &: \quad f(v_{t-1},\theta) = 1 - \exp[-\theta(v_{t-1} - c)^2] \\ \text{LSTAR} \quad &: \quad f(v_{t-1},\theta) = \frac{2}{1 + \exp[(-\theta(v_{t-1} - c)]} - 1 \end{aligned}$$

where c is a non-centrality parameter. A second order expansion of the form given in (2) is usually considered sufficient to capture the essential nonlinear features of models such as ESTAR and LSTAR.

In this framework, the null hypothesis of linearity and alternative of nonlinearity can be expressed as

$$\begin{aligned} H_{0,0} \quad &: \quad \delta_2 = \delta_3 = 0 \\ H_{1,0} \quad &: \quad \delta_2 \neq 0 \text{ and/or } \delta_3 \neq 0 \end{aligned}$$

where $H_{.,0}$ denotes a hypothesis under the assumption of y_t being $I(0)$. For the purposes of testing, we can rewrite the DGP (2) as a regression model in terms of the observed y_t

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-1}^2 + \beta_3 y_{t-1}^3 + \varepsilon_t \tag{3}$$

where

$$\beta_0 = \mu(1 - \delta_1 + \delta_2\mu - \delta_3\mu^2), \quad \beta_1 = \delta_1 - 2\delta_2\mu + 3\delta_3\mu^2, \quad \beta_2 = \delta_2 - 3\delta_3\mu, \; \beta_3 = \delta_3.$$

In terms of (3), the null and alternative hypotheses can therefore be stated as

$$\begin{aligned} H_{0,0} \quad &: \quad \beta_2 = \beta_3 = 0 \\ H_{1,0} \quad &: \quad \beta_2 \neq 0 \text{ and/or } \beta_3 \neq 0. \end{aligned}$$

The standard Wald statistic for testing these restrictions is given by

$$W_0 = T\left(\frac{RSS_0^r}{RSS_0^u} - 1\right)$$

where RSS_0^u and RSS_0^r denote, respectively, the residual sums of squares from the unrestricted OLS regression (3) and a restricted OLS regression imposing $\beta_2 = \beta_3 = 0$ in (3), i.e.

$$y_t = \beta_0 + \beta_1 y_{t-1} + \varepsilon_t. \tag{4}$$

Standard large sample theory shows that W_0 will follow an asymptotic $\chi^2(2)$ distribution under the null $H_{0,0}$, and will diverge at the rate $O_p(T)$ under the nonlinear alternative $H_{1,0}$.

Now consider a corresponding nonlinear AR(1) model for an $I(1)$ series, where the nonlinearity is assumed to enter through the first differences of y_t

$$\begin{aligned} y_t &= \mu + v_t, \\ \Delta v_t &= \phi \Delta v_{t-1} + \lambda f(\Delta v_{t-1}, \theta) \Delta v_{t-1} + \varepsilon_t \end{aligned} \tag{5}$$

where ϕ, λ and the function $f(.,\theta)$ are again chosen such that Δv_t is globally stationary. The function $f(.,\theta)$ is used in a generic sense but needs not be identical to that in (1). If the function again allows a Taylor series expansion around $\theta = 0$, (5) can be approximated to the second order by

$$\begin{aligned} y_t &= \mu + v_t, \\ \Delta v_t &= \lambda_1 \Delta v_{t-1} + \lambda_2 (\Delta v_{t-1})^2 + \lambda_3 (\Delta v_{t-1})^3 + \varepsilon_t. \end{aligned} \tag{6}$$

The null of linearity and alternative of nonlinearity are here given by

$$\begin{aligned} H_{0,1} &: \lambda_2 = \lambda_3 = 0 \\ H_{1,1} &: \lambda_2 \neq 0 \text{ and/or } \lambda_3 \neq 0 \end{aligned}$$

where $H_{.,1}$ denotes a hypothesis under the assumption of y_t being $I(1)$. As with the analysis of the $I(0)$ case above, the DGP (6) can be rewritten as a regression model

$$\Delta y_t = \lambda_1 \Delta y_{t-1} + \lambda_2 (\Delta y_{t-1})^2 + \lambda_3 (\Delta y_{t-1})^3 + \varepsilon_t. \tag{7}$$

Notice that since $\Delta y_t = \Delta v_t$, (7) and (6) are essentially identical, thus the null and alternative $H_{0,1}$ and $H_{1,1}$ are as given above.

The corresponding Wald statistic based on (7) is

$$W_1 = T \left(\frac{RSS_1^r}{RSS_1^u} - 1 \right)$$

where RSS_1^u and RSS_1^r denote, respectively, the residual sums of squares from the unrestricted OLS regression (7) and a restricted OLS regression imposing $\lambda_2 = \lambda_3 = 0$ in (7), i.e.

$$\Delta y_t = \lambda_1 \Delta y_{t-1} + \varepsilon_t. \tag{8}$$

Standard theory again shows that W_1 follows an asymptotic $\chi^2(2)$ distribution under the null $H_{0,1}$, and diverges at the rate $O_p(T)$ under the alternative $H_{1,1}$.

When conducting tests such as W_0 and W_1 in practice, we also need to admit the possibility of more general autoregressive structures. Following Kapetanios *et al.* (2003) and Harvey and Leybourne (2007), we assume that this additional serial correlation enters linearly so that the regression models (3), (4), (7) and (8) are replaced with

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-1}^2 + \beta_3 y_{t-1}^3 + \sum_{j=1}^{p} \beta_{4,j} \Delta y_{t-j} + \varepsilon_t \tag{9}$$

$$y_t = \beta_0 + \beta_1 y_{t-1} + \sum_{j=1}^{p} \beta_{4,j} \Delta y_{t-j} + \varepsilon_t \tag{10}$$

$$\Delta y_t = \lambda_1 \Delta y_{t-1} + \lambda_2 (\Delta y_{t-1})^2 + \lambda_3 (\Delta y_{t-1})^3 + \sum_{j=2}^{p} \lambda_{4,j} \Delta y_{t-j} + \varepsilon_t \tag{11}$$

and

$$\Delta y_t = \lambda_1 \Delta y_{t-1} + \sum_{j=2}^{p} \lambda_{4,j} \Delta y_{t-j} + \varepsilon_t \tag{12}$$

respectively. In what follows, p is determined using a general-to-specific methodology, with sequential testing performed at the 10%-level with $p_{\max} = [8(T/100)^{1/4}]$, where $[.]$ denotes integer part.

