Granger Centre
for time series econometrics

Granger Centre Discussion Paper Series

Testing for nonlinear trends when the order of integration is unknown

by

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

Granger Centre Discussion Paper No. 09/04

The University of Nottingham

Testing for Nonlinear Trends when the Order of Integration is Unknown

David I. Harvey[†,*], Stephen J. Leybourne[†] and Lisa Xiao[‡]
[†]School of Economics, University of Nottingham
[‡]Royal London Asset Management Limited

August 2009

Abstract

We consider testing for the presence of nonlinearities in the mean and/or trend of a time series, approximating the potential nonlinear behaviour using a Fourier function expansion. In contrast to procedures that are currently available, we develop tests that are robust to the order of integration, in the sense that they are asymptotically correctly sized regardless of whether the stochastic component of the series is stationary or contains a unit root. The tests we propose take the form of Wald statistics based on cumulated series, together with a correction factor to line up the asymptotic critical values across the I(0) and I(1) environments. The local asymptotic power and finite sample properties of the tests are evaluated using various different correction factors. We envisage that the testing procedure we recommend should be very useful to applied researchers wishing to draw robust inference regarding the presence of nonlinear trend components in a series.

Keywords: Trend function testing; Robust tests; Fourier approximation.

JEL Classifications: C22.

1 Introduction

The issue of correctly identifying the underlying deterministic component in a financial or economic time series has received much attention in the econometrics literature, in recognition of the importance the problem assumes in applied modelling and forecasting. The effectiveness of both policy modelling and prediction is reliant on correct identification of the underlying trend function, and correctly specifying the trend function is also of crucial importance in the context of unit root and stationarity testing.

[*]Corresponding author: School of Economics, University of Nottingham, University Park, Nottingham, NG7 2RD, UK. Email: dave.harvey@nottingham.ac.uk.

Traditionally, consideration of the deterministic trend function has been restricted to evaluating whether the series simply contains either a constant or a constant plus linear trend, although more recent work has focused on the possibility of changes in level and/or trend, partly due to the critical role that modelling breaks can play in achieving unit root and stationarity tests with reliable size and power. A common feature of the vast majority of this literature is that structural changes are typically modelled as occurring instantaneously. However, it has recently been recognised that an assumption of instantaneous change might not adequately capture the true evolutionary pattern of the deterministic component present in the data. For example, changes in economic aggregates are affected by the responses of a large number of individuals, who would not be expected to react simultaneously to shocks. Consequently, models of the deterministic component that permit smoothly evolving (nonlinear) trends can often provide a better explanation of the underlying trend function.

One way of capturing nonlinear behaviour in the underlying trend is to approximate the trend by using frequency components of the Fourier function. In recent work, Becker *et al.* (2004) and Becker *et al.* (2006) show that a Fourier approximation is sufficiently flexible to provide a decent approximation to a wide variety of functions, even for those that are not periodic. Becker *et al.* (2004) provide a test for the presence of nonlinear trend components based on a Fourier approximation, under the assumption that the stochastic component of the series is stationary. Enders and Lee (2008) and Becker *et al.* (2006) also adopt a Fourier approximation for the trend, employing such an approach in the context of testing for, respectively, a unit root and stationarity in a time series allowing for possibly nonlinear deterministics.

Formal testing of whether a time series contains a nonlinear Fourier-approximated trend is greatly complicated by the fact that in practice it is not known whether the trend is embedded in a stationary I(0) or unit root I(1) series. If one knew that the shocks were stationary, then one could proceed with tests such as those of Becker *et al.* (2004); however, if the shocks in fact follow a unit root process, such tests will, as noted by Becker *et al.* (2006), be over-sized, and consequently might spuriously suggest the presence of nonlinear trend terms when in fact none exist. Reliance on tests that lack robustness to the order of integration can therefore lead to an over-specified model of the trend, subsequently reducing the efficacy of modelling and forecasting efforts. Equally, an unnecessarily elaborate trend specification incorporated into unit root or stationarity tests will inevitably result in procedures with much reduced power.

In view of these considerations, it is important to have available tests for nonlinear trend components of a series that are not reliant on an assumption of stationary shocks. Enders and Lee (2004) propose an alternative test for nonlinear trend terms under the assumption that the data contain a unit root, although this test is then not appropriate if the shocks are in fact stationary. Given that knowledge of the order of integration can never in practice be taken as known, it is highly desirable to develop nonlinear trend tests which are robust to the order of integration, in the sense that they are correctly sized regardless of whether the stochastic component of the series is truly I(0) or I(1). The aim of this paper is to provide a test procedure with precisely these properties, that can then be used in applications without recourse to making assumptions or drawing unit root/stationarity test inferences regarding the order of integration (the latter themselves being dependent on an assumption regarding the presence or otherwise of nonlinear trend terms).

The plan of the paper is as follows: in section 2, we introduce the Fourier function-based nonlinear trend model, and describe assumptions regarding the stochastic component, allowing either I(0) or I(1) behaviour in the model errors. Section 3 presents our robust testing methodology, which is a development of the seminal work of Vogelsang (1998), whereby Wald tests based on cumulated series are constructed, and a multiplicative correction is applied to enable asymptotically robust inference across I(0) and I(1) errors to be conducted. Here we also derive the local asymptotic power of the modified Wald tests to detect the presence of a nonlinear Fourier trend function, considering the behaviour of a number of different specifications for the multiplicative correction statistic. In section 4 we provide the results from a set of finite sample Monte Carlo simulation experiments, which together document the proposed tests' size and power behaviour in small samples. Some conclusions are offered in section 5, and proofs of the main technical results in this paper are given in an appendix. In what follows, $\xrightarrow{d}$ denotes weak convergence and $\lfloor \cdot \rfloor$ the integer part of an argument.

2 The Model

Consider the following model for the time series y_t, $t = 1, ..., T$

$$y_t = \mu_t + \lambda_{t,T} + u_t \tag{1}$$

where μ_t and $\lambda_{t,T}$ are, respectively, linear and nonlinear deterministic components of y_t, while u_t is stochastic. With regard to the precise specifications of μ_t and $\lambda_{t,T}$, we consider two separate cases according to whether or not y_t is assumed to contain a linear trend. In the "mean case" we let the linear component be simply a constant, i.e. $\mu_t = \alpha$, while the nonlinear component is specified using the Fourier series expansion

$$\lambda_{t,T} = \sum_{f=1}^{n} \gamma_{1f,T} \sin\left(\frac{2\pi f t}{T}\right) + \sum_{f=1}^{n} \gamma_{2f,T} \cos\left(\frac{2\pi f t}{T}\right) \tag{2}$$

where n dictates the number of frequencies contained in the expansion, and $f \in \mathbb{Z}^+$ denotes a particular frequency. In the "trend case" we admit a linear trend in μ_t, i.e. $\mu_t = \alpha + \beta t$ and also allow for smoothly evolving nonlinear trending components in $\lambda_{t,T}$

$$\begin{aligned}\lambda_{t,T} \quad = \quad & \sum_{f=1}^{n} \gamma_{1f,T} \sin\left(\frac{2\pi f t}{T}\right) + \sum_{f=1}^{n} \gamma_{2f,T} \cos\left(\frac{2\pi f t}{T}\right) \\ & + \sum_{f=1}^{n} \gamma_{3f,T} \sin\left(\frac{2\pi f t}{T}\right) t + \sum_{f=1}^{n} \gamma_{4f,T} \cos\left(\frac{2\pi f t}{T}\right) t. \end{aligned} \tag{3}$$

As regards the specification of the parameters $\gamma_{1f,T}, ..., \gamma_{4f,T}$, $f = 1, .., n$, and the stochastic component of (1), u_t, we will consider one or other of the following two assumptions:

Assumption I(0) *The stochastic process u_t is such that $u_t = v_t$ where $v_t = C(L)\varepsilon_t$, $C(L) = \sum_{j=0}^{\infty} C_j L^j$ with $C(1)^2 > 0$ and $\sum_{i=0}^{\infty} i|C_i| < \infty$, and where ε_t is an IID sequence with mean zero, variance σ_ε^2 and finite fourth moment. The long run variance of v_t is defined as $\omega_v^2 = \lim_{T\to\infty} T^{-1} E(\sum_{t=1}^{T} v_t)^2 = \sigma_\varepsilon^2 C(1)^2$. Also, for $f = 1, .., n$*

$$\gamma_{1f,T} = \gamma_{1f}\omega_v T^{-1/2}, \quad \gamma_{2f,T} = \gamma_{2f}\omega_v T^{-1/2}$$
$$\gamma_{3f,T} = \gamma_{3f}\omega_v T^{-3/2}, \quad \gamma_{4f,T} = \gamma_{4f}\omega_v T^{-3/2}.$$

Assumption I(1) *The stochastic process u_t is such that $\Delta u_t = v_t$ where v_t is as defined in Assumption I(0). Also, for $f = 1, .., n$*

$$\gamma_{1f,T} = \gamma_{1f}\omega_v T^{1/2}, \quad \gamma_{2f,T} = \gamma_{2f}\omega_v T^{1/2}$$
$$\gamma_{3f,T} = \gamma_{3f}\omega_v T^{-1/2}, \quad \gamma_{4f,T} = \gamma_{4f}\omega_v T^{-1/2}.$$

Remark 1 Note that under Assumption I(0), the series y_t is a stationary stochastic process about a potentially nonlinearly evolving mean/trend, while under Assumption I(1), y_t has the same deterministic features but also has as its stochastic component a unit root process. The various T-scalings on the Fourier coefficient terms are analytical devices that yield Pitman drifts appropriate to a local asymptotic power analysis of the nonlinear trend test statistics we consider in the next section. The scaling by ω_v is simply a convenience device to allow it to be factored out of the limit distributions that arise.

Remark 2 The conditions on v_t are such that it follows a conventional stationary, invertible linear process, and therefore admits the functional central limit theorem

$$T^{-1/2} \sum_{t=1}^{\lfloor Tr \rfloor} v_t \xrightarrow{d} \omega_v W(r)$$

where $W(r)$ is a standard Brownian motion process on $[0, 1]$.

In practice, the order of integration of the stochastic component of y_t cannot be taken as a given, even in the absence of nonlinear deterministics. Indeed, identification of the form of the deterministic component is often a precursor to conducting unit root and/or stationarity tests to determine the order of integration. Given this lack of knowledge concerning the integration properties of the data, in the next section we examine an approach that allows testing the null of a linear trend specification in y_t against a nonlinear deterministic component, which is asymptotically valid under both Assumption I(0) and Assumption I(1), thereby allowing robust testing for nonlinear trends.

