doi:10.1111/j.1467-9892.2006.00478.x

A STATIONARITY TEST IN THE PRESENCE OF AN UNKNOWN NUMBER OF SMOOTH BREAKS

By Ralf Becker, Walter Enders and Junsoo Lee

University of Alabama and University of Manchester

First Version received December 2004

Abstract. Macroeconomic variables have been shown to display a wide variety of structural breaks of unknown number, duration and form. This poses a challenge since improperly modelled breaks can result in a seriously misspecified model. In this paper, we develop a new test for stationarity that approximates the unknown form of structural breaks using a selected frequency component from a Fourier approximation. Our proposed test performs quite well when breaks are gradual, and shows reasonable power. The appropriate use of the test is illustrated by examining real exchange rates in the post-Bretton Woods period.

Keywords. Structural breaks; stationarity tests; Fourier approximation.
JEL classifications numbers. C12, C22, E17.

1. INTRODUCTION

Perron's (1989) seminal paper made clear the importance of properly modelling structural breaks in testing for a unit root. It is now well known that stationarity and unit-root tests can be invalid when a researcher improperly specifies the nature of any structural breaks present in the data-generating process (DGP).[1] This poses a serious problem for practitioners since important macroeconomic variables can display a wide variety of structural breaks of unknown number, duration and form. A researcher who is unsure about whether a series is stationary is unlikely to know the proper way to model the potential breaks. A complicating factor is that a break occurring at time *t* need not manifest itself contemporaneously. Even major breaks, such as the stock market crash of 1929 and the oil price shocks of the 1970s, did not display their full impacts immediately.

There are papers that allow for the possibility of a pre-specified number of structural breaks. For example, Clemente *et al.* (1998), Vogelsang and Perron (1998), Sen (2003), and Lee and Strazicich (2003) develop unit-root tests that use dummy variables to capture the possibility of changes in the level and trend of a series. Busetti and Harvey (2001), Kurozumi (2002), Presno and Lopez (2003), Harvey and Mills (2003), and Busetti and Taylor (2003) modify the standard Kwaitowski, Phillips, Schmidt and Shin (KPSS) (1992) test by including dummy variables to capture changes in the level and trend. Note that all these papers require that the break(s) be sharp. Since time dummies might not capture the nature of the breaks, Luukkonen *et al.* (1988), Kapetanios *et al.* (2003), and Leybourne *et al.* (1998) develop tests for a unit root allowing for a single break such that the deterministic component of the model is a smooth transition process. Similarly, Harvey and Mills (2004) posit a stationarity test with a smooth transition in the linear trend. These papers assume a pre-specified number of sharp breaks or one specific type of nonlinearity. However, the actual nature of the break(s) is generally unknown, and there is no specific guide as to where and how many breaks to use in testing for a unit root or stationarity. Using an incorrect specification for the form and/or number of breaks can be as problematic as ignoring the breaks altogether.

In order to mitigate the problem of controlling for breaks of an unknown form and number, we develop a stationarity test that uses a selected frequency component of a Fourier function to approximate the deterministic components of the model. Several authors, including Gallant (1984), Davies (1987), Gallant and Souza (1991), and Becker *et al.* (2004), show that a Fourier approximation can often capture the behaviour of an unknown function even if the function itself is not periodic. In Section 2, we illustrate this point by showing that a series with various types of breaks can often be captured using a selected frequency component of a Fourier approximation. Although our methodology can detect sharp breaks, it is designed to work best when breaks are gradual.[2] Moreover, our approximation has good power to detect u-shaped breaks and smooth breaks located near the end of a series. Hence, instead of selecting specific break dates, the number of breaks, and the form of the breaks, the specification problem is transformed into selecting the proper frequency component to be included in the estimating equation.

Since we use trigonometric terms to capture unknown nonlinearities, our work parallels the unit-root tests suggested by Bierens (1997) and Enders and Lee (2005). Here, we develop a KPSS-type stationarity test since tests with the null of a unit root have low power with stationary, but persistent data. The problem of low power is exacerbated when a theory, such as purchasing power parity or the convergence of growth rates across nations, can be more naturally tested under the null of stationarity. Stationarity tests are also useful since they can be used to confirm results from unit-root tests with a stationary alternative.