3 Linearity Testing when the Order of Integration is Unknown

If the order of integration of the time series y_t is known, then the (asymptotically) size controlled and efficient testing strategy is simply to apply W_0 if the series is $I(0)$, and W_1 if the series is $I(1)$. However, the focus of this paper is the situation where it is not known whether y_t is stationary or unit root. We therefore consider an approach which asymptotically selects the W_0 statistic when the data are stationary, and the W_1 statistic when the data have a unit root. This can be achieved by use of a weighted average statistic

$$W_\lambda = \{1 - \lambda\} W_0 + \lambda W_1 \tag{13}$$

where λ is some function that converges in probability (at sufficiently fast rates) to zero when y_t is $I(0)$ and to one when y_t is $I(1)$. This approach has been used by Harvey *et al.* (2007a, 2007b) in the contexts of robust testing for the presence of a trend and a break in trend, respectively. To be operational, a suitable function for λ must be chosen. While many functions could be employed, here we follow Harvey *et al.* (2007a) in using information provided by both unit root and stationarity statistics, along with the following functional form for λ

$$\lambda(U, S) = \exp\left(-g\left(\frac{U}{S}\right)^2\right) \tag{14}$$

where g is some finite positive constant and U and S denote suitably chosen unit root and stationarity statistics. Heuristically, when the series is stationary, $(U/S)^2$ diverges, resulting in λ converging to zero, and when the series is unit root, $(U/S)^2$ converges to zero resulting in λ converging to one. This ensures that W_0 or W_1 is selected by W_λ in (13) as appropriate for the order of integration.

For our purposes, we specify U and S as, respectively, the standard Dickey-Fuller unit root statistic (allowing for a constant) and the nonparametric stationarity statistic proposed by Harris *et al.* (2003). The augmented Dickey-Fuller statistic U is the t-ratio for testing $\pi_1 = 0$ in the following regression

$$\Delta y_t = \pi_0 + \pi_1 y_{t-1} + \sum_{j=1}^{p} \gamma_j \Delta y_{t-j} + \varepsilon_t. \tag{15}$$

As with the Wald tests W_0 and W_1, the number of lagged difference terms used in (15) is determined using a general-to-specific methodology, with sequential testing performed at the 10%-level with $p_{\max} = [8(T/100)^{1/4}]$.

The S statistic is given by

$$S = \frac{T^{-1/2} \sum_{t=k+1}^{T} \tilde{y}_t \tilde{y}_{t-k}}{\hat{\omega}\{\tilde{y}_t \tilde{y}_{t-k}\}} \tag{16}$$

where $\tilde{y}_t$ denotes the deviation y_t from its mean, and $\hat{\omega}^2\{a_{t,k}\}$ is the Bartlett kernel-based long run variance estimator of a sequence of variables $a_{1,k}, ..., a_{T,k}$ defined by

$$\hat{\omega}\{a_{t,k}\}^2 = \hat{\gamma}_0\{a_{t,k}\} + 2\sum_{j=1}^{l}\left(1 - \frac{j}{l}\right)\hat{\gamma}_j\{a_{t,k}\}, \quad \hat{\gamma}_j\{a_{t,k}\} = T^{-1}\sum_{t=j+k+1}^{T} a_{t,k}a_{t-j,k} \tag{17}$$

with $a_{t,k} = \tilde{y}_t \tilde{y}_{t-k}$ for (16). In the computation of S, we set $k = (2T)^{1/2}$ and $l = 12(T/100)^{1/4}$ in (16) and (17), rounded to the nearest integer.

The asymptotic behaviour of W_0, W_1 and $\lambda(U, S)$ is given by the following Lemmas, the proofs of which are contained in the Appendix

Lemma 1 *(a) If y_t is $I(0)$, then under $H_{0,0}$ and $H_{1,0}$: $|U|$ diverges to $+\infty$, $|S| = O_p(1)$, and $\lambda(U, S) \xrightarrow{p} 0$ at a rate faster than $O_p(T^{-\gamma})$ for any finite $\gamma > 0$; (b) If y_t is $I(1)$, then under $H_{0,1}$ and $H_{1,1}$: $|U| = O_p(1)$, $|S|$ diverges to $+\infty$, and $\lambda(U, S) \xrightarrow{p} 1$.*

Lemma 2 *(a) If y_t is $I(0)$, then (i) under $H_{0,0}$: $W_0 \Rightarrow \chi^2(2)$ and $W_1 = O_p(T)$, (ii) under $H_{1,0}$: $W_0 = O_p(T)$ and $W_1 = O_p(T)$; (b) If y_t is $I(1)$, then (i) under $H_{0,1}$: $W_0 = O_p(1)$ and $W_1 \Rightarrow \chi^2(2)$, (ii) under $H_{1,1}$: $W_0 = O_p(1)$ and $W_1 = O_p(T)$.*

The limit behaviour of the test statistic W_λ can then be summarised as follows

Theorem 1 *(a) Under the null of $I(0)$ linearity ($H_{0,0}$), $W_\lambda = W_0 + o_p(1) \Rightarrow \chi^2(2)$ and under $I(1)$ linearity ($H_{0,1}$), $W_\lambda = W_1 + o_p(1) \Rightarrow \chi^2(2)$; (b) Under the alternative of $I(0)$ nonlinearity ($H_{1,0}$), $T^{-1}W_\lambda = T^{-1}W_0 + o_p(1)$ and under $I(1)$ nonlinearity ($H_{1,1}$), $T^{-1}W_\lambda = T^{-1}W_1 + o_p(1)$.*

The proof is given in the Appendix. This implies that under the null of either $I(0)$ or $I(1)$ linearity, W_λ selects the efficient test in the limit, and is asymptotically distributed $\chi^2(2)$. Under the nonlinear alternative of either $I(0)$ or $I(1)$ nonlinearity,

W_λ again selects the efficient test in the limit in each case, and is thus consistent at the rate $O_p(T)$.

The value of the constant g in (14) has no effect on the asymptotic properties of W_λ, however, its setting controls the rate at which W_λ switches between W_0 and W_1 in finite samples. As a result, its value needs to be calibrated on the basis of finite sample simulations. We simulated the size of the W_λ test for a grid of g values, for both $I(0)$ and $I(1)$ processes, and for the sample sizes $T = 150$ and $T = 300$. Overall, we found that the empirical sizes were closest to nominal size for the value $g = 0.1$; this value is hence recommended and employed throughout the remainder of the paper.

4 Finite Sample Properties

In this section we use Monte Carlo simulation to investigate the finite sample size and power behaviour of the W_λ linearity test. In order to gauge the performance of this new test, we also report results for W_0 and W_1, the robust test of Harvey and Leybourne (2007), W^*, and a two-step procedure comprised of (i) pre-testing for the order of integration, and (ii) conducting either W_0 or W_1, depending on the outcome of the pre-test. For the latter, in the first step we employed the stationarity test of Harris *et al.* (2003), given by (16) above, which has the attractive property of being asymptotically standard normally distributed under the null of linear or nonlinear stationarity ($H_{0,0}$ or $H_{1,0}$); see the proof of Lemma 1 (a) in the Appendix. If S was found to reject the null (using one-sided upper-tail standard normal 5%-level critical values), then W_1 was conducted; if no rejection was observed, W_0 was selected. This pre-test procedure is denoted W_P. All Monte Carlo simulations reported in this section were computed using 20,000 replications, for linearity tests conducted at the nominal 5%-level.