3 Testing for Nonlinear Trends

Our focus in this paper is to devise a test for the presence of nonlinear deterministic components, abstracting from the issue of whether the series under consideration is stationary or contains a unit root. In the context of (1), we are concerned with testing the null (H_0) of a linear mean/trend against the alternative (H_1) of a nonlinear mean/trend. Specifically, in the mean case we wish to test

$$\begin{aligned} H_0^{\mu} &: \ \gamma_{1f} = \gamma_{2f} = 0, \ f = 1, ..., n \\ H_1^{\mu} &: \ \text{at least one of } \gamma_{1f}, \gamma_{2f} \neq 0, \ f = 1, ..., n \end{aligned}$$

and in the trend case

$$\begin{aligned} H_0^{\tau} &: \ \gamma_{1f} = \gamma_{2f} = \gamma_{3f} = \gamma_{4f} = 0, \ f = 1, ..., n \\ H_1^{\tau} &: \ \text{at least one of } \gamma_{1f}, \gamma_{2f}, \gamma_{3f}, \gamma_{4f} \neq 0, \ f = 1, ..., n. \end{aligned}$$

Standard (autocorrelation corrected) Wald-type tests of these hypotheses based on regression (1) will have the usual asymptotic χ^2 null distributions provided Assumption I(0) holds. However, it is straightforward to demonstrate that under Assumption I(1), these same tests will *diverge* to $+\infty$ under the null as the sample size increases. They therefore spuriously indicate the presence of nonlinear deterministics when the errors are I(1), thereby displaying a highly unsatisfactory lack of robustness.

To proceed in developing tests that are robust to stationary and unit root errors, we adapt the approach proposed by Vogelsang (1998) for trend function hypothesis testing in the presence of uncertainty regarding the order of integration. Consider first the partially summed counterpart to regression (1)

$$z_t = \sum_{s=1}^{t} \mu_s + \sum_{s=1}^{t} \lambda_{s,T} + \eta_t \tag{4}$$

where $z_t = \sum_{s=1}^{t} y_s$ and $\eta_t = \sum_{s=1}^{t} u_s$. In the mean case, (4) is given by

$$z_t = \alpha t + \sum_{f=1}^{n} \gamma_{1f,T} \sum_{s=1}^{t} \sin\left(\frac{2\pi f s}{T}\right) + \sum_{f=1}^{n} \gamma_{2f,T} \sum_{s=1}^{t} \cos\left(\frac{2\pi f s}{T}\right) + \eta_t \tag{5}$$

and in the trend case,

$$\begin{aligned} z_t \ = \ & \alpha t + \beta \sum_{s=1}^{t} s + \sum_{f=1}^{n} \gamma_{1f,T} \sum_{s=1}^{t} \sin\left(\frac{2\pi f s}{T}\right) + \sum_{f=1}^{n} \gamma_{2f,T} \sum_{s=1}^{t} \cos\left(\frac{2\pi f s}{T}\right) \\ & + \sum_{f=1}^{n} \gamma_{3f,T} \sum_{s=1}^{t} \sin\left(\frac{2\pi f s}{T}\right) s + \sum_{f=1}^{n} \gamma_{4f,T} \sum_{s=1}^{t} \cos\left(\frac{2\pi f s}{T}\right) s + \eta_t. \end{aligned} \tag{6}$$

Next, consider the standard Wald statistics for testing H_0^μ in the mean case and H_0^τ in the trend case (denoted W_T^μ and W_T^τ respectively) based on the partially summed regressions, that is

$$\begin{aligned} W_T^\mu &= \frac{RSS_R^\mu - RSS_U^\mu}{RSS_U^\mu/T} \\ W_T^\tau &= \frac{RSS_R^\tau - RSS_U^\tau}{RSS_U^\tau/T} \end{aligned}$$

where RSS_R^μ denotes the residual sum of squares from an OLS regression of z_t on t, RSS_R^τ denotes the residual sum of squares from an OLS regression of z_t on t and $\sum_{s=1}^{t} s$ (restricted regressions), and RSS_U^μ and RSS_U^τ denote the residual sums of squares from OLS estimation of (5) and (6), respectively (unrestricted regressions).

The advantage of considering partially summed regressions in this way is that the resulting Wald statistics have (when scaled by T) well-defined asymptotic distributions under both Assumption I(0) and Assumption I(1). These limit distributions are detailed in the following two theorems.

Theorem 1 *Let y_t be generated by (1) and let Assumption I(0) hold. Then:*
(i) In the mean case ($\mu_t = \alpha$),

$$T^{-1}W_T^\mu \xrightarrow{d} \frac{\int_0^1 L_R^\mu(r,\boldsymbol{\gamma}^\mu)^2 dr}{\int_0^1 L_U^\mu(r)^2 dr} - 1 \equiv D_0^\mu(\boldsymbol{\gamma}^\mu)$$

where, with $\boldsymbol{\gamma}^\mu = [\gamma_{11}, ..., \gamma_{1n}, \gamma_{21}, ..., \gamma_{2n}]$, $L_R^\mu(r,\boldsymbol{\gamma}^\mu)$ denotes the continuous time residuals from the projection of $\sum_{f=1}^{n} \gamma_{1f} m_{1f}(r) + \sum_{f=1}^{n} \gamma_{2f} m_{2f}(r) + W(r)$ onto the space spanned by r, and $L_U^\mu(r)$ denotes the continuous time residuals from the projection of $W(r)$ onto the space spanned by $\{r, m_{11}(r), ..., m_{1n}(r), m_{21}(r), ..., m_{2n}(r)\}$, where

$$\begin{aligned} m_{1f}(r) &= \frac{1}{2\pi f}(1-\cos(2\pi f r)) \\ m_{2f}(r) &= \frac{1}{2\pi f}\sin(2\pi f r). \end{aligned}$$

(ii) In the trend case ($\mu_t = \alpha + \beta t$),

$$T^{-1}W_T^\tau \xrightarrow{d} \frac{\int_0^1 L_R^\tau(r,\boldsymbol{\gamma}^\tau)^2 dr}{\int_0^1 L_U^\tau(r)^2 dr} - 1 \equiv D_0^\tau(\boldsymbol{\gamma}^\tau)$$

where, with $\boldsymbol{\gamma}^\tau = [\gamma_{11}, ..., \gamma_{1n}, \gamma_{21}, ..., \gamma_{2n}, \gamma_{31}, ..., \gamma_{3n}, \gamma_{41}, ..., \gamma_{4n}]$, $L_R^\tau(r,\boldsymbol{\gamma}^\tau)$ denotes the continuous time residuals from the projection of $\sum_{f=1}^{n} \gamma_{1f} m_{1f}(r) + \sum_{f=1}^{n} \gamma_{2f} m_{2f}(r) + \sum_{f=1}^{n} \gamma_{3f} m_{3f}(r) + \sum_{f=1}^{n} \gamma_{4f} m_{4f}(r) + W(r)$ onto the space spanned by $\{r, r^2\}$, and $L_U^\tau(r)$ denotes the continuous time residuals from the projection of $W(r)$ onto the space *spanned by* $\{r, r^2, m_{11}(r), ..., m_{1n}(r), m_{21}(r), ..., m_{2n}(r), m_{31}(r), ..., m_{3n}(r), m_{41}(r), ..., m_{4n}(r)\}$*, where* $m_{1f}(r)$ *and* $m_{2f}(r)$ *are as defined in (i), and*

$$\begin{aligned} m_{3f}(r) &= \frac{1}{(2\pi f)^2}\sin(2\pi fr) - \frac{r}{2\pi f}\cos(2\pi fr) \\ m_{4f}(r) &= \frac{1}{(2\pi f)^2}\cos(2\pi fr) + \frac{r}{2\pi f}\sin(2\pi fr) - \frac{1}{(2\pi f)^2}. \end{aligned}$$

Theorem 2 *Let* y_t *be generated by (1) and let Assumption I(1) hold. Then:*

(i) In the mean case ($\mu_t = \alpha$)*,*

$$T^{-1}W_T^{\mu} \xrightarrow{d} \frac{\int_0^1 N_R^{\mu}(r, \boldsymbol{\gamma}^{\mu})^2 dr}{\int_0^1 N_U^{\mu}(r)^2 dr} - 1 \equiv D_1^{\mu}(\boldsymbol{\gamma}^{\mu})$$

where $N_R^{\mu}(r, \boldsymbol{\gamma}^{\mu})$ *denotes the continuous time residuals from the projection of* $\sum_{f=1}^{n}\gamma_{1f}m_{1f}(r)+\sum_{f=1}^{n}\gamma_{2f}m_{2f}(r)+\int_0^r W(s)ds$ *onto the space spanned by* r*, and* $N_U^{\mu}(r)$ *denotes the continuous time residuals from the projection of* $\int_0^r W(s)ds$ *onto the space spanned by* $\{r, m_{11}(r), ..., m_{1n}(r), m_{21}(r), ..., m_{2n}(r)\}$*, with* $\boldsymbol{\gamma}^{\mu}$*,* $m_{1f}(r)$ *and* $m_{2f}(r)$ *as defined in Theorem 1.*

(ii) In the trend case ($\mu_t = \alpha + \beta t$)*,*

$$T^{-1}W_T^{\tau} \xrightarrow{d} \frac{\int_0^1 N_R^{\tau}(r, \boldsymbol{\gamma}^{\tau})^2 dr}{\int_0^1 N_U^{\tau}(r)^2 dr} - 1 \equiv D_1^{\tau}(\boldsymbol{\gamma}^{\tau})$$

where $N_R^{\tau}(r, \boldsymbol{\gamma}^{\tau})$ *denotes the continuous time residuals from the projection of* $\sum_{f=1}^{n}\gamma_{1f}m_{1f}(r) + \sum_{f=1}^{n}\gamma_{2f}m_{2f}(r) + \sum_{f=1}^{n}\gamma_{3f}m_{3f}(r) + \sum_{f=1}^{n}\gamma_{4f}m_{4f}(r) + \int_0^r W(s)ds$ *onto the space spanned by* $\{r, r^2\}$*, and* $N_U^{\tau}(r)$ *denotes the continuous time residuals from the projection of* $\int_0^r W(s)ds$ *onto the space spanned by* $\{r, r^2, m_{11}(r), ..., m_{1n}(r), m_{21}(r), ..., m_{2n}(r), m_{31}(r), ..., m_{3n}(r), m_{41}(r), ..., m_{4n}(r)\}$*, with* $\boldsymbol{\gamma}^{\tau}$*,* $m_{1f}(r)$*,* $m_{2f}(r)$*,* $m_{3f}(r)$ *and* $m_{4f}(r)$ *as defined in Theorem 1.*

Remark 3 To obtain the asymptotic distributions of the tests under the null hypotheses, i.e. H_0^{μ} and H_0^{τ}, we simply set $\boldsymbol{\gamma}^{\mu} = 0$ and $\boldsymbol{\gamma}^{\tau} = 0$, respectively, in Theorems 1 and 2.