2. TEST STATISTICS

We consider the following DGP:

$$\begin{aligned} y_t &= X_t'\beta + Z_t'\gamma + r_t + \varepsilon_t \\ r_t &= r_{t-1} + u_t, \end{aligned} \tag{1}$$

where ε_t are stationary errors and u_t are independent and identically distributed (i.i.d.) with variance σ_u^2. Here, we use $X_t = [1]$ for a level-stationary process for y_t and $X_t = [1, t]'$ for a trend-stationary process. We choose $Z_t = [\sin(2\pi kt/T), \cos(2\pi kt/T)]'$ to capture a break (or other type of unattended nonlinearity) in the deterministic term, where k is the frequency and T the sample size. Under the null hypothesis $\sigma_u^2 = 0$, so that the process described by eqn (1) is stationary.

The rationale for selecting $Z_t = [\sin(2\pi kt/T), \cos(2\pi kt/T)]'$ is based on the fact that a Fourier expansion is capable of approximating absolutely integrable functions to any desired degree of accuracy. Specifically, let $\alpha(t)$ denote a function with an unknown number of breaks of unknown form(s). Regardless of the nature of the breaks, under very weak conditions, $\alpha(t)$ can be approximated to any degree of accuracy by the sufficiently long Fourier series:

$$\alpha(t) = \alpha_0 + \sum_{k=1}^{n} a_k \sin\left(\frac{2\pi kt}{T}\right) + \sum_{k=1}^{n} b_k \cos\left(\frac{2\pi kt}{T}\right); \quad n < \frac{T}{2},$$

where n is the number of frequencies contained in the approximation and k represents a particular frequency.

As is well known, beginning with $n = 1$, it is always possible to improve the approximation by using additional frequencies. When $n = T/2$ is reached, the fit of $\alpha(t)$ will be perfect. However, to keep the problem tractable, we begin by considering a Fourier approximation using a single-frequency component, so that

$$\alpha(t) \cong Z_t'\gamma = \gamma_1 \sin\left(\frac{2\pi kt}{T}\right) + \gamma_2 \cos\left(\frac{2\pi kt}{T}\right) \tag{2}$$

where k represents the frequency selected for the approximation, and $\gamma = [\gamma_1, \gamma_2]'$ measures the amplitude and displacement of the frequency component.

A desirable feature of eqn (2) is that the standard linear specification emerges as a special case by setting $\gamma_1 = \gamma_2 = 0$. It also follows that at least one frequency component must be present if there is a structural break. Hence, if it is possible to reject the null hypothesis $\gamma_1 = \gamma_2 = 0$, the series must have a nonlinear component. Becker *et al.* (2004) use this property of eqn (2) to develop a test that can have more power to detect breaks of an unknown form than the standard Bai and Perron (1998) test.

It should be clear that eqn (2) can best mimic a break when $\alpha(t)$ is smooth. Nevertheless, to gain some intuition concerning eqn (2), consider the solid line in panel 1 of Figure 1 that depicts a sharp u-shaped break in the intercept of a series $\{y_t\}$. The break is such that $y_t = 0.80$ for $33 \leqslant t \leqslant 66$ and $y_t = 1.0$ otherwise. The dashed line in the panel shows the approximation to the series obtained by setting $k = 1$, $\gamma_1 = -0.004$, $\gamma_2 = 0.112$, $X_t = [1]$, and $\beta = 0.932$. The values of β, γ_1 and γ_2 were selected by regressing y_t on Z_t and a constant for each integer frequency in the interval (1, 5). The frequency $k = 1$ was selected as it provided the smallest value of the sum of squared residuals (SSR = 0.275). In contrast, if we use only an intercept term, SSR = 0.898. Panel 2 shows the effect of moving the break towards the beginning of the