4.1 Size

First we consider empirical size behaviour using the following DGP

$$(1 - \rho L)y_t = (1 - \psi L)\varepsilon_t$$

with $\varepsilon_t \sim NID(0, 1)$. Simulations were conducted for the sample sizes $T = \{150, 300\}$, with the parameter settings $\rho = \{0, 0.8, 0.9, 0.95, 1\}$ and $\psi = \{0, \pm 0.3, \pm 0.6, \pm 0.9\}$. Table 1 reports the results. The W_λ test is very well behaved under the null hypothesis of linearity, with very little evidence of size distortion for either the $I(0)$ or $I(1)$ cases, even for $T = 150$. Across the range of DGP parameters considered, there are a few combinations of ρ and ψ which give rise to empirical size deviating a little from nominal size, but these deviations are infrequent and never greater than $\pm 2\%$. Overall, therefore, we consider that the test is robust to the order of integration and to general forms of serial correlation, displaying very attractive size properties.

By contrast, the W_0 and W_1 tests that assume knowledge of the order of integration are not robust to the degree of persistence in the series. For the sample sizes considered, the W_0 test is approximately correctly sized when $\rho = 0$, but becomes increasingly over-sized as ρ approaches one (the only exception is when $\psi = 0.9$ where (near) cancelling roots are present). In the $I(1)$ case this over-size is quite severe, being around 15% for the central case of $\psi = 0$. On the other hand, the W_1 test has approximately correct size in the $I(1)$ case ($\rho = 1$) but suffers from under-size when $\rho < 1$. However, given that the test does not exhibit over-size, the W_1 test could be viewed as a conservative test for linearity when uncertainty exists as to the order of integration.

The W^* test of Harvey and Leybourne (2007) displays generally good size control, as was shown in the simulation results of that paper. However, for $T = 150$, some noticeable size distortions are present, and although the degree of oversize is relatively modest, W_λ is shown to exhibit better size behaviour here. For the pre-test procedure W_P, substantial over-size is seen for $T = 150$, particularly for unit root and near unit root series. Although this improves for $T = 300$, size distortions still persist. In addition, due to the inevitable Type I error in the first stage of the pre-test when the series is $I(0)$, there is no guarantee that W_P will be asymptotically correctly sized for linear stationary time series. These undesirable size properties raise significant questions over the value of using such a pre-test procdure in practice. On the basis of these finite sample size simulations, it is clearly the W_λ test that is the best performing test.

4.2 Power

We also examine the power of the linearity tests against a number of nonlinear alternatives. First we consider both forms of nonlinearity discussed earlier, namely LSTAR and ESTAR. As with the size simulations, we consider both $I(0)$ and $I(1)$ data, generated according to the following processes

$$\begin{aligned} I(0) \quad &: \quad (1-\phi L)y_t = \delta f(y_{t-1}, \theta) y_{t-1} + \varepsilon_t, \\ I(1) \quad &: \quad (1-\phi L)\Delta y_t = \lambda f(\Delta y_{t-1}, \theta)\Delta y_{t-1} + \varepsilon_t. \end{aligned}$$

Note that here we abstract from the effects of moving average errors, and we assumed knowledge of the dynamic structure when computing the tests by setting $p = 0$ in (9) and (15) when the data are $I(0)$, and $p = 1$ when the series are $I(1)$. A range of nonlinearity parameter values θ are reported for several representative combinations of ϕ and δ, λ, all for the sample sizes $T = \{150, 300\}$. The ϕ, δ and λ parameters are chosen to ensure global stationarity of y_t (Δy_t), i.e. given $\phi \geq 0$, we ensure $|\phi \pm \delta| < 1$ ($|\phi \pm \lambda| < 1$) for stationary (unit root) LSTAR models and $|\phi + \delta| < 1$ ($|\phi + \lambda| < 1$) for stationary (unit root) ESTAR models. The ESTAR condition allows for locally unit root ($\phi = 1$) or even explosive ($\phi > 1$) behaviour, while maintaining global stationarity.

The results are reported in Tables 2 and 3 for LSTAR and ESTAR models respectively. For the LSTAR models, when y_t is $I(0)$, W_λ has power almost identical to that of the efficient test W_0, even for the smaller sample size $T = 150$. The pre-test procedure W_P also displays good power performance, although the rejection frequencies are almost always below those of W_λ, and the W_P test was also shown to have poor size control. The power of the W^* test lies substantially below that of W_0 and W_λ, and while the W_1 test can be considered as a conservative robust linearity test, its power performance is extremely poor, as might be anticipated. When y_t is $I(1)$, the power of W_λ is very slightly lower than that of W_1 in the smaller sample simulations, but, as would be expected given its asymptotic efficiency properties, the powers are again very close for $T = 300$. In contrast, W^* has power well below that of W_λ and W_1 for all sample sizes, and in many cases the relative power losses are considerable. W_P again performs well, being broadly competitive with the better sized W_λ, although as before, W_P is almost always outperformed by W_λ. The W_0 test has very low power when y_t is $I(1)$ and the results clearly highlight its lack of consistency in this situation, as shown in Lemma 2 (b) (ii).

Turning now to the results for ESTAR models, in the $I(0)$ cases, much the same can be said as for the LSTAR models, with W_λ having power close to that of the efficient test W_0 and clear superiority over W^*, while the conservative W_1 test again displays very low rejection frequencies. W_P is more competitive for $T = 150$ than was observed for LSTAR models, and actually outpeforms W_λ on many occasions. Again however, it must be stressed that W_P lacks proper size control; in addition, there are regions of the parameter space where W_λ has very large power gains over W_P: see results for $\phi = 1.5$, $\delta = -1.0$, $\theta = 0.1$. When y_t is $I(1)$, the power of W_λ is again markedly superior to that of W^*, and compares very well with the efficient test W_1. In fact, there are a number of cases where the power of W_λ exceeds that of W_1. This finite sample artefact arises because W_0, despite its lack of consistency, has substantial finite sample power for many of the processes considered here, thus W_λ, derived as a weighted average of W_0 and W_1, can on occasion achieve additional rejections of the null hypothesis than would be realised by simply using W_1. On comparing W_λ with W_P, it can be seen that W_λ now often demonstrates significant power advantages, up to the order of around 25%.