Although the limit distributions of $T^{-1}W_T^{\mu}$ and $T^{-1}W_T^{\tau}$ depend on the order of integration of u_t, a modification can be applied to ensure that the critical values for a given test and significance level coincide under Assumption I(0) and Assumption I(1). Specifically, consider a statistic, denoted generically by J_T, that satisfies the following two conditions: (i) J_T converges in probability to zero when the series u_t is stationary, and (ii) J_T weakly converges to a non-degenerate pivotal limit distribution J when a unit root is present in u_t. Then, using the mean case as an example, given a suitable choice for J_T, one could consider the following Vogelsang (1998)-type modified statistic for testing the null H_0^μ

$$MW_T^\mu(J_T) = T^{-1} W_T^\mu \exp(-b_\xi J_T)$$

where b_ξ is a finite positive scaling constant whose role will now be made precise. Straightforward application of the continuous mapping theorem shows that, assuming $J_T \xrightarrow{p} 0$ under Assumption I(0), $MW_T^\mu(J_T) \xrightarrow{d} D_0^\mu(\boldsymbol{\gamma}^\mu)$, and that, assuming $J_T \xrightarrow{d} J$ under Assumption I(1), $MW_T^\mu(J_T) \xrightarrow{d} D_1^\mu(\boldsymbol{\gamma}^\mu)\exp(-b_\xi J)$. The value of b_ξ can then be chosen so that, for a given significance level ξ, $MW_T^\mu(J_T)$ has the same asymptotic critical values under both Assumptions I(0) and I(1), i.e. b_ξ is chosen such that

$$\Pr(D_0^\mu(\mathbf{0}) > cv_\xi^\mu) = \Pr(D_1^\mu(\mathbf{0})\exp(-b_\xi J) > cv_\xi^\mu) = \xi$$

where cv_ξ^μ denotes the ξ-level asymptotic critical value obtained from $D_0^\mu(\mathbf{0})$. The modified statistic $MW_T^\mu(J_T)$ compared with critical values from $D_0^\mu(\mathbf{0})$ would then provide a feasible test of the null H_0^μ that is applicable without requiring knowledge concerning the order of integration of u_t. In the trend case, similar arguments can be applied to give rise to a corresponding modified statistic

$$MW_T^\tau(J_T) = T^{-1} W_T^\tau \exp(-b_\xi J_T)$$

noting that b_ξ and J_T are used in a generic sense here, since different specifications for J_T, as well as different values for b_ξ, will clearly be appropriate in the mean and trend cases.

To be operational, the procedures require choices for the statistic J_T. A number of candidate statistics which possess the properties required of J_T in the simpler linear trend testing framework are considered by Harvey *et al.* (2006). One is the unit root statistic of Park and Choi (1988) and Park (1990) that is suggested by Vogelsang (1998), while other candidates are functions of unit root statistics such as those proposed by Dickey and Fuller (1979) and Breitung (2002). Since we naturally wish J_T to be invariant to the presence of any nonlinear deterministic components in y_t, as well as the linear component μ_t, we consider a number of functions of unit root statistics which are invariant to the parameters involved in both μ_t and $\lambda_{t,T}$.

Specifically, denoting the mean case by $i = \mu$ and the trend case by $i = \tau$, let $\hat{u}_t^i$ denote the OLS residuals from the regression (1) with $\mu_t = \alpha$ and $\lambda_{t,T}$ given by (2) when $i = \mu$, and $\mu_t = \alpha + \beta t$ and $\lambda_{t,T}$ given by (3) when $i = \tau$. We then consider J_T to be one of the following family of statistics

$$DF_T^{i,j} = \left|DF_T^i\right|^{-j}$$

for $j = 0.5, 1, 2$. Here, DF_T^i is the standard Dickey-Fuller t-ratio unit root statistic

$$DF_T^i = \hat{\rho}/\text{s.e.}(\hat{\rho})$$

where the parameter estimates are obtained from the OLS regression

$$\Delta \hat{u}_t^i = \rho \hat{u}_{t-1}^i + \sum_{p=1}^{k} \delta_p \Delta \hat{u}_{t-p}^i + e_t. \tag{7}$$

The lag truncation parameter k is assumed to satisfy the standard condition that as $T \to \infty$, $1/k + k^2/T \to 0$. As a point of comparison we also consider employing the unit root statistic of Breitung (2002) but based on the residuals $\hat{u}_t^i$. This is given by

$$B_T^i = s_u^{-2} T^{-3} \sum_{t=1}^{T} (\sum_{s=1}^{t} \hat{u}_s^i)^2$$

where $s_u^2 = T^{-1} \sum_{t=1}^{T} (\hat{u}_t^i)^2$.

Remark 4 Notice that we do not consider a generalised version of the Park and Choi (1988) and Park (1990) unit root statistic for J_T, since their approach uses a statistic for testing for the presence of additional polynomial trend terms in y_t, and these terms conflict with those in the Fourier expansion already admitted in the deterministic component of y_t.

In Theorem 3 below we provide the asymptotic distributions of the $MW_T^{\mu}(J_T)$ and $MW_T^{\tau}(J_T)$ statistics for each of the above specifications for J_T.

Theorem 3 *Let y_t be generated by (1). Then:*

(a) Under Assumption I(0):

(i) In the mean case ($\mu_t = \alpha$),

$$MW_T^{\mu}(DF_T^{\mu,j}), MW_T^{\mu}(B_T^{\mu}) \xrightarrow{d} D_0^{\mu}(\boldsymbol{\gamma}^{\mu})$$

for $j = 0.5, 1, 2$, where $D_0^{\mu}(\boldsymbol{\gamma}^{\mu})$ is as defined in Theorem 1.

(ii) In the trend case ($\mu_t = \alpha + \beta t$),

$$MW_T^{\tau}(DF_T^{\tau,j}), MW_T^{\tau}(B_T^{\tau}) \xrightarrow{d} D_0^{\tau}(\boldsymbol{\gamma}^{\tau})$$

for $j = 0.5, 1, 2$, where $D_0^{\tau}(\boldsymbol{\gamma}^{\tau})$ is as defined in Theorem 1.

(b) Under Assumption I(1):

(i) In the mean case ($\mu_t = \alpha$),

$$\begin{aligned} MW_T^{\mu}(DF_T^{\mu,j}) &\xrightarrow{d} D_1^{\mu}(\boldsymbol{\gamma}^{\mu}) \exp(-b_{\xi} \left| DF^{\mu} \right|^{-j}) \\ MW_T^{\mu}(B_T^{\mu}) &\xrightarrow{d} D_1^{\mu}(\boldsymbol{\gamma}^{\mu}) \exp(-b_{\xi} B^{\mu}) \end{aligned}$$

for $j = 0.5, 1, 2$, where $D_1^{\mu}(\boldsymbol{\gamma}^{\mu})$ is as defined in Theorem 2 and

$$DF^{\mu} = \frac{K^{\mu}(1)^2 - K^{\mu}(0)^2 - 1}{2\sqrt{\int_0^1 K^{\mu}(r)^2 dr}}$$
$$B^{\mu} = \frac{\int_0^1 \left(\int_0^r K^{\mu}(s)ds\right)^2 dr}{\int_0^1 K^{\mu}(r)^2 dr}$$

with $K^{\mu}(r)$ the continuous time residuals from the projection of $W(r)$ onto the space spanned by $\{1, \sin(2\pi r), ..., \sin(2\pi n r), \cos(2\pi r), ..., \cos(2\pi n r)\}$.

(ii) In the trend case ($\mu_t = \alpha + \beta t$),

$$\begin{aligned} MW_T^{\tau}(DF_T^{\tau,j}) &\xrightarrow{d} D_1^{\tau}(\boldsymbol{\gamma}^{\tau}) \exp(-b_{\xi} |DF^{\tau}|^{-j}) \\ MW_T^{\tau}(B_T^{\tau}) &\xrightarrow{d} D_1^{\tau}(\boldsymbol{\gamma}^{\tau}) \exp(-b_{\xi} B^{\tau}) \end{aligned}$$

for $j = 0.5, 1, 2$, where $D_1^{\tau}(\boldsymbol{\gamma}^{\tau})$ is as defined in Theorem 2 and

$$DF^{\tau} = \frac{K^{\tau}(1)^2 - K^{\tau}(0)^2 - 1}{2\sqrt{\int_0^1 K^{\tau}(r)^2 dr}}$$
$$B^{\tau} = \frac{\int_0^1 \left(\int_0^r K^{\tau}(s)ds\right)^2 dr}{\int_0^1 K^{\tau}(r)^2 dr}$$

with $K^{\tau}(r)$ the continuous time residuals from the projection of $W(r)$ onto the space spanned by $\{1, r, \sin(2\pi r), ..., \sin(2\pi n r), \cos(2\pi r), ..., \cos(2\pi n r), r\sin(2\pi r), ..., r\sin(2\pi n r), r\cos(2\pi r), ..., r\cos(2\pi n r)\}$.

Remark 5 As before, to obtain the asymptotic distributions of the modified tests under the null hypotheses, i.e. H_0^{μ} and H_0^{τ}, we set $\boldsymbol{\gamma}^{\mu} = 0$ and $\boldsymbol{\gamma}^{\tau} = 0$, respectively, in Theorem 3.

To proceed further with our asymptotic analysis, we need to decide on suitable values that might be considered for the maximum number of frequencies in the Fourier series expansion, i.e. n in (2) and (3). Becker *et al.* (2006) argue that using $n = 1$ or $n = 2$ should prove sufficient to capture the important features of the underlying trend in most circumstances, thus we proceed by considering these two settings. Of course, it is entirely straightforward to extend our analysis to larger values of n, should the need arise.