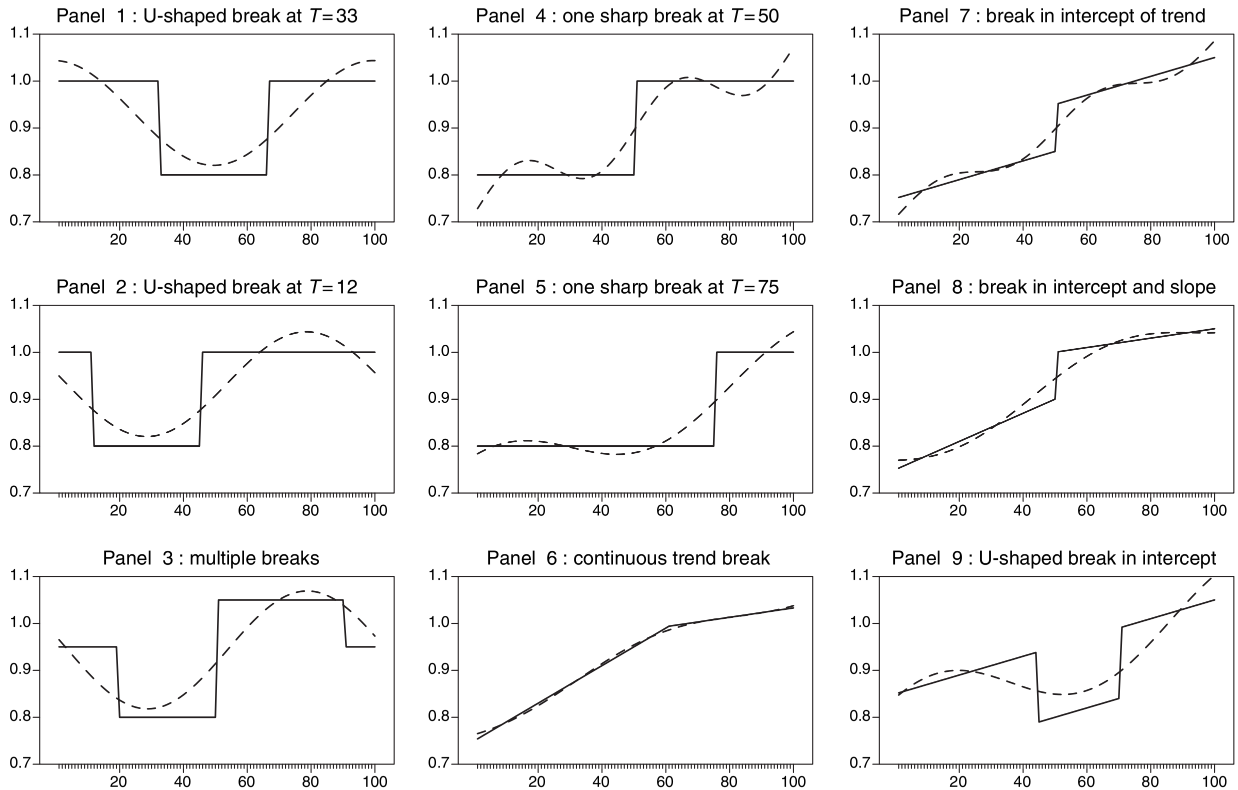


FIGURE 1. Fourier approximations. Panel 1: $y_t = 0.8$ if $33 \leqslant t \leqslant 66$, else $y_t = 1.0$; panel 2: $y_t = 0.8$ if $12 \leqslant t \leqslant 45$, else $y_t = 1.0$; panel 3: $y_t = 0.8$ if $20 \leqslant t \leqslant 50$, $y_t = 1.05$ if $51 \leqslant t \leqslant 90$, else $y_t = 0.95$; panel 4: $y_t = 0.8$ if $t \leqslant 50$, else $y_t = 1.0$; panel 5: $y_t = 0.8$ if $t \leqslant 75$, else $y_t = 1.0$; panel 6: $y_t = 0.75 + 0.004t$ if $t \leqslant 60$, else $y_t = 0.933 + 0.001t$; panel 7: $y_t = 0.75 + 0.002t$ if $t \leqslant 50$, else $y_t = 0.85 + 0.002t$; panel 8: $y_t = 0.75 + 0.003t$ if $t \leqslant 50$, else $y_t = 0.95 + 0.002t$; panel 9: $y_t = 0.70 + 0.002t$ if $45 \leqslant t \leqslant 70$, else $y_t = 0.85 + 0.002t$.

sample. It is interesting to note that the fit of the approximation used in panel 2 is identical to that of panel 1 (SSR = 0.275) although the values of β, γ_1 and γ_2 differ. For a given size and duration, the location of the break will not affect the sum of squared residuals obtained by regressing a single-frequency component on the series in question. This observation is important since some tests for breaks, such as the Bai and Perron (1998) test, have little power to detect u-shaped breaks or breaks located near the end of a series. Panel 3 shows multiple breaks in the intercept such that the starting and ending values are unchanged. Panels 4 to 9 show breaks in the intercept and/or breaks in the slope of a trending function. In these last cases, the dashed lines were obtained by setting $X_t = [1, t]'$ and selecting the integer frequency yielding the best fit. Of course, reordering the entries of the series from high to low or changing the sign of the breaks will not affect the ability of the approximation to capture the essential details of the breaks.