We now consider the behaviour of the linearity tests under a different nonlinear alternative specification, one that is not nested in our model framework. Specifically, we consider a self-exciting threshold autoregressive (SETAR) model, for both $I(0)$ and $I(1)$ cases, given by

$$I(0) \quad : \quad \begin{array}{ll} y_t = \phi y_{t-1} + \varepsilon_t & \text{if } y_{t-1} \leq 0 \\ y_t = \delta y_{t-1} + \varepsilon_t & \text{if } y_{t-1} > 0 \end{array},$$

$$I(1) \quad : \quad \begin{array}{ll} \Delta y_t = \phi \Delta y_{t-1} + \varepsilon_t & \text{if } \Delta y_{t-1} \leq 0 \\ \Delta y_t = \lambda \Delta y_{t-1} + \varepsilon_t & \text{if } \Delta y_{t-1} > 0 \end{array}.$$

The SETAR model can be obtained as an LSTAR as $\theta \to \infty$, but the nonlinear component cannot be written using a function $f(., \theta)$ that admits a Taylor series expansion around $\theta = 0$, as specified in the model setup. It is interesting, therefore, to examine whether the tests have power against such alternatives, acting simply as tests of model mis-specification.

Table 4 reports results for the parameter values $\phi = 0$ and $\delta, \lambda = \{0.3, 0.5, 0.7, 0.9\}$ (ensuring stationarity of y_t for $I(0)$ series and Δy_t for $I(1)$ series), for the sample sizes

$T = \{150, 300\}$. First, it can be seen that the tests do have power against SETAR alternatives. As with the results for the LSTAR and ESTAR simulations, highest power is observed for the W_0 test when the series is $I(0)$, and for the W_1 test when the series is $I(1)$, as might be expected, but these tests perform very poorly when applied in the reverse contexts, i.e. W_0 for $I(1)$ data, and W_1 for $I(0)$ data. Once again, W_λ has power that is very close to W_0 for $I(0)$ series and W_1 for $I(1)$ series, making it the best-performing test. The W_P procedure is also very competitive (but as we have shown earlier lacks size control), while W^* suffers power losses relative to W_λ.

In summary, our finite sample simulations allow us to conclude that W_λ displays good finite sample size control, has power that compares very favourably with the efficient tests that assume knowledge of the order of integration, and offers substantial power improvements over the robust W^* test proposed by Harvey and Leybourne (2007). We therefore recommend that the new W_λ test be used in practice.

5 Empirical Application

In this section we test for linearity in short-term interest rates. Many empirical studies have found evidence for nonlinearity in the evolution of short-term interest rate series over time; see, *inter alia*, Hamilton (1988), Gray (1996), Barkoulas et al. (1997), Ang and Bekaert (2002a, 2002b) and Audrino (2006). However, it is unclear as to whether such series are best modelled by unit root or stationary processes, thus application of our robust tests to such series seems particularly apposite, since they remain agnostic as to the true order of integration.

The data we consider are 3-month interest rates for Australia, Canada, New Zealand, Switzerland, UK, and US, monthly over the period 1978:1–2006:12 (348 observations). The data were obtained from the OECD Statistics database. We apply the new test W_λ and the W^* test of Harvey and Leybourne (2007) to each series at the 10%, 5% and 1% significance levels.[1] Note that the W^* test requires re-computation of the test statistic when run at different significance levels, since the construction of the statistic involves use of a significance level-dependent parameter.

The results are reported in Table 5. We find that both W_λ and W^* reject the null of linearity at the 1%-level for Australia, Canada, New Zealand and US, providing strong evidence of nonlinearity for short-term interest rates in these countries. Of particular interest are the results for the UK, where W_λ rejects at the 5%-level, but no rejection is obtained, even at the 10%-level, when using W^*. On the basis of the simulation evidence of the previous section, where W_λ was found to display better size and more power than W^*, we would infer that it is the result of the W_λ test that is to be believed here, indicating nonlinearity for the UK short-term interest rate. In the case of Switzerland, no evidence of nonlinearity is detected by either of the tests at conventional significance levels, thus Swiss interest rates appear to exist as an exception

[1] The W_0 and W_1 tests are not applied since they lack robustness to the order of integration; also, W_P is not considered since it lacks proper size control.

to the general finding of short-term interest rate nonlinearity.

6 Conclusion

In this paper we have considered the issue of testing the null hypothesis of linearity against a nonlinear alternative, in the practical scenario where the order of integration of the data is unknown. We provide a new test, W_λ, that is asymptotically efficient, in the sense that for both $I(0)$ and $I(1)$ series, W_λ is asymptotically equivalent to the efficient test that assumes knowledge of the order of integration. Monte Carlo simulation shows that the new test has very attractive size properties for varying degrees of persistence, while at the same time forsaking very little in terms of finite sample power relative to the efficient tests. Moreover, it is shown to have superior finite sample size and power properties to the recently proposed robust test of Harvey and Leybourne (2007), as well as outperforming an alternative procedure based on pre-testing for the order of integration. We therefore recommend that the W_λ test proposed in this paper be used in practical applications.

Throughout this paper we have assumed that the delay parameter in the function $f(.,\theta)$ is one period, and that additional serial correlation enters the models linearly rather than nonlinearly. These assumptions make the analysis tractable and are consistent with related work such as Kapetanios *et al.* (2003) and Harvey and Leybourne (2007), also allowing straightforward comparisons with the latter. However, in future research it would also be interesting to consider developing tests where these assumptions are relaxed. For example, the general approach taken in this paper to achieving robust linearity tests could also be applied to models where the delay parameter is treated as unknown (but bounded). Equivalent tests to W_0 and W_1 could first be defined using regressions along the lines of those considered by Luukkonen *et al.* (1988), and then a robust statistic that switches between them in the limit could be constructed using the same principle as in W_λ of this paper.

Appendix

Proof of Lemma 1

(a) Under $H_{0,0}$, the series is linear and stationary, and so $U = O_p(T^{1/2})$ by standard results. Under $H_{1,0}$, the series is nonlinear and stationary; however, since $E(\Delta y_t) = 0$ still obtains, we find for the sample moments involved in the construction of the Dickey-Fuller statistic U, the following orders hold

$$\begin{aligned}
\textstyle\sum_{t=p+1}^{T} \left(\Delta y_{t-i} - \overline{\Delta y_{-i}}\right)\left(y_{t-1} - \overline{y_{-1}}\right) &= O_p(T) \quad \forall i \leq p, \\
\textstyle\sum_{t=p+1}^{T} \left(y_{t-1} - \overline{y_{-1}}\right)^2 &= O_p(T), \\
\textstyle\sum_{t=p+1}^{T} \left(\Delta y_{t-i} - \overline{\Delta y_{-i}}\right)\left(\Delta y_{t-j} - \overline{\Delta y_{-j}}\right) &= O_p(T) \quad \forall i, j \leq p.
\end{aligned}$$

These combine to show that $U = O_p(T^{1/2})$. Thus, $|U|$ diverges to $+\infty$ as $T \to \infty$ under both $H_{0,0}$ and $H_{1,0}$.

Under $H_{0,0}$, Theorem 8 (with $n = 1$) of Harris *et al.* (2003) shows that $S \Rightarrow N(0, 1)$. The same result can be obtained for S under $H_{1,0}$ on replacing Assumption LP with a suitable mixingale condition.