3.1 Asymptotic critical values and b_{ξ} values

In Table 1 we present the asymptotic critical values and associated b_{ξ} values for the $MW_T^{\mu}(DF_T^{\mu,j})$ and $MW_T^{\tau}(DF_T^{\tau,j})$ tests, $j = 0.5, 1, 2$, and also the $MW_T^{\mu}(B_T^{\mu})$ and

$MW_T^\tau(B_T^\tau)$ tests, for both $n = 1$ and $n = 2$, at the $\xi = 0.10$, $\xi = 0.05$ and $\xi = 0.01$ significance levels. Notice that for a given correction statistic, the values of b_ξ are dependent on the choice of n as well as the significance level. The critical values were obtained from direct simulation of $D_0^\mu(\mathbf{0})$ and $D_0^\tau(\mathbf{0})$ in Theorem 1, for the mean and trend cases respectively, approximating the Brownian motion process using $NIID(0,1)$ random variates, and with the integrals approximated by normalized sums of 1000 steps. Here and throughout the paper, simulations were based on 50,000 replications using Gauss 9.0. The b_ξ values were obtained as follows. The limit expressions in Theorem 3(b) were simulated with the settings $\boldsymbol{\gamma}^\mu = \mathbf{0}$ and $\boldsymbol{\gamma}^\tau = \mathbf{0}$ for a grid of possible b_ξ in steps of 0.001. The values of b_ξ reported in Table 1 are then those which give the desired sizes to an accuracy of 3 decimal places. Use of these values yields tests that are asymptotically correctly sized under both I(0) and I(1) errors.

3.2 Asymptotic power

We now consider the asymptotic local power performance of the nonlinear trend tests under both Assumption I(0) and Assumption I(1). Asymptotic power can be simulated using the results of Theorem 3. Due to the multi-parameter nature of the alternative hypothesis, we employ the simplifying expedient of setting $\gamma_{11} = ... = \gamma_{1n} = \gamma_{21} = ... = \gamma_{2n} = \gamma$ for the mean case, and $\gamma_{11} = ... = \gamma_{1n} = \gamma_{21} = ... = \gamma_{2n} = \gamma_{31} = ... = \gamma_{3n} = \gamma_{41} = ... = \gamma_{4n} = \gamma$ for the trend case. Figure 1 reports the asymptotic local power (at the nominal 0.05 significance level) of all the tests as functions of γ, for the mean case. Four sub-figures are reported, corresponding to (a) I(0) errors with $n = 1$, (b) I(1) errors with $n = 1$, (c) I(0) errors with $n = 2$, and (d) I(1) errors with $n = 2$. Figure 2 presents the equivalent power results for the trend case.

In Figure 1(a) only one curve is visible, since all the tests have the same asymptotic distribution under Assumption I(0). Here we observe the reassuring feature that power is monotonically increasing in γ, that is, with increasing departure from the null hypothesis. In contrast, Figure 1(b) shows that while the asymptotic powers of the tests under Assumption I(1) still increase monotonically in γ, they do so in quite different ways. The three tests that make use of the DF_T^μ correction statistic all substantially outperform $MW_T^\mu(B_T^\mu)$ in terms of power, across all values of γ. Among the three DF_T^μ-based tests, the power ranking is $MW_T^\mu(DF_T^{\mu,0.5})$, followed by $MW_T^\mu(DF_T^{\mu,1})$, followed by $MW_T^\mu(DF_T^{\mu,2})$, although it is noteworthy that the difference between $MW_T^\mu(DF_T^{\mu,0.5})$ and $MW_T^\mu(DF_T^{\mu,1})$ is rather less marked than that between $MW_T^\mu(DF_T^{\mu,1})$ and $MW_T^\mu(DF_T^{\mu,2})$. For the case $n = 2$ in Figures 1(c)-1(d), almost identical comments apply *vis-à-vis* the relative power rankings of the tests; if anything, the superiority of the DF_T^μ-based tests is even further highlighted.

In the trend case (Figure 2), the curves again demonstrate monotonic power increases for all the tests in γ. Interestingly, in comparison to the mean case results in Figure 1, the test powers rise more rapidly in γ in the trend case under Assumption I(1), but less quickly under Assumption I(0), for both $n = 1$ and $n = 2$. Under I(1) errors, the same test rankings are observed as in the mean case, with the only noticeable difference being that $MW_T^\mu(B_T^\mu)$ loses less relative to the DF_T^μ-based tests than was seen in Figure 1.

Remark 6 Although our asymptotic analysis has focused on local power with the relevant T-scalings on the Fourier coefficient terms as outlined under Assumption I(0) and Assumption I(1), it should be clear that the tests will be consistent against fixed alternatives in all cases where the Pitman drifts require scaling by T^{-d}, $d > 0$, i.e. in the mean and trend cases under Assumption I(0) and in the trend case under Assumption I(1) (provided $\gamma_{3f,T}$ and/or $\gamma_{4f,T}$ are non-zero). In the remaining case, i.e. the mean case under Assumption I(1), the tests will not be consistent under a fixed alternative, since to obtain non-negligible local asymptotic power, the nonlinear trend coefficients must be scaled *up* by $T^{1/2}$. This feature is in keeping with the well-known result in the setting of testing for instantaneous level breaks in time series, that an unscaled level break in an I(1) process is not asymptotically detectable.

On the basis of the findings of our asymptotic analysis, the obvious recommendation would be to use the $MW_T^\mu(DF_T^{\mu,0.5})$ test, since it outperforms the alternative procedures considered here in terms of power. Whether such a clear preference is also manifest in finite samples is an issue we investigate in the next section.

4 Finite sample simulations

4.1 Size

To assess the finite sample size of the tests, we use the following DGP which allows for ARMA(1,1) and ARIMA(0,1,1) errors

$$\begin{aligned} y_t &= u_t, \quad t = 1, ..., T \\ u_t &= \phi u_{t-1} + \varepsilon_t - \theta \varepsilon_{t-1}, \quad t = 2, ..., T \end{aligned}$$

with $u_1 = \varepsilon_1$ and $\varepsilon_t \sim NIID(0, 1)$. Table 2 reports nominal 0.05-level size results for $T = \{150, 300, 600\}$ and the autocorrelation parameter settings $\phi = \{0.00, 0.50, 0.70, 0.90, 0.95, 1.00\}$, $\theta = \{-0.5, 0.0, 0.5\}$, for the mean case statistics with $n = 1$. In the estimation of regression (7), k was chosen using the MAIC procedure of Ng and Perron (2001) with maximum lag length set at $k_{\max} = \lfloor 12(T/100)^{1/4} \rfloor$, using the modification suggested by Perron and Qu (2007).

The finite sample size of all the tests is very reasonable when $\phi = 1.00$, even for $T = 150$, with the exception of $MW_T^\mu(B_T^\mu)$ when $\theta = 0.5$, where a moderate degree of over-size is observed in the smaller samples. In the stationary cases, the tests are seen to be under-sized to varying degrees, particularly $MW_T^\mu(DF_T^{\mu,0.5})$ and $MW_T^\mu(DF_T^{\mu,1})$ which remain markedly under-sized even for $T = 600$. This pattern of under-size arises since the rate of divergence of $|DF_T^\mu|^{0.5}$ and $|DF_T^\mu|^1$ in T to $+\infty$ is slower than for $|DF_T^\mu|^2$; consequently, the speed of convergence of $DF_T^{\mu,0.5}$ and $DF_T^{\mu,1}$ to zero, and hence the speed of convergence of $\exp(-b_\xi DF_T^{\mu,0.5})$ and $\exp(-b_\xi DF_T^{\mu,1})$ to one, is also slower. As a result, in small and moderately sized samples, the statistics are still contaminated by the random variables embedded in the correction terms, preventing the nominal size from being achieved. Qualitatively similar results were obtained for $n = 2$, and also in the trend case for $n = 1$ and $n = 2$. In view of this observed finite sample under-sizing in the stationary case, it is important to now assess how this impacts on the finite sample power of the tests, since it is no longer clear that our asymptotic preference for $MW_T^\mu(DF_T^{\mu,0.5})$ will carry over to samples of size typically encountered in practice.

4.2 Power

We now turn our attention to the finite sample power properties of the tests, again for the representative mean case with $n = 1$. Specifically we simulate from the following DGP

$$\begin{aligned} y_t &= \gamma \sin\left(\frac{2\pi t}{T}\right) + \gamma \cos\left(\frac{2\pi t}{T}\right) + u_t, \quad t = 1, ..., T \\ u_t &= \phi u_{t-1} + \varepsilon_t, \quad t = 2, ..., T \end{aligned}$$

with $u_1 = \varepsilon_1$ and $\varepsilon_t \sim NIID(0, 1)$. Figure 3 reports the results for nominal 0.05-level tests using $T = 150$, $\phi = \{0.00, 0.50, 0.70, 0.90, 0.95, 1.00\}$ and a grid of γ values from 0 to $\gamma_{\max}$ in 50 steps, where the choices $\gamma_{\max} = \{2, 4, 8, 20, 20, 100\}$ correspond to the six values of ϕ considered. We associate larger values of $\gamma_{\max}$ with larger values of ϕ in this way so that meaningful power curves are obtained.

The first observation that can be made is that in the I(0) cases, i.e. Figure 3(a)-3(e), the tests display quite different power curves to each other in finite samples, despite being asymptotically equivalent in these stationary AR(1) environments. In the pure noise error case of Figure 3(a), we see that $MW_T^\mu(B_T^\mu)$ always provides the most powerful test while $MW_T^\mu(DF_T^{\mu,0.5})$ generally has the least power. Although this would not be predicted from our asymptotic analysis, this arises because $MW_T^\mu(B_T^\mu)$ has empirical size closest to the nominal 0.05 level, while $MW_T^\mu(DF_T^{\mu,0.5})$ has size essentially zero. The power curves of $MW_T^\mu(DF_T^{\mu,1})$ and $MW_T^\mu(DF_T^{\mu,2})$ generally lie inbetween those of $MW_T^\mu(B_T^\mu)$ and $MW_T^\mu(DF_T^{\mu,0.5})$, in keeping with their respective size behaviour. At the other extreme, where the errors are a pure random walk, i.e. Figure 3(f), the power rankings are the reverse of those seen in Figure 3(a), that is $MW_T^\mu(DF_T^{\mu,0.5})$ is the most powerful test, followed by $MW_T^\mu(DF_T^{\mu,1})$, then $MW_T^\mu(DF_T^{\mu,2})$ and finally $MW_T^\mu(B_T^\mu)$, exactly in line with our asymptotic results for the I(1) case. Of course, this is unsurprising since the tests are all approximately correctly sized in this situation, even for $T = 150$.

For $0 < \phi < 1$, Figures 3(b)-3(d) show that as ϕ increases, the tests' power rankings migrate from the noise case to the random walk case in a fairly predictable way. Notice that the DF_T^μ-based tests generally outperform the test based on B_T^μ for $\phi = 0.9$ and above. Given that this region in ϕ might be considered of particular empirical relevance, especially in a world where uncertainty exists regarding the order of integration, these results would indicate a preference for the DF_T^{μ}-based tests, in line with the asymptotic findings. Among these procedures, the low finite sample power of $MW_T^{\mu}(DF_T^{\mu,0.5})$ in more stationary cases renders it somewhat unattractive, and, taking our asymptotic power and finite sample size and power analyses together, our recommendation would be for the less sensitive variant $MW_T^{\mu}(DF_T^{\mu,1})$. Unreported simulations for the case where $n = 2$, and also the trend case for $n = 1$ and $n = 2$, confirmed that the same rankings essentially arise, and thus our general recommendation would be for $MW_T^{\mu}(DF_T^{\mu,1})$ and $MW_T^{\tau}(DF_T^{\tau,1})$ in the mean and trend cases respectively.