Although the approximation seems to work quite well, our aim is not to develop another test for a break or to explicitly model the form of the break(s). In fact, if the nature of the breaks is known, it might be preferable to use an alternative specification for Z_t. Instead, our goal is to develop a test for stationarity that treats the number and form of breaks as unknown nuisance parameters. We use the fact that $Z_t'\gamma$ in eqn (1) can mimic various nonlinearities as the basis for the stationarity test. Below, we show that the test for $\sigma_u^2 = 0$ is invariant to the values of β, γ_1 and γ_2 present in the DGP. As such, it becomes possible to test for stationarity in the presence of otherwise neglected breaks or nonlinearities.

When Z_t is absent, the DGP in eqn (1) corresponds to that of KPSS. They suggest testing the null hypothesis of stationarity (i.e. $\sigma_\mu^2 = 0$) using the sample statistic

$$\tau_{\text{KPSS}} = \frac{1}{T^2}\frac{\sum_{t=1}^{T}\hat{S}_t^2}{\hat{\sigma}^2}$$

where $\hat{S}_t = \Sigma_{j=1}^{t}\hat{e}_j$ and $\hat{e}_t$ are the ordinary least square (OLS) residuals obtained from regressing y_t on X_t. The estimate of the long-run error variance,

$$\sigma^2 = \lim_{T\to\infty} T^{-1}E(S_T^2), \quad \text{where } S_T = \sum_{t=1}^{T}\varepsilon_t,$$

can be obtained using a nonparametric correction for non-i.i.d. errors. We need only a slight modification of the KPSS statistic when allowing for the existence of a time-varying intercept under the null hypothesis. Simply, let $\tilde{e}_t$ denote the residuals from the following regression:

$$y_t = \alpha + \gamma_1 \sin\left(\frac{2\pi kt}{T}\right) + \gamma_2 \cos\left(\frac{2\pi kt}{T}\right) + e_t \tag{3a}$$

or

$$y_t = \alpha + \beta t + \gamma_1 \sin\left(\frac{2\pi kt}{T}\right) + \gamma_2 \cos\left(\frac{2\pi kt}{T}\right) + e_t \tag{3b}$$

As such, it is possible obtain the following test statistics:

$$\tau_\mu(k) \text{ or } \tau_\tau(k) = \frac{1}{T^2}\frac{\sum_{t=1}^{T} \tilde{S}_t(k)^2}{\tilde{\sigma}^2} \tag{4}$$

where $\tilde{S}_t(k) = \Sigma_{j=1}^{t}\tilde{e}_j$ and $\tilde{e}_j$ are the OLS residuals from the regression (3a) for $\tau_\mu(k)$ or (3b) for $\tau_\tau(k)$. As in KPSS, a nonparametric estimate $\tilde{\sigma}^2$of the long-run variance can be obtained by choosing a truncation lag parameter l and a set of weights w_j, $j = 1, \ldots, l$:

$$\tilde{\sigma}^2 = \tilde{\gamma}_0 + 2\sum w_j \tilde{\gamma}_j,$$

where $\tilde{\gamma}_j$ is the jth sample autocovariance of the residuals $\tilde{e}_t$ from eqn (3a) or (3b).

To obtain the asymptotic distribution of our test statistics, $\tau_\mu(k)$ or $\tau_\tau(k)$, we need the following results, where we let $[rT]$, $r \in [0, 1]$, be an integer close to rT.[3]

PROPOSITION 1

$$(a) \quad \frac{1}{T}\sum_{t=1}^{T} \sin\left(\frac{2\pi kj}{T}\right) \to \frac{1}{2\pi k}(1 - \cos(2\pi k)) \equiv s_0$$

$$(b) \quad \frac{1}{T}\sum_{t=1}^{T} \cos\left(\frac{2\pi kj}{T}\right) \to \frac{1}{2\pi k}\sin(2\pi k) \equiv c_0$$

$$(c) \quad \frac{1}{T}\sum_{j=1}^{[rT]} \sin\left(\frac{2\pi kj}{T}\right) \to \frac{r}{2\pi k}(1 - \cos(2\pi kr)) \equiv s_r$$

$$(d) \quad \frac{1}{T}\sum_{j=1}^{[rT]} \cos\left(\frac{2\pi kj}{T}\right) \to \frac{r}{2\pi k}\sin(2\pi kr) \equiv c_r$$