It therefore follows under both $H_{0,0}$ and $H_{1,0}$ that

$$\begin{aligned}
\lambda(U, S) &= \exp\left(-g\left(\frac{U}{S}\right)^2\right) \\
&= \frac{1}{1 + g\left(\frac{U}{S}\right)^2 + \frac{g^2}{2}\left(\frac{U}{S}\right)^4 + \frac{g^3}{6}\left(\frac{U}{S}\right)^6 + \ldots} \\
&= \frac{1}{1 + g\left|O_p(T)\right| + \frac{g^2}{2}\left|O_p(T^2)\right| + \frac{g^3}{6}\left|O_p(T^3)\right| + \ldots} \\
&\xrightarrow{p} 0.
\end{aligned}$$

The infinite expansion in the denominator shows that $\lambda(U, S) \xrightarrow{p} 0$ at an exponential rate, that is, a rate faster than $O_p(T^{-\gamma})$ for any finite $\gamma > 0$.

(b) Under $H_{0,1}$, the series is linear and unit root, and so $U = O_p(1)$, since it converges to the Dickey-Fuller limit distribution. Under $H_{1,1}$, the series is nonlinear and unit root. In this case, $E(\Delta y_t) \neq 0$, because expanding for $E(\Delta y_t)$ using (7) involves terms such as $E[(\Delta y_{t-1})^2]$. This has the effect of inducing a drift into y_t, which the Dickey-Fuller statistic neglects to take into consideration. The corresponding sample moments involved in U now have the following orders

$$\begin{aligned}
\textstyle\sum_{t=p+1}^{T} \left(\Delta y_{t-i} - \overline{\Delta y_{-i}}\right)\left(y_{t-1} - \overline{y_{-1}}\right) &= O_p(T^{3/2}) \quad \forall i \leq p, \\
\textstyle\sum_{t=p+1}^{T} \left(y_{t-1} - \overline{y_{-1}}\right)^2 &= O_p(T^3), \\
\textstyle\sum_{t=p+1}^{T} \left(\Delta y_{t-i} - \overline{\Delta y_{-i}}\right)\left(\Delta y_{t-j} - \overline{\Delta y_{-j}}\right) &= O_p(T) \quad \forall i, j \leq p.
\end{aligned}$$

which combine to give $U = O_p(1)$.[2] Thus, $|U| = O_p(1)$ under both $H_{0,1}$ and $H_{1,1}$.

Under $H_{0,1}$, Lemma 1 (with $N = 1$) of Harris *et al.* (2005) shows that $|S|$ diverges to $+\infty$. The same result can be shown to hold under $H_{1,1}$ due to the neglected drift term induced by $E(\Delta y_t) \neq 0$.

Then, under both $H_{0,1}$ and $H_{1,1}$

$$\begin{aligned} \lambda(U,S) &= \exp\left(-g\left(\frac{U}{S}\right)^2\right) \\ &= \exp\left(-\left|o_p(1)\right|\right) \\ &\xrightarrow{p} 1 \end{aligned}$$

as stated.

Proof of Lemma 2

(a) (i) Under $H_{0,0}$, the series is stationary and $\beta_2 = \beta_3 = 0$, and so $W_0 \Rightarrow \chi^2(2)$ by standard central limit theory. With regard to W_1, the unrestricted model (11) and restricted model (12) are constructed using first differences, and are therefore mis-specified since neither nests the true stationary DGP. This results in the population values λ_2 and λ_3 being non-zero, and thus $\hat{\lambda}_2$ and $\hat{\lambda}_3$ converge in probability to non-zero constants. As a result, $RSS_1^r/RSS_1^u = O_p(1)$ but does not converge in probability to one. Thus

$$\begin{aligned} W_1 &= T[O_p(1) - 1] \\ &= O_p(T). \end{aligned}$$

(a) (ii) Under $H_{1,0}$, the series is stationary and nonlinear, with $\beta_2 \neq 0$ and/or $\beta_3 \neq 0$, and so $W_0 = O_p(T)$ by standard consistency arguments. For W_1, the unrestricted and restricted models are again mis-specified, with neither nesting the stationary DGP. As in (a) (i) above, λ_2 and λ_3 are non-zero, and $\hat{\lambda}_2$ and $\hat{\lambda}_3$ converge in probability to non-zero values. It again follows that $RSS_1^r/RSS_1^u = O_p(1)$, not converging in probability to one, and hence $W_1 = O_p(T)$.

(b) (i) Under $H_{0,1}$, the series is unit root and $\lambda_2 = \lambda_3 = 0$, and so $W_1 \Rightarrow \chi^2(2)$ by standard limit theory. Although the series is unit root, the regressions involved in the computation of W_0, i.e. (9) and (10), nest the DGP, and hence W_0 has the standard order $O_p(1)$ (but does not converge in distribution to a $\chi^2(2)$ random variable because of the unit root in y_t).

(b) (ii) Under $H_{1,1}$, the series is unit root and nonlinear, with $\lambda_2 \neq 0$ and/or $\lambda_3 \neq 0$, and so $W_1 = O_p(T)$ by standard consistency arguments. In respect of W_0, both the

[2] This result is analagous to that of West (1988), where the effects of a neglected drift on the Dickey-Fuller statistic are considered.

unrestricted model (9) and restricted model (10) are mis-specified since neither nests the true DGP. However, in this case, the population values β_2 and β_3 are zero. This arises since the DGP behaves as a unit root model with drift (due to $E(\Delta y_t) \neq 0$), and although the unrestricted regression model is mis-specified, it does contain a constant and y_{t-1}, the corresponding coefficients of which (β_0 and β_1) account for the drift and unit root. Thus the mis-specification involves an attempt to model nonlinear dynamics in the (stationary) differences ($\Delta y_{t-1}^2, \Delta y_{t-1}^3$) using the (unit root) levels (y_{t-1}^2, y_{t-1}^3). The regressors are then of a different order to the variables in the DGP and the relevant coefficients are zero in population. As a consequence, $\hat{\beta}_2 \xrightarrow{p} 0$ and $\hat{\beta}_3 \xrightarrow{p} 0$, and $RSS_0^r - RSS_0^u = O_p(1)$. Since $RSS_0^u = O_p(T)$, we then have

$$\begin{aligned} W_0 &= \frac{RSS_0^r - RSS_0^u}{T^{-1}RSS_0^u} \\ &= \frac{O_p(1)}{O_p(1)} \\ &= O_p(1). \end{aligned}$$

Proof of Theorem 1

(a) Under $H_{0,0}$, the results of Lemma 1 (a) show that $\lambda(U,S) \xrightarrow{p} 0$ at an exponential rate, and the results of Lemma 2 (a) (i) give $W_0 \Rightarrow \chi^2(2)$ and $W_1 = O_p(T)$. It then follows that

$$\begin{aligned} W_\lambda &= \{1 - \lambda(U,S)\}W_0 + \lambda(U,S)W_1 \\ &= \{1 + o_p(T^{-\gamma})\}W_0 + o_p(T^{-\gamma})O_p(T) \qquad \forall \gamma > 0 \\ &= W_0 + o_p(1) \\ &\Rightarrow \chi^2(2). \end{aligned}$$