5 Conclusion

In this paper we have provided tests for the presence of nonlinear deterministic components of a time series, as approximated by a Fourier expansion. We focus on procedures that are valid when the order of the stochastic component of the series is unknown, that is, it is unclear whether the underlying shocks are I(0) or I(1). The tests we propose are of the Wald type, and are based on cumulated series, together with a correction factor to line up the asymptotic critical values across the I(0) and I(1) environments, in the spirit of Vogelsang (1998). The behaviour of various different correction factors were investigated in terms of their asymptotic local power and also finite sample size and power behaviour. The particular correction factor that we recommend makes use of a modified Dickey-Fuller unit root statistic, which allows for the potentially nonlinear trend components. The resulting test displays good asymptotic and finite sample performance, and should prove to be a useful tool for applied researchers, offering a robust approach to testing for nonlinearities in the deterministic specification of an autocorrelated series in the presence of uncertainty regarding its order of integration.

Appendix

Since the Wald statistics are invariant to μ_t, we may set $\mu_t = 0$ in (1) without loss of generality in what follows.

Proof of Theorem 1

(i) First write

$$\begin{aligned} T^{-1}W_T^{\mu} &= \frac{T^{-2}RSS_R^{\mu}}{T^{-2}RSS_U^{\mu}} - 1 \\ &= \frac{T^{-1}\sum_{t=1}^{T}(T^{-1/2}e_{Rt})^2}{T^{-1}\sum_{t=1}^{T}(T^{-1/2}e_{Ut})^2} - 1 \end{aligned}$$

where, respectively, e_{Rt} and e_{Ut} denote OLS residuals from a regression of z_t on t and from the regression (5).

Now, for $t = \lfloor Tr \rfloor$, the limit process for $T^{-1/2}e_{Rt}$ is the same as the limit of the residual process when $T^{-1/2}z_{\lfloor Tr \rfloor}$ is regressed on r. We find that

$$
\begin{aligned}
T^{-1/2}z_{\lfloor Tr \rfloor} &= T^{-1/2}\sum_{f=1}^{n}\gamma_{1f,T}\sum_{s=1}^{\lfloor Tr \rfloor}\sin\left(\frac{2\pi fs}{T}\right) + T^{-1/2}\sum_{f=1}^{n}\gamma_{2f,T}\sum_{s=1}^{\lfloor Tr \rfloor}\cos\left(\frac{2\pi fs}{T}\right) + T^{-1/2}\eta_{\lfloor Tr \rfloor} \\
&= \omega_v\sum_{f=1}^{n}\gamma_{1f}T^{-1}\sum_{s=1}^{\lfloor Tr \rfloor}\sin\left(\frac{2\pi fs}{T}\right) + \omega_v\sum_{f=1}^{n}\gamma_{2f}T^{-1}\sum_{s=1}^{\lfloor Tr \rfloor}\cos\left(\frac{2\pi fs}{T}\right) + T^{-1/2}\eta_{\lfloor Tr \rfloor} \\
&\xrightarrow{d} \omega_v\sum_{f=1}^{n}\gamma_{1f}\int_0^r \sin(2\pi fs)\,ds + \omega_v\sum_{f=1}^{n}\gamma_{2f}T^{-1}\int_0^r \cos(2\pi fs)\,ds + \omega_v W(r) \\
&= \omega_v\sum_{f=1}^{n}\gamma_{1f}m_{1f}(r) + \omega_v\sum_{f=1}^{n}\gamma_{2f}m_{2f}(r) + \omega_v W(r) \qquad (8)
\end{aligned}
$$

where the last line follows on noting that

$$
\begin{aligned}
\int_0^r \sin(2\pi fs)\,ds &= \frac{1}{2\pi f}(1-\cos(2\pi fr)) \\
\int_0^r \cos(2\pi fs)\,ds &= \frac{1}{2\pi f}\sin(2\pi fr)
\end{aligned}
$$

A continuous time regression of (8) on r yields the limit residual process $\omega_v L_R^{\mu}(r, \boldsymbol{\gamma}^{\mu})$. Entirely similar reasoning establishes that the limit process for $T^{-1/2}e_{Ut}$ is the continuous time residual process $\omega_v L_U^{\mu}(r, \boldsymbol{\gamma}^{\mu})$. An application of the continuous mapping theorem (CMT) then verifies that

$$
\begin{aligned}
\frac{T^{-1}\sum_{t=1}^{T}(T^{-1/2}e_{Rt})^2}{T^{-1}\sum_{t=1}^{T}(T^{-1/2}e_{Ut})^2} - 1 &\xrightarrow{d} \frac{\omega_v^2\int_0^1 L_R^{\mu}(r,\boldsymbol{\gamma}^{\mu})^2dr}{\omega_v^2\int_0^1 L_U^{\mu}(r)^2dr} - 1 \\
&= \frac{\int_0^1 L_R^{\mu}(r,\boldsymbol{\gamma}^{\mu})^2dr}{\int_0^1 L_U^{\mu}(r)^2dr} - 1
\end{aligned}
$$

(ii) Again write

$$
T^{-1}W_T^{\tau} = \frac{T^{-1}\sum_{t=1}^{T}(T^{-1/2}e_{Rt})^2}{T^{-1}\sum_{t=1}^{T}(T^{-1/2}e_{Ut})^2} - 1
$$

where now, respectively, e_{Rt} and e_{Ut} denote OLS residuals from a regression of z_t on t and $\sum_{s=1}^{t} s$ and from the regression (6).

Now, for $t = \lfloor Tr \rfloor$, the limit process for $T^{-1/2}e_{Rt}$ is the same as the limit of the residual process when $T^{-1/2}z_{\lfloor Tr \rfloor}$ is regressed on r and $\int_0^r sds = r^2/2$ or, equally, when

$T^{-1/2}z_{\lfloor Tr\rfloor}$ is regressed on r and r^2. Here we find that

$$\begin{aligned}
T^{-1/2}z_{\lfloor Tr\rfloor} &= T^{-1/2}\sum_{f=1}^{n}\gamma_{1f,T}\sum_{s=1}^{\lfloor Tr\rfloor}\sin\left(\frac{2\pi fs}{T}\right)+T^{-1/2}\sum_{f=1}^{n}\gamma_{2f,T}\sum_{s=1}^{\lfloor Tr\rfloor}\cos\left(\frac{2\pi fs}{T}\right)\\
&\quad +T^{-1/2}\sum_{f=1}^{n}\gamma_{3f,T}\sum_{s=1}^{\lfloor Tr\rfloor}\sin\left(\frac{2\pi fs}{T}\right)s+T^{-1/2}\sum_{f=1}^{n}\gamma_{4f,T}\sum_{s=1}^{\lfloor Tr\rfloor}\cos\left(\frac{2\pi fs}{T}\right)s\\
&\quad +T^{-1/2}\eta_{\lfloor Tr\rfloor}\\
&= \omega_v\sum_{f=1}^{n}\gamma_{1f}T^{-1}\sum_{s=1}^{\lfloor Tr\rfloor}\sin\left(\frac{2\pi fs}{T}\right)+\omega_v\sum_{f=1}^{n}\gamma_{2f}T^{-1}\sum_{s=1}^{\lfloor Tr\rfloor}\cos\left(\frac{2\pi fs}{T}\right)\\
&\quad +\omega_v\sum_{f=1}^{n}\gamma_{3f}T^{-2}\sum_{s=1}^{\lfloor Tr\rfloor}\sin\left(\frac{2\pi fs}{T}\right)s+\omega_v\sum_{f=1}^{n}\gamma_{4f}T^{-2}\sum_{s=1}^{\lfloor Tr\rfloor}\cos\left(\frac{2\pi fs}{T}\right)s\\
&\quad +T^{-1/2}\eta_{\lfloor Tr\rfloor}\\
&\xrightarrow{d} \omega_v\sum_{f=1}^{n}\gamma_{1f}m_{1f}(r)+\omega_v\sum_{f=1}^{n}\gamma_{2f}m_{2f}(r)\\
&\quad +\omega_v\sum_{f=1}^{n}\gamma_{3f}\int_0^r\sin(2\pi fs)\,s\,ds+\omega_v\sum_{f=1}^{n}\gamma_{4f}\int_0^r\cos(2\pi fs)\,s\,ds\\
&\quad +\omega_v W(r)
\end{aligned}$$

Then, since

$$\begin{aligned}
\int_0^r\sin(2\pi fs)\,s\,ds &= \frac{1}{(2\pi f)^2}\sin(2\pi fr)-\frac{r}{2\pi f}\cos(2\pi fr)\\
\int_0^r\cos(2\pi fs)\,s\,ds &= \frac{1}{(2\pi f)^2}\cos(2\pi fr)+\frac{r}{2\pi f}\sin(2\pi fr)-\frac{1}{(2\pi f)^2}
\end{aligned}$$

we find that

$$\begin{aligned}
T^{-1/2}z_{\lfloor Tr\rfloor} &\xrightarrow{d} \omega_v\sum_{f=1}^{n}\gamma_{1f}m_{1f}(r)+\omega_v\sum_{f=1}^{n}\gamma_{2f}m_{2f}(r)\\
&\quad +\omega_v\sum_{f=1}^{n}\gamma_{3f}m_{3f}(r)+\omega_v\sum_{f=1}^{n}\gamma_{4f}m_{4f}(r)\\
&\quad +\omega_v W(r) \qquad (9)
\end{aligned}$$

A continuous time regression of (9) on r and r^2 yields the limit residual process $\omega_v L_R^\tau(r,\boldsymbol{\gamma}^\tau)$. As before, similar reasoning establishes that the limit process for $T^{-1/2}e_{Ut}$ is the continuous time residual process $\omega_v L_U^\tau(r,\boldsymbol{\gamma}^\tau)$ and it follows from the CMT that

$$\frac{T^{-1}\sum_{t=1}^{T}(T^{-1/2}e_{Rt})^2}{T^{-1}\sum_{t=1}^{T}(T^{-1/2}e_{Ut})^2}-1\xrightarrow{d}\frac{\int_0^1 L_R^\mu(r,\boldsymbol{\gamma}^\mu)^2dr}{\int_0^1 L_U^\mu(r)^2dr}-1$$

Proof of Theorem 2

(i) Now write

$$
\begin{aligned}
T^{-1}W_T^{\mu} &= \frac{T^{-4}RSS_R^{\mu}}{T^{-4}RSS_U^{\mu}} - 1 \\
&= \frac{T^{-1}\sum_{t=1}^{T}(T^{-3/2}e_{Rt})^2}{T^{-1}\sum_{t=1}^{T}(T^{-3/2}e_{Ut})^2} - 1
\end{aligned}
$$

where, respectively, e_{Rt} and e_{Ut} denote OLS residuals from a regression of z_t on t and from the regression (5).