$$(e) \quad \frac{1}{T^2}\sum_{t=1}^{T} t \cdot \sin\left(\frac{2\pi kj}{T}\right) \to \frac{1}{(2\pi k)^2}\sin(2\pi k) - \frac{1}{2\pi k}\cos(2\pi k) \equiv s_1$$

$$(f) \quad \frac{1}{T^2}\sum_{t=1}^{T} t \cdot \cos\left(\frac{2\pi kj}{T}\right) \to \frac{1}{(2\pi k)^2}[\cos(2\pi k) + (2\pi k)\sin(2\pi k) - 1] \equiv c_1$$

$$(g) \quad \frac{1}{T}\sum_{t=1}^{T} \sin^2\left(\frac{2\pi kt}{T}\right) \to \frac{1}{2} - \frac{\sin(4\pi k)}{4\pi k} \equiv s_2$$

$$(h) \quad \frac{1}{T}\sum_{t=1}^{T} \cos^2\left(\frac{2\pi kt}{T}\right) \to \frac{1}{2} + \frac{\sin(4\pi k)}{4\pi k} \equiv c_2$$

$$(i) \quad \frac{1}{\sqrt{T}}\sum_{t=1}^{T} e_t \to \sigma W(1) \equiv \sigma f_1$$

$$(j) \quad \frac{1}{T^{1.5}}\sum_{t=1}^{T} t \cdot e_t \to \sigma\left[W(1) - \int_0^1 W(r)\mathrm{d}r\right] \equiv \sigma f_2$$

$$(k) \quad \frac{1}{\sqrt{T}}\sum_{t=1}^{T} e_t \sin\left(\frac{2\pi kt}{T}\right) \to \sigma(2\pi k)\left[W(1) + (2\pi k)\int_0^1 \sin(2\pi kr)W(r)\mathrm{d}r\right] \equiv \sigma f_3$$

$$(l) \quad \frac{1}{\sqrt{T}}\sum_{t=1}^{T} e_t \cos\left(\frac{2\pi kt}{T}\right) \to \sigma(2\pi k)^2\left[\int_0^1 \cos(2\pi kr)W(r)\mathrm{d}r\right] \equiv \sigma f_4.$$

In cases where k is restricted to be an integer, it is possible to simplify terms such that: $s_0 = 0$, $c_0 = 0$, $s_1 = -1/(2\pi k)$, $c_1 = 0$, $s_2 = 1/2$ and $c_2 = 1/2$. Utilizing the above results, we can show that the asymptotic distribution of $\tau_i(k)$, $i = \mu, \tau$, is given as follows.

THEOREM 1. *Suppose that* y_t *is generated by the DGP in eqn* (1) *with* $\sigma_u^2 = 0$ *under the null, and one adopts the testing regression* (3a) *or* (3b). *Then, under the null hypothesis:*

$$\tau_i(k) \to \int_0^1 \underline{V}_i(r)^2\mathrm{d}r, \quad i = \mu, \tau \tag{5}$$

where $\underline{V}_i(r)$ *is the projection of the Wiener process* $W(r)$ *on the orthogonal complement of the space spanned by the nonlinear function* $f_i(k, r)$, $i = \mu, \tau$, *with*

$$f_\mu(k,r) = [1, \sin(2\pi kr), \cos(2\pi kr)]' \quad \text{and} \quad f_\tau(k,r) = [1, r, \sin(2\pi kr), \cos(2\pi kr)]',$$

as defined over the interval $r \in [0, 1]$, *such that*

$$\underline{V}_i(r) = W(r) - f_i(k,r)\hat{\gamma}, \quad \textit{with } \hat{\gamma} = \text{argmin} \int_0^1 (W(r) - f_i(k,r)\gamma)^2 \mathrm{d}r.$$

The precise expressions for $\underline{V}_i(r)$, $i = \mu, \tau$, *are relegated to eqns* (A.2) *and* (A.4) *in the Appendix.*

It is clear that the asymptotic distributions of the resulting test statistics depend on the frequency k, but are invariant to the other parameters in the DGP. To obtain critical values via simulations, we employ the DGP in eqn (1) with $\beta = \gamma = 0$. However, the invariance of the test statistics to β and γ means that any other choice would lead to the identical critical values. Pseudo-i.i.d. $N(0, 1)$ random numbers were generated using the Gauss procedure RNDN and all calculations were conducted using the Gauss software version 6.0.10. The initial values y_0 and ε_0 are assumed to be random, and we set $\sigma_\varepsilon^2 = 1$. The critical values are reported in Table Ia for the sample sizes $T = 100$, 500 and 1000. The critical values with $T = 1000$ can be considered as asymptotic values that can be used when $T > 1000$.[4] These critical values were calculated using 50,000 Monte Carlo replications for values of $k = 1, \ldots, 5$.