Under $H_{0,1}$, Lemma 1 (b) gives $\lambda(U,S) \xrightarrow{p} 1$, and Lemma 2 (b) (i) shows $W_0 = O_p(1)$ and $W_1 \Rightarrow \chi^2(2)$. So

$$\begin{aligned} W_\lambda &= o_p(1)O_p(1) + \{1 + o_p(1)\}W_1 \\ &= W_1 + o_p(1) \\ &\Rightarrow \chi^2(2). \end{aligned}$$

(b) Under $H_{1,0}$, the results of Lemma 1 (a) show that $\lambda(U,S) \xrightarrow{p} 0$ at an exponential rate, and Lemma 2 (a) (ii) gives $W_0 = O_p(T)$ and $W_1 = O_p(T)$. We find

$$\begin{aligned} T^{-1}W_\lambda &= \{1 - \lambda(U,S)\}T^{-1}W_0 + \lambda(U,S)T^{-1}W_1 \\ &= \{1 + o_p(T^{-\gamma})\}T^{-1}W_0 + o_p(T^{-\gamma})O_p(1) \\ &= T^{-1}W_0 + o_p(1). \end{aligned}$$

Under $H_{1,1}$, Lemma 1 (b) gives $\lambda(U,S) \xrightarrow{p} 1$, and Lemma 2 (b) (ii) shows $W_0 = O_p(1)$ and $W_1 = O_p(T)$. Hence

$$\begin{aligned} T^{-1}W_\lambda &= o_p(1)O_p(T^{-1}) + \{1 + o_p(1)\}T^{-1}W_1 \\ &= T^{-1}W_1 + o_p(1). \end{aligned}$$

References

Ang, A. and Bekeart, G. (2002a), "Regime switches in interest rates," *Journal of Business and Economic Statistics* 20, 163-182.

Ang, A. and Bekeart, G. (2002b), "Short rate nonlinearities and regime switches," *Journal of Economic Dynamics and Control* 26, 1243-1274.

Audrino, F. (2006), "Tree-structured multiple regimes in interest rates," *Journal of Business and Economic Statistics* 24, 338-353.

Barkoulas, J. T., Baum, C. F. and Onochie, J. (1997), "A nonparametric investigation of the 90-day T-bill rate," *Review of Financial Economics* 6, 187-198.

Gray, S. F. (1996), "Modeling the conditional distribution of interest rates as a regime-switching process," *Journal of Financial Economics* 42, 27-62.

Hamilton, J. D. (1988), "Rational-expectations econometric analysis of changes in regime: An investigation of the term structure of interest rates," *Journal of Economic Dynamics and Control* 12, 385-423.

Harris, D., Leybourne, S. J. and McCabe, B. P. (2005), "Panel stationarity tests for purchasing power parity with cross-sectional dependence," *Journal of Business and Economic Statistics* 23, 395-409.

Harris, D., McCabe, B. P. and Leybourne, S. J. (2003), "Some limit theory for autocovariances whose order depends on sample size," *Econometric Theory* 19, 829-864.

Harvey, D. I. and Leybourne, S. J. (2007), "Testing for time series linearity," *Econometrics Journal*, 10, 149-165.

Harvey, D. I., Leybourne, S. J. and Taylor (2007a), "A simple, robust and powerful test of the trend hypothesis," *Journal of Econometrics* 141, 1302-1330.

Harvey, D. I., Leybourne, S. J. and Taylor (2007b), "Simple, robust and powerful tests of the breaking trend hypothesis," *Econometric Theory*, forthcoming.

Kapetanios, G., Shin, Y. and Snell, A. (2003), "Testing for a unit root in the nonlinear STAR framework," *Journal of Econometrics* 112, 359-379.

Luukkonen, R., Saikkonen, P. and Teräsvirta, T. (1988), "Testing linearity against smooth transition autoregressive models," *Biometrika* 75, 491-499.

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

Vogelsang, T. J. (1998), "Trend function hypothesis testing in the presence of serial correlation," *Econometrica* 66, 123-148.

West, K. D. (1988), "Asymptotic normality, when regressors have a unit root," *Econometrica* 56, 1397-1417.

Table 1. Empirical sizes of nominal 5%-level linearity tests.

		$T = 150$					$T = 300$				
ρ	ψ	W_λ	W_P	W^*	W_0	W_1	W_λ	W_P	W^*	W_0	W_1
0.00	-0.9	5.0	6.2	5.9	6.3	3.6	4.4	5.2	5.0	5.3	3.2
	-0.6	4.8	5.5	5.9	5.6	3.5	4.8	5.2	5.0	5.3	3.2
	-0.3	4.9	5.5	6.2	5.6	3.9	4.7	5.0	5.1	5.0	3.3
	0.0	5.8	6.2	6.2	6.3	3.6	5.2	5.3	5.2	5.4	3.2
	0.3	6.0	6.3	6.3	6.4	3.6	5.6	5.7	5.2	5.7	3.2
	0.6	5.9	6.0	6.3	6.2	3.3	5.5	5.6	5.6	5.7	3.2
	0.9	6.3	6.4	6.9	6.5	3.3	5.7	5.7	5.7	5.8	3.2
0.80	-0.9	5.3	7.9	6.5	8.0	3.7	4.4	6.3	5.2	6.3	3.2
	-0.6	5.2	7.5	6.3	7.6	3.4	4.8	6.2	5.2	6.3	3.0
	-0.3	5.5	7.5	6.3	7.5	3.7	5.0	6.2	5.3	6.3	3.0
	0.0	4.8	6.4	5.8	6.4	3.6	4.6	5.6	5.2	5.7	3.1
	0.3	4.6	5.9	5.6	6.0	3.8	4.3	5.1	5.0	5.2	3.4
	0.6	4.8	5.6	5.7	5.7	3.7	4.6	5.2	4.9	5.3	3.4
	0.9	6.1	6.4	6.6	6.5	3.6	5.4	5.5	5.4	5.7	3.3
0.90	-0.9	5.6	9.4	7.2	9.7	3.7	4.6	6.9	5.7	7.3	3.3
	-0.6	5.5	9.2	7.0	9.5	3.5	4.8	7.0	5.6	7.3	3.1
	-0.3	5.8	9.2	7.0	9.6	3.5	4.9	6.9	5.5	7.3	3.0
	0.0	4.9	7.8	6.4	8.2	3.7	4.3	6.4	5.4	6.7	3.2
	0.3	4.7	6.9	5.7	7.2	3.7	3.9	5.6	5.2	5.8	3.3
	0.6	4.0	5.3	5.3	5.4	3.7	3.6	4.7	4.6	4.8	3.3
	0.9	5.8	6.2	6.2	6.3	3.6	5.2	5.3	5.2	5.4	3.2
0.95	-0.9	5.7	10.4	7.9	12.3	3.8	4.4	7.4	6.0	9.0	3.3
	-0.6	5.5	10.0	7.8	12.0	3.5	4.4	7.4	6.1	9.2	3.1
	-0.3	5.3	9.9	7.9	12.1	3.3	4.5	7.4	6.1	9.4	3.1
	0.0	4.8	8.6	7.3	10.4	3.9	4.0	6.8	5.9	8.5	3.1
	0.3	4.6	7.9	6.6	9.5	3.7	3.7	6.2	5.4	7.4	3.3
	0.6	3.6	5.4	5.3	5.9	3.6	3.1	4.7	4.5	5.1	3.4
	0.9	5.3	6.0	6.0	6.1	3.7	4.7	5.2	5.1	5.4	3.3
1.00	-0.9	6.6	9.5	7.7	17.4	6.5	5.4	6.6	6.5	15.9	5.6
	-0.6	5.9	9.0	7.5	17.0	5.9	5.2	6.3	6.5	15.9	5.3
	-0.3	5.8	8.8	7.6	17.2	5.8	5.0	6.0	6.6	15.9	5.1
	0.0	5.8	8.4	7.3	15.7	6.1	5.1	6.2	6.3	15.2	5.4
	0.3	5.8	8.2	7.0	14.9	5.9	5.1	6.0	6.5	14.9	5.3
	0.6	4.9	6.8	5.5	9.8	5.5	4.9	5.8	5.2	11.3	5.3
	0.9	4.3	5.4	5.5	6.0	4.1	3.4	4.6	4.2	5.3	3.9