Now, for $t = \lfloor Tr \rfloor$, the limit process for $T^{-1/2}e_{Rt}$ is the same as the limit of the residual process when $T^{-3/2}z_{\lfloor Tr \rfloor}$ is regressed on r. We find

$$
\begin{aligned}
T^{-3/2}z_{\lfloor Tr \rfloor} &= T^{-3/2}\sum_{f=1}^{n}\gamma_{1f,T}\sum_{s=1}^{\lfloor Tr \rfloor}\sin\left(\frac{2\pi fs}{T}\right) + T^{-3/2}\sum_{f=1}^{n}\gamma_{2f,T}\sum_{s=1}^{\lfloor Tr \rfloor}\cos\left(\frac{2\pi fs}{T}\right) + T^{-3/2}\eta_{\lfloor Tr \rfloor} \\
&= \omega_v\sum_{f=1}^{n}\gamma_{1f}T^{-1}\sum_{s=1}^{\lfloor Tr \rfloor}\sin\left(\frac{2\pi fs}{T}\right) + \omega_v\sum_{f=1}^{n}\gamma_{2f}T^{-1}\sum_{s=1}^{\lfloor Tr \rfloor}\cos\left(\frac{2\pi fs}{T}\right) + T^{-3/2}\eta_{\lfloor Tr \rfloor} \\
&\xrightarrow{d} \omega_v\sum_{f=1}^{n}\gamma_{1f}m_{1f}(r) + \omega_v\sum_{f=1}^{n}\gamma_{2f}m_{2f}(r) + \omega_v\int_0^r W(s)ds \qquad (10)
\end{aligned}
$$

A continuous time regression of (10) on r yields the limit residual process $\omega_v N_R^{\mu}(r, \boldsymbol{\gamma}^{\mu})$. Similar reasoning establishes that the limit process for $T^{-1/2}e_{Ut}$ is the continuous time residual process $\omega_v N_U^{\mu}(r, \boldsymbol{\gamma}^{\mu})$ and an application of the CMT shows that

$$
\frac{T^{-1}\sum_{t=1}^{T}(T^{-3/2}e_{Rt})^2}{T^{-1}\sum_{t=1}^{T}(T^{-3/2}e_{Ut})^2} - 1 \xrightarrow{d} \frac{\int_0^1 N_R^{\mu}(r, \boldsymbol{\gamma}^{\mu})^2 dr}{\int_0^1 N_U^{\mu}(r)^2 dr} - 1
$$

(ii) Again write

$$
T^{-1}W_T^{\tau} = \frac{T^{-1}\sum_{t=1}^{T}(T^{-3/2}e_{Rt})^2}{T^{-1}\sum_{t=1}^{T}(T^{-3/2}e_{Ut})^2} - 1
$$

where now, respectively, e_{Rt} and e_{Ut} denote OLS residuals from a regression of z_t on t and $\sum_{s=1}^{t} s$ and from the regression (6).

Now, for $t = \lfloor Tr \rfloor$, the limit process for $T^{-3/2}e_{Rt}$ is the same as the limit of the residual process when $T^{-3/2}z_{\lfloor Tr\rfloor}$ is regressed on r and r^2. Now,

$$\begin{aligned}
T^{-1/2}z_{\lfloor Tr\rfloor} &= T^{-3/2}\sum_{f=1}^{n}\gamma_{1f,T}\sum_{s=1}^{\lfloor Tr\rfloor}\sin\left(\frac{2\pi fs}{T}\right)+T^{-3/2}\sum_{f=1}^{n}\gamma_{2f,T}\sum_{s=1}^{\lfloor Tr\rfloor}\cos\left(\frac{2\pi fs}{T}\right)\\
&\quad +T^{-3/2}\sum_{f=1}^{n}\gamma_{3f,T}\sum_{s=1}^{\lfloor Tr\rfloor}\sin\left(\frac{2\pi fs}{T}\right)s+T^{-3/2}\sum_{f=1}^{n}\gamma_{4f,T}\sum_{s=1}^{\lfloor Tr\rfloor}\cos\left(\frac{2\pi fs}{T}\right)s\\
&\quad +T^{-3/2}\eta_{\lfloor Tr\rfloor}\\
&= \omega_v\sum_{f=1}^{n}\gamma_{1f}T^{-1}\sum_{s=1}^{\lfloor Tr\rfloor}\sin\left(\frac{2\pi fs}{T}\right)+\omega_v\sum_{f=1}^{n}\gamma_{2f}T^{-1}\sum_{s=1}^{\lfloor Tr\rfloor}\cos\left(\frac{2\pi fs}{T}\right)\\
&\quad +\omega_v\sum_{f=1}^{n}\gamma_{3f}T^{-2}\sum_{s=1}^{\lfloor Tr\rfloor}\sin\left(\frac{2\pi fs}{T}\right)s+\omega_v\sum_{f=1}^{n}\gamma_{4f}T^{-2}\sum_{s=1}^{\lfloor Tr\rfloor}\cos\left(\frac{2\pi fs}{T}\right)s\\
&\quad +T^{-3/2}\eta_{\lfloor Tr\rfloor}\\
&\xrightarrow{d} \omega_v\sum_{f=1}^{n}\gamma_{1f}m_{1f}(r)+\omega_v\sum_{f=1}^{n}\gamma_{2f}m_{2f}(r)\\
&\quad +\omega_v\sum_{f=1}^{n}\gamma_{3f}m_{3f}(r)+\omega_v\sum_{f=1}^{n}\gamma_{4f}m_{4f}(r)\\
&\quad +\omega_v\int_0^r W(s)ds \qquad (11)
\end{aligned}$$

A continuous time regression of (11) on r and r^2 yields the limit residual process $\omega_v N_R^{\tau}(r,\boldsymbol{\gamma}^{\tau})$. Again, similar reasoning establishes that the limit process for $T^{-3/2}e_{Ut}$ is the continuous time residual process $\omega_v N_U^{\tau}(r,\boldsymbol{\gamma}^{\tau})$ and then from the CMT

$$\frac{T^{-1}\sum_{t=1}^{T}(T^{-3/2}e_{Rt})^2}{T^{-1}\sum_{t=1}^{T}(T^{-3/2}e_{Ut})^2}-1\xrightarrow{d}\frac{\int_0^1 N_R^{\tau}(r,\boldsymbol{\gamma}^{\tau})^2dr}{\int_0^1 N_U^{\tau}(r)^2dr}-1$$

Proof of Theorem 3

(a) Parts (i) and (ii) follow because it is easily shown that when the model errors are I(0) the unit root statistics $|DF_T^i|$ diverge to $+\infty$, while the B_T^i unit root statistics converge in probability to zero. This is because these statistics are constructed from the residuals $\hat{u}_t^i$ which are based on an OLS regression with a fully specified (or over-specified) deterministic component. In turn, our chosen exponential functions of these statistics all converge in probability to one because we assume $j>0$ and $b_\xi>0$. As a consequence, the original and modified Wald statistics are asymptotically equivalent.

(b) In part (i), since $\hat{u}_t^{\mu}$ are residuals from the OLS regression (1) with $\mu_t=\alpha$ and $\lambda_{t,T}$

given by (2), it is easily shown that

$$T^{-1/2}\hat{u}^{\mu}_{\lfloor Tr \rfloor} \xrightarrow{d} K^{\mu}(r).$$

Similarly, in part (ii), $\hat{u}^{\tau}_t$ are the OLS residuals from (1) with $\mu_t = \alpha + \beta t$ and $\lambda_{t,T}$ given by (3), and hence

$$T^{-1/2}\hat{u}^{\tau}_{\lfloor Tr \rfloor} \xrightarrow{d} K^{\tau}(r).$$

The stated limit distributions of DF^i_T and B^i_T then follow from a straightforward extension of results in Harvey *et al.* (2006). The distributional results for the modified Wald statistics follow via joint convergence and the CMT.

References

Becker, R., Enders, W. and Hurn, S. (2004), "A general test for time dependence in parameters," *Journal of Applied Econometrics* 19, 899–906.

Becker, R., Enders, W. and Lee, J. (2006), "A stationarity test in the presence of an unknown number of smooth breaks," *Journal of Time Series Analysis* 27, 381–409.

Breitung, J. (2002), "Nonparametric tests for unit roots and cointegration," *Journal of Econometrics* 108, 343–363.

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

Enders, W. and Lee, J. (2008), "The flexible Fourier form and testing for unit roots: an example of the term structure of interest rates," Mimeo, University of Alabama.

Harvey, D. I., Leybourne, S. J. and Taylor, A. M. R. (2006), "On robust trend function hypothesis testing," *Studies in Nonlinear Dynamics and Econometrics* 10, Issue 1, Article 1.

Ng, S. and Perron, P. (2001), "Lag length selection and the construction of unit root tests with good size and power," *Econometrica* 69, 1519–1554.

Perron, P. and Qu, Z. (2007), "A simple modification to improve the finite sample properties of Ng and Perron's unit root tests," *Economics Letters* 94, 12–19.

Park, J. Y. (1990), "Testing for unit roots and cointegration by variable addition," in Fomby, T. and Rhodes, F. (eds.), *Advances in Econometrics: Cointegration, Spurious Regression and Unit Roots*, Jai Press, Greenwich.

Park, J. Y. and Choi, B. (1988), “A new approach to testing for a unit root,” CAE Working Paper 88-23, Cornell University.

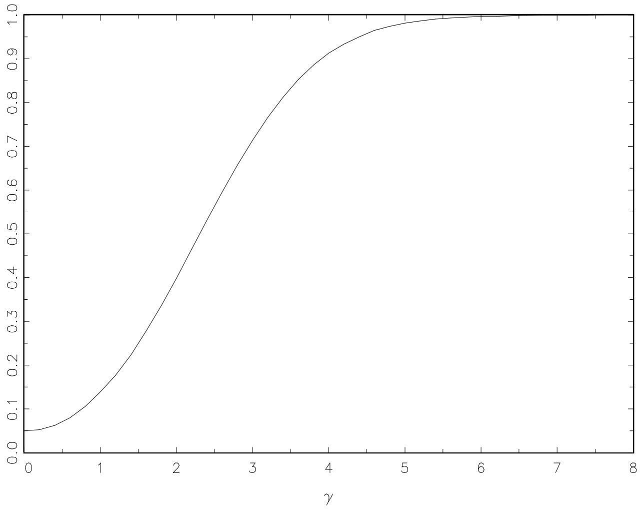
Vogelsang, T. J. (1998), “Trend function hypothesis testing in the presence of serial correlation,” *Econometrica* 66, 123–148.

Table 1. Asymptotic critical values and b_ξ values for ξ-level tests

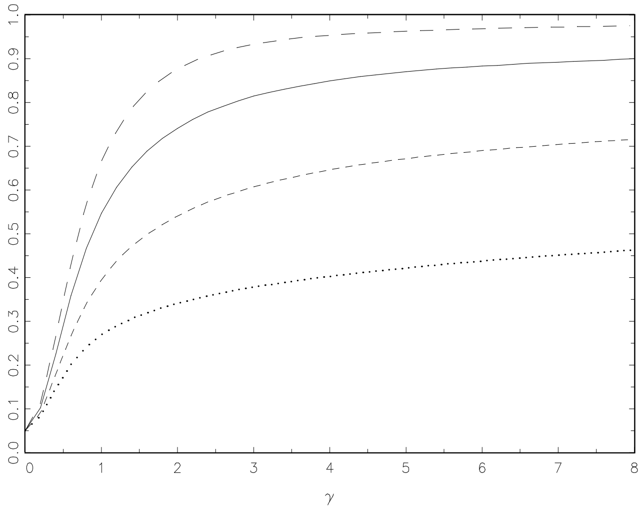
Panel A. Mean case						
	$n=1$			$n=2$		
	$\xi=0.10$	$\xi=0.05$	$\xi=0.01$	$\xi=0.10$	$\xi=0.05$	$\xi=0.01$
Critical value	5.268	7.439	13.370	9.337	12.754	21.495
b_ξ: $MW_T^\mu(DF_T^{\mu,0.5})$	3.692	3.962	4.668	5.548	5.835	6.534
b_ξ: $MW_T^\mu(DF_T^{\mu,1})$	6.316	7.096	8.973	10.582	11.554	13.688
b_ξ: $MW_T^\mu(DF_T^{\mu,2})$	19.604	23.700	34.262	40.477	47.387	62.423
b_ξ: $MW_T^\mu(B_T^\mu)$	406.247	566.707	1065.633	1246.742	1740.261	2890.065
Panel B. Trend case						
	$n=1$			$n=2$		
	$\xi=0.10$	$\xi=0.05$	$\xi=0.01$	$\xi=0.10$	$\xi=0.05$	$\xi=0.01$
Critical value	4.466	5.859	9.472	7.821	10.026	15.512
b_ξ: $MW_T^\tau(DF_T^{\tau,0.5})$	5.968	6.369	7.124	8.459	8.857	9.652
b_ξ: $MW_T^\tau(DF_T^{\tau,1})$	12.191	13.330	15.616	19.309	20.607	23.172
b_ξ: $MW_T^\tau(DF_T^{\tau,2})$	52.474	59.878	77.079	102.937	114.129	138.625
b_ξ: $MW_T^\tau(B_T^\tau)$	3537.707	4192.037	5748.313	10999.097	12534.431	15986.885

Table 2. Finite sample sizes of nominal 0.05-level $MW_T^{\mu}(\cdot)$ tests, mean case, $n = 1$, $T = 150$

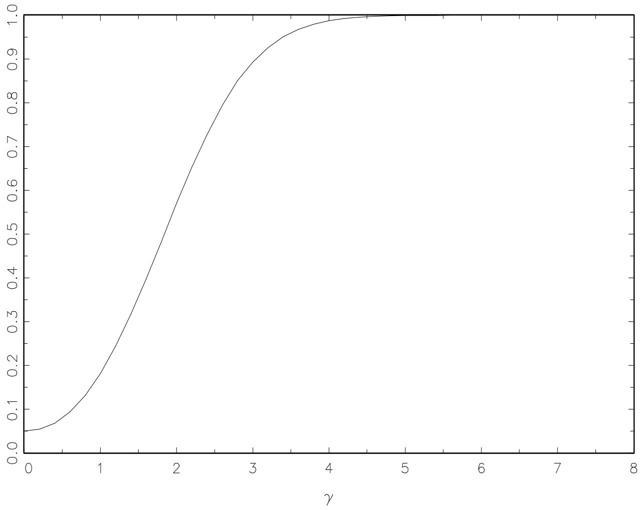
		$T = 150$				$T = 300$				$T = 600$			
ϕ	θ	$DF_T^{\mu,0.5}$	$DF_T^{\mu,1}$	$DF_T^{\mu,2}$	B_T^{μ}	$DF_T^{\mu,0.5}$	$DF_T^{\mu,1}$	$DF_T^{\mu,2}$	B_T^{μ}	$DF_T^{\mu,0.5}$	$DF_T^{\mu,1}$	$DF_T^{\mu,2}$	B_T^{μ}
0.00	−0.50	0.001	0.006	0.026	0.032	0.001	0.010	0.034	0.039	0.002	0.016	0.039	0.044
	0.00	0.001	0.009	0.027	0.035	0.002	0.012	0.031	0.042	0.003	0.015	0.035	0.046
	0.50	0.000	0.004	0.013	0.014	0.002	0.011	0.025	0.025	0.005	0.022	0.035	0.035
0.50	−0.50	0.001	0.003	0.013	0.029	0.001	0.004	0.022	0.034	0.001	0.008	0.030	0.041
	0.00	0.001	0.007	0.025	0.033	0.001	0.008	0.029	0.038	0.002	0.011	0.033	0.043
	0.50	0.001	0.009	0.027	0.035	0.002	0.012	0.031	0.042	0.003	0.015	0.035	0.046
0.70	−0.50	0.001	0.002	0.008	0.028	0.001	0.003	0.014	0.031	0.001	0.005	0.025	0.038
	0.00	0.002	0.006	0.022	0.032	0.001	0.006	0.025	0.035	0.001	0.008	0.031	0.041
	0.50	0.002	0.008	0.023	0.045	0.001	0.006	0.022	0.045	0.001	0.006	0.025	0.048
0.90	−0.50	0.006	0.009	0.015	0.040	0.002	0.005	0.014	0.032	0.001	0.004	0.017	0.033
	0.00	0.008	0.015	0.027	0.045	0.002	0.007	0.020	0.035	0.001	0.005	0.023	0.034
	0.50	0.008	0.015	0.028	0.073	0.002	0.007	0.018	0.052	0.001	0.004	0.019	0.045
0.95	−0.50	0.021	0.024	0.028	0.057	0.006	0.011	0.018	0.042	0.002	0.005	0.016	0.034
	0.00	0.024	0.030	0.040	0.063	0.008	0.013	0.025	0.045	0.002	0.007	0.021	0.036
	0.50	0.023	0.031	0.044	0.103	0.007	0.014	0.025	0.065	0.002	0.007	0.020	0.047
1.00	−0.50	0.044	0.039	0.035	0.047	0.044	0.040	0.037	0.047	0.047	0.045	0.044	0.048
	0.00	0.048	0.047	0.045	0.051	0.048	0.047	0.046	0.049	0.049	0.048	0.046	0.049
	0.50	0.049	0.049	0.051	0.079	0.049	0.048	0.049	0.066	0.050	0.050	0.050	0.058