Table 2. Estimated powers of nominal 5%-level linearity tests: LSTAR, $\phi = 0$.

Panel A. $I(0)$

		$T=150$					$T=300$				
δ	θ	W_λ	W_P	W^*	W_0	W_1	W_λ	W_P	W^*	W_0	W_1
0.7	0.1	8.1	7.9	7.0	8.1	3.3	10.9	10.6	8.7	10.9	4.1
	0.3	31.9	31.2	24.2	32.0	8.4	58.2	56.4	46.2	58.2	14.7
	0.5	66.9	65.2	55.4	67.0	18.0	93.2	90.8	87.5	93.2	35.7
	0.7	87.1	85.2	79.3	87.2	29.4	99.3	97.6	98.3	99.3	56.7
	0.9	95.1	93.3	90.5	95.2	38.8	99.9	98.7	99.8	99.9	70.2
0.9	0.1	9.8	9.6	8.3	9.8	3.7	14.8	14.4	11.6	14.8	5.0
	0.3	49.4	48.2	39.3	49.5	12.6	80.7	78.2	70.7	80.7	24.1
	0.5	87.3	85.4	79.5	87.4	30.1	99.4	97.7	98.3	99.4	57.8
	0.7	97.8	96.3	95.5	97.8	47.1	100.0	99.2	99.9	100.0	79.7
	0.9	99.6	98.7	99.0	99.7	57.6	100.0	99.6	100.0	100.0	87.8

Panel B. $I(1)$

		$T=150$					$T=300$				
λ	θ	W_λ	W_P	W^*	W_0	W_1	W_λ	W_P	W^*	W_0	W_1
0.7	0.1	8.4	10.0	6.9	13.6	9.1	13.0	13.9	8.4	13.6	13.9
	0.3	42.3	41.0	19.9	15.8	45.2	75.3	74.9	34.3	15.8	76.7
	0.5	82.0	78.5	40.5	17.3	84.5	98.8	98.4	66.1	14.6	99.0
	0.7	96.2	94.2	58.9	18.6	97.1	100.0	99.9	81.6	14.9	100.0
	0.9	99.4	98.5	70.1	20.4	99.6	100.0	100.0	87.0	17.3	100.0
0.9	0.1	10.9	12.4	7.9	13.9	12.0	19.1	19.8	10.3	13.8	20.0
	0.3	63.8	60.8	30.0	16.9	67.0	93.0	92.5	51.2	15.5	93.8
	0.5	95.8	93.3	60.0	19.2	96.7	100.0	99.9	82.2	15.6	100.0
	0.7	99.7	99.2	78.6	27.4	99.8	100.0	100.0	90.0	25.2	100.0
	0.9	100.0	99.9	85.9	40.2	100.0	100.0	100.0	92.6	40.7	100.0

Table 3. Estimated powers of nominal 5%-level linearity tests: ESTAR.

Panel A. $I(0)$												
			$T=150$					$T=300$				
ϕ	δ	θ	W_λ	W_P	W^*	W_0	W_1	W_λ	W_P	W^*	W_0	W_1
0.0	0.7	0.1	19.8	19.4	15.3	20.0	1.7	38.5	37.0	29.1	38.6	1.5
		0.3	38.7	38.6	30.6	39.8	1.7	74.4	71.9	63.1	74.6	1.8
		0.5	28.6	29.3	22.7	30.1	2.0	62.5	60.8	49.4	63.0	2.0
		0.7	17.3	18.4	14.3	18.8	2.1	41.1	40.3	30.9	41.9	2.1
		0.9	10.6	11.6	9.6	11.8	2.2	25.5	25.3	18.5	26.3	2.3
	0.9	0.1	31.6	30.9	25.0	31.9	1.6	60.1	57.9	49.5	60.2	1.5
		0.3	60.0	62.0	49.0	63.6	2.1	95.0	92.3	88.7	95.3	2.1
		0.5	29.5	34.1	24.5	34.9	2.5	63.8	65.5	50.1	67.7	2.6
		0.7	12.8	16.2	11.8	16.6	2.6	26.1	29.0	19.6	30.1	2.5
		0.9	6.9	9.2	7.4	9.5	2.7	11.1	13.0	9.3	13.4	2.5
1.0	−0.7	0.1	57.0	61.5	50.2	63.2	5.5	89.6	87.7	82.4	90.7	5.9
		0.3	57.0	57.8	46.8	59.3	6.4	87.1	84.6	78.0	87.5	7.7
		0.5	44.9	45.1	35.4	46.3	6.1	74.1	72.1	62.6	74.5	7.1
		0.7	34.4	34.5	26.7	35.4	5.5	58.5	56.9	47.1	58.8	6.0
		0.9	26.4	26.6	20.4	27.3	5.0	45.4	44.1	35.1	45.6	5.3
	−0.9	0.1	71.3	73.2	63.6	75.2	7.0	96.2	93.4	92.2	96.6	8.0
		0.3	74.8	74.0	65.1	76.0	9.1	96.7	93.8	92.7	96.8	12.3
		0.5	63.8	62.9	53.1	64.6	8.2	90.9	88.1	83.6	91.0	10.3
		0.7	51.4	50.6	41.0	51.9	6.9	80.2	77.7	69.7	80.2	8.1
		0.9	40.8	40.1	31.6	41.2	5.8	67.2	65.0	55.4	67.2	6.5
1.5	−1.0	0.1	94.2	79.5	97.9	98.9	7.8	100.0	74.3	100.0	100.0	8.0
		0.3	91.1	92.2	87.5	95.4	6.7	99.9	96.6	99.5	99.9	7.8
		0.5	74.0	76.8	66.1	78.9	6.3	96.8	94.1	92.9	97.2	7.8
		0.7	56.1	58.6	46.9	60.1	5.7	86.1	83.9	76.9	86.9	6.8
		0.9	41.8	43.8	33.9	45.0	5.0	70.9	69.4	59.5	71.8	5.8
	−1.4	0.1	99.8	94.5	99.9	100.0	15.0	100.0	94.8	100.0	100.0	19.7
		0.3	99.4	97.3	98.4	99.5	19.8	100.0	97.4	100.0	100.0	32.6
		0.5	95.9	93.9	91.6	96.3	16.2	100.0	97.4	99.8	100.0	25.9
		0.7	87.1	85.3	79.0	87.6	12.3	99.1	96.3	97.7	99.1	18.3
		0.9	75.1	73.6	64.4	75.7	9.6	95.9	92.9	91.6	95.9	13.1