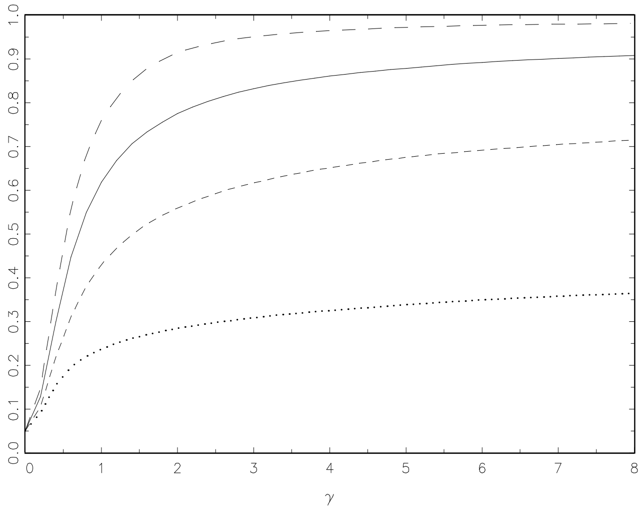
(a) I(0) errors, $n = 1$



(b) I(1) errors, $n = 1$

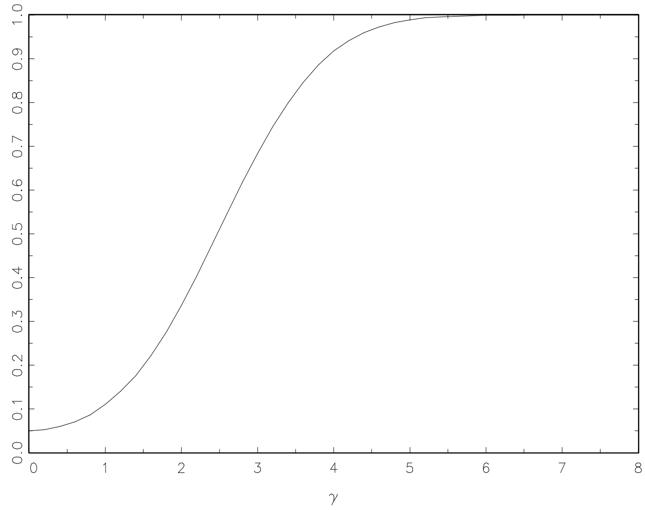


(c) I(0) errors, $n = 2$

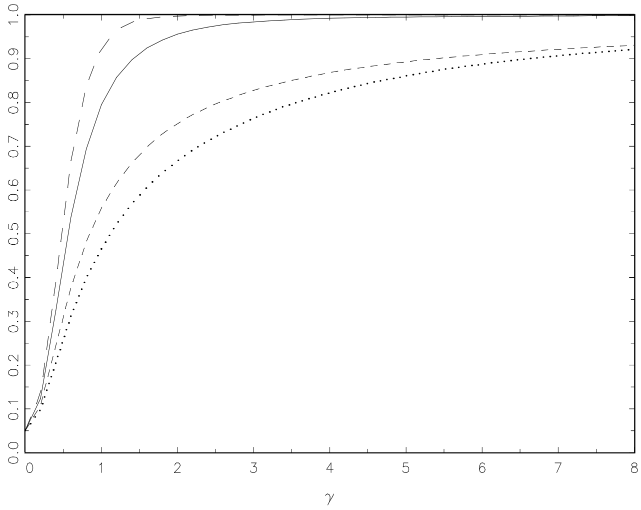


(d) I(1) errors, $n = 2$

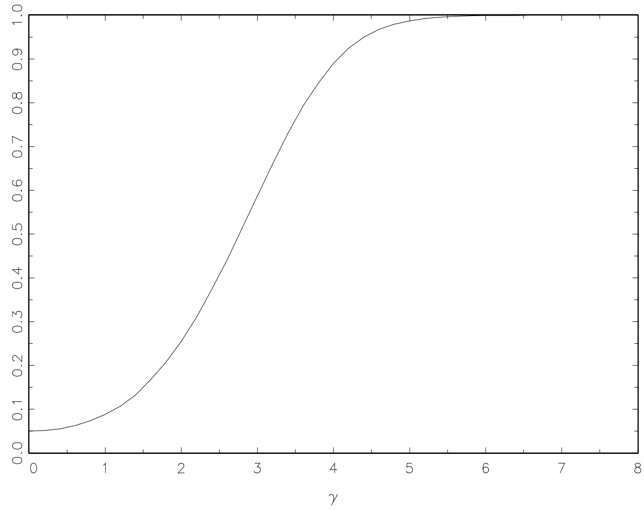
Figure 1. Asymptotic local power of tests, mean case:
$MW_T^{\mu}(DF_T^{\mu,0.5})$: $--$, $MW_T^{\mu}(DF_T^{\mu,1})$: —— , $MW_T^{\mu}(DF_T^{\mu,2})$: - - - , $MW_T^{\mu}(B_T^{\mu})$: $\cdots$



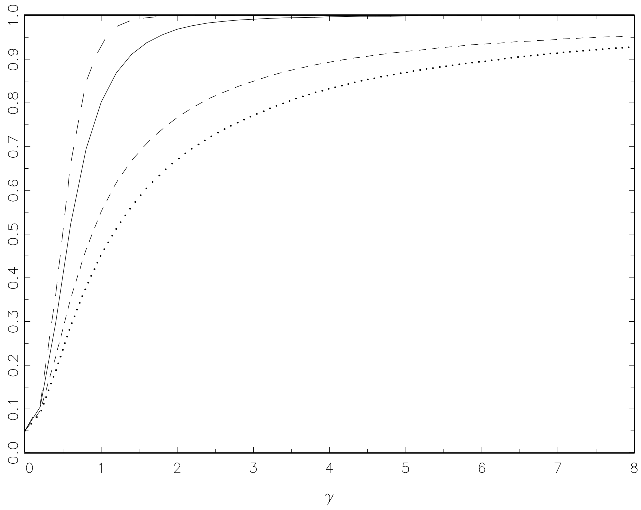
(a) I(0) errors, $n = 1$



(b) I(1) errors, $n = 1$

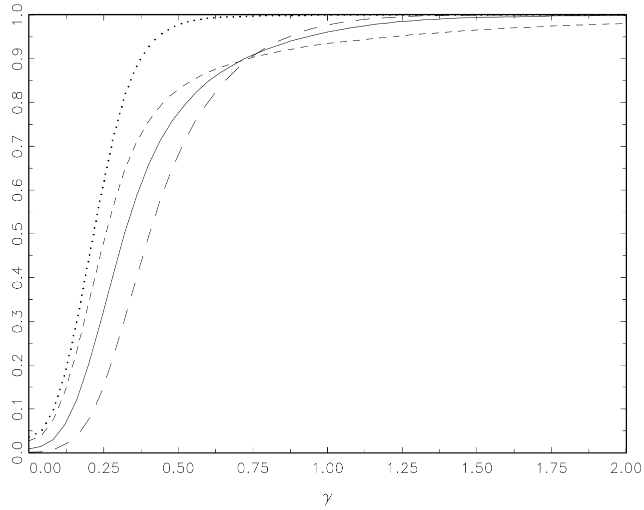


(c) I(0) errors, $n = 2$

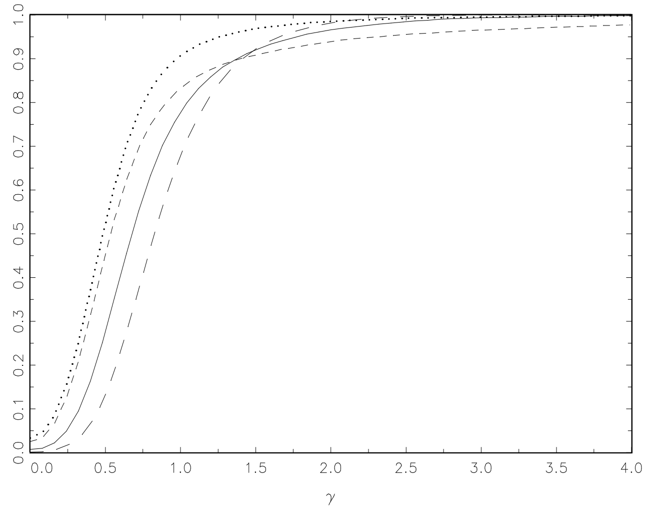


(d) I(1) errors, $n = 2$

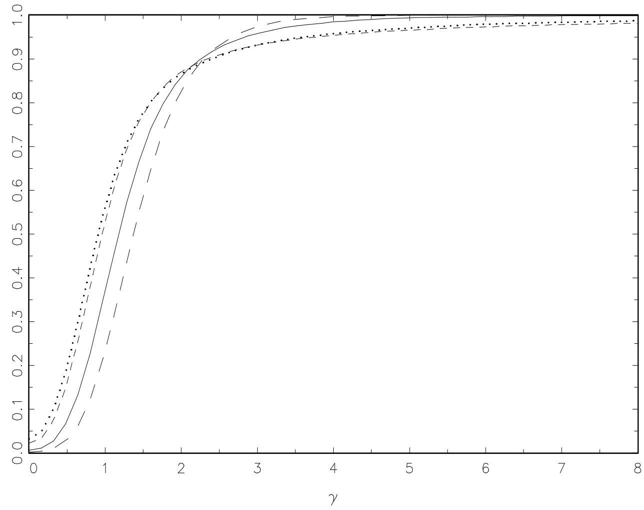
Figure 2. Asymptotic local power of tests, trend case:
$MW_T^{\tau}(DF_T^{\tau,0.5})$: – – , $MW_T^{\tau}(DF_T^{\tau,1})$: —— , $MW_T^{\tau}(DF_T^{\tau,2})$: - - - , $MW_T^{\tau}(B_T^{\tau})$: $\cdots$



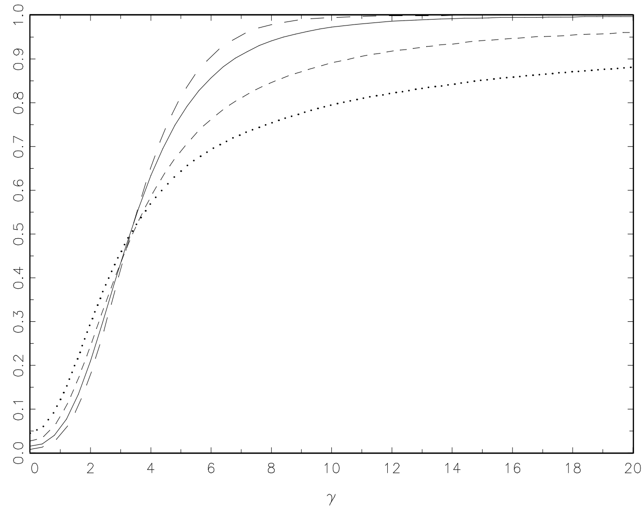
(a) $\phi = 0.00$



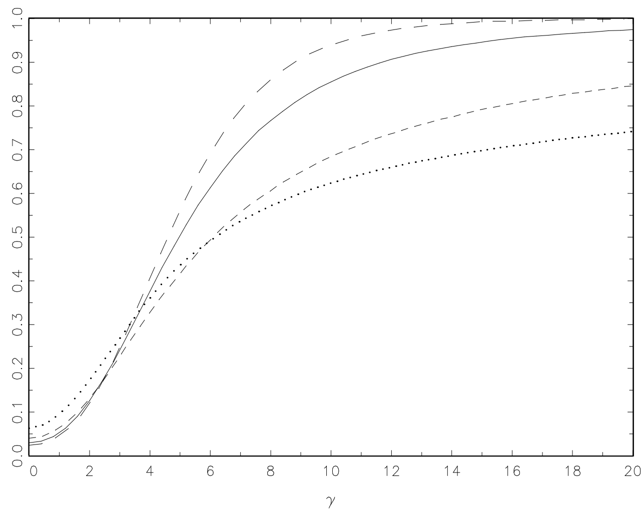
(b) $\phi = 0.50$



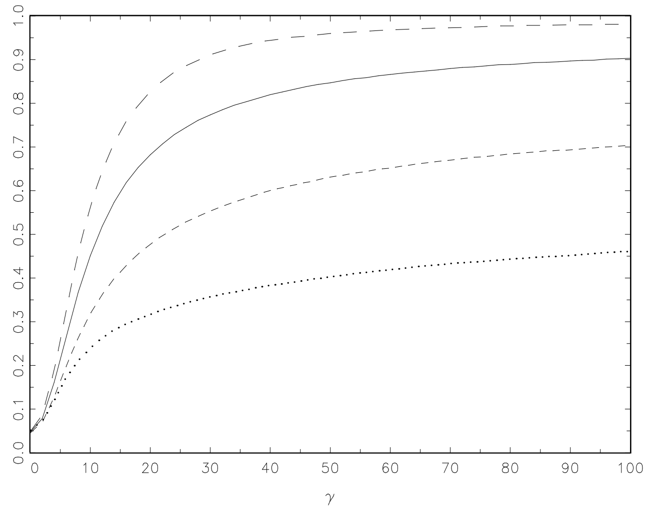
(c) $\phi = 0.70$



(d) $\phi = 0.90$



(e) $\phi = 0.95$



(f) $\phi = 1.00$

Figure 3. Finite sample power of tests, mean case, n=1, $T = 150$:
$MW_T^\mu(DF_T^{\mu,0.5})$: $--$, $MW_T^\mu(DF_T^{\mu,1})$: — , $MW_T^\mu(DF_T^{\mu,2})$: - - - , $MW_T^\mu(B_T^\mu)$: $\cdots$