Table 3. Continued.

Panel B. $I(1)$												
			$T = 150$					$T = 300$				
ϕ	λ	θ	W_λ	W_P	W^*	W_0	W_1	W_λ	W_P	W^*	W_0	W_1
0.0	0.7	0.1	16.7	15.5	13.8	8.8	20.2	35.0	34.5	25.8	8.4	38.3
		0.3	33.8	28.9	26.8	10.0	40.7	70.0	66.4	55.4	8.8	75.0
		0.5	26.5	23.9	20.8	12.7	31.8	58.8	57.0	43.2	10.9	63.8
		0.7	17.4	16.6	14.2	14.0	20.6	39.1	38.8	27.5	12.1	42.9
		0.9	12.0	12.2	10.2	14.6	13.8	24.9	25.2	17.6	13.0	27.4
	0.9	0.1	26.8	23.1	21.9	8.7	32.4	55.8	53.7	43.4	8.4	60.3
		0.3	58.2	50.0	45.0	29.0	65.5	92.2	85.8	81.2	28.7	95.4
		0.5	38.8	36.2	25.6	40.4	39.2	67.4	65.1	47.0	37.9	70.0
		0.7	24.6	24.3	14.5	43.5	21.6	34.3	33.1	20.5	40.1	34.0
		0.9	18.6	18.4	10.8	44.4	14.2	19.1	18.2	11.0	40.5	17.6
1.0	−0.7	0.1	55.7	46.3	47.1	16.6	64.5	87.0	81.6	75.8	16.2	90.9
		0.3	50.0	40.4	42.8	9.3	59.7	82.9	77.6	71.0	9.1	87.5
		0.5	37.9	31.0	32.4	7.7	46.2	69.4	65.7	55.5	7.6	74.4
		0.7	28.7	24.3	24.6	7.0	35.7	53.8	51.9	41.4	6.9	58.6
		0.9	21.8	19.0	19.0	6.8	27.2	41.1	40.4	31.0	6.7	45.5
	−0.9	0.1	66.2	52.4	58.4	13.5	75.8	93.4	86.3	86.1	13.3	96.6
		0.3	65.8	50.6	58.5	7.3	76.3	93.2	85.3	86.2	7.0	96.7
		0.5	54.3	42.4	47.2	6.4	64.6	86.5	80.2	76.0	6.3	91.0
		0.7	42.9	34.5	36.4	6.6	52.0	75.2	70.6	61.8	6.6	80.2
		0.9	33.2	27.5	27.8	6.9	41.0	61.9	59.1	48.6	6.7	66.9
1.5	−1.0	0.1	99.4	94.1	90.8	44.8	100.0	99.9	98.6	96.2	49.1	100.0
		0.3	88.7	70.6	81.0	20.9	95.6	98.0	91.3	95.0	21.2	99.9
		0.5	69.8	55.3	60.3	14.9	79.3	93.7	87.1	86.0	14.6	97.2
		0.7	51.2	41.6	43.1	12.4	60.5	82.3	77.6	69.7	12.1	87.0
		0.9	37.3	31.5	31.6	11.1	45.4	66.8	64.0	52.8	11.1	71.7
	−1.4	0.1	97.4	81.7	95.2	28.6	100.0	99.3	94.2	97.2	29.6	100.0
		0.3	92.2	66.3	93.1	10.2	99.6	98.0	89.0	97.1	9.8	100.0
		0.5	86.6	62.7	84.0	7.0	96.2	97.4	88.3	95.8	7.3	99.9
		0.7	76.8	56.6	71.2	6.4	87.7	96.0	87.5	92.3	6.6	99.1
		0.9	64.8	49.4	57.0	6.4	75.7	92.1	84.5	84.5	6.4	95.9

Table 4. Estimated powers of nominal 5%-level linearity tests: SETAR, $\phi = 0$.

	$T = 150$					$T = 300$				
Panel A. $I(0)$										
δ	W_λ	W_P	W^*	W_0	W_1	W_λ	W_P	W^*	W_0	W_1
0.3	13.5	13.3	11.1	13.6	4.7	24.1	23.4	18.0	24.1	6.8
0.5	31.8	31.4	24.0	32.1	7.9	59.0	57.3	47.0	59.1	12.8
0.7	57.6	57.6	46.4	59.0	11.2	89.0	86.5	80.6	89.2	19.4
0.9	73.3	77.2	67.0	79.3	11.3	97.4	95.3	94.6	97.8	17.8
Panel B. $I(1)$										
	$T = 150$					$T = 300$				
λ	W_λ	W_P	W^*	W_0	W_1	W_λ	W_P	W^*	W_0	W_1
0.3	28.1	28.5	9.6	13.6	29.7	54.6	55.2	12.6	14.7	55.7
0.5	69.0	68.4	16.3	21.1	70.2	95.6	95.7	27.3	22.0	95.7
0.7	94.6	94.5	28.7	38.9	94.8	100.0	100.0	54.3	39.7	100.0
0.9	94.9	94.3	47.1	72.6	94.4	99.7	99.6	75.0	73.1	99.6

Table 5. Application to short-term interest rate series, 1978:1–2006:12.

Country	W_λ	$W^*_{10\%}$	$W^*_{5\%}$	$W^*_{1\%}$
Australia	20.27***	37.35*	37.70**	38.31***
Canada	22.09***	25.38*	25.56**	25.88***
New Zealand	54.03***	71.76*	72.45**	73.69***
Switzerland	1.59	5.03	5.06	5.11
UK	6.32**	7.31	7.38	7.50
US	30.20***	36.45*	36.75**	37.27***

Note: *, ** and *** denote rejection at the 10%-, 5%- and 1%-levels, respectively.