We consider three different ways to select frequency component(s) to include in the testing equation. As illustrated in Figure 1, selecting the value $k = 1$ (or possibly $k = 2$) is sufficient to replicate the essential details of a large number of breaks. Given this pre-specified value of k, it is possible to estimate eqn (3a) or (3b), calculate the value of $\tau_\mu(k)$ or $\tau_\tau(k)$, and perform the stationarity test using the critical values in Table Ia. For several reasons, we do not recommend using higher frequencies when pre-specifying the value of k. First, the approximation is likely to work best for data that have several smooth breaks of an unknown variety. In particular, the so-called 'long-swings' often identified in real exchange rate data seem to be particularly suitable to our approximation. Moreover, since breaks shift the spectral density function towards frequency zero, the most appropriate frequency for a break is likely to be at the low end of the spectrum. In fact, Becker *et al.* (2004) show that the higher frequencies are likely to be associated with stochastic parameter variability (not structural breaks). Hence, it is the low frequencies that are the most likely ones to interfere with a test for stationarity versus nonstationarity.

Instead of using a single pre-specified frequency component, Gallant (1984) and Bierens (1997) use cumulative frequencies to estimate the unknown functional form. After all, if the use of $k = 1$ can reasonably represent an unknown

TABLE I

(a) Critical Values for $\tau_\mu(k)$ and $\tau_\tau(k)$, Single Frequency; (b) Critical Values for $\tau_\mu(n)$ and $\tau_\tau(n)$, Cumulative Frequency; (c) Critical Values for $F_i(\hat{k})$

		$\tau_\mu(k)$-level			$\tau_\tau(k)$-trend		
T	k	10%	5%	1%	10%	5%	1%
(a)							
100	1	0.1318	0.1720	0.2699	0.0471	0.0546	0.0716
	2	0.3150	0.4152	0.6671	0.1034	0.1321	0.2022
	3	0.3393	0.4480	0.7182	0.1141	0.1423	0.2103
	4	0.3476	0.4592	0.7222	0.1189	0.1478	0.2170
	5	0.3518	0.4626	0.7386	0.1201	0.1484	0.2177
500	1	0.1294	0.1696	0.2709	0.0463	0.0539	0.0720
	2	0.3053	0.4075	0.6615	0.0995	0.1278	0.1968
	3	0.3309	0.4424	0.7046	0.1123	0.1404	0.2091
	4	0.3369	0.4491	0.7152	0.1155	0.1441	0.2111
	5	0.3415	0.4571	0.7344	0.1178	0.1465	0.2178
1000	1	0.1295	0.1704	0.2706	0.0461	0.0538	0.0718
	2	0.3050	0.4047	0.6526	0.0994	0.1275	0.1959
	3	0.3304	0.4388	0.7086	0.1117	0.1398	0.2081
	4	0.3355	0.4470	0.7163	0.1149	0.1436	0.2139
	5	0.3422	0.4525	0.7297	0.1163	0.1451	0.2153

		$\tau_\mu(n)$-level			$\tau_\tau(n)$-trend		
T	n	10%	5%	1%	10%	5%	1%
(b)							
100	1	0.1323	0.1735	0.2700	0.0472	0.0548	0.0718
	2	0.0800	0.1048	0.1638	0.0282	0.0318	0.0399
	3	0.0589	0.0769	0.1203	0.0201	0.0222	0.0268
	4	0.0461	0.0599	0.0925	0.0155	0.0169	0.0201
	5	0.0384	0.0499	0.0777	0.0126	0.0136	0.0158
500	1	0.1290	0.1688	0.2696	0.0462	0.0538	0.0714
	2	0.0778	0.1023	0.1614	0.0276	0.0312	0.0397
	3	0.0553	0.0729	0.1157	0.0193	0.0216	0.0265
	4	0.0433	0.0568	0.0901	0.0148	0.0162	0.0196
	5	0.0354	0.0461	0.0723	0.0119	0.0130	0.0154
1000	1	0.1289	0.1691	0.2671	0.0461	0.0541	0.0719
	2	0.0772	0.1020	0.1623	0.0274	0.0311	0.0395
	3	0.0554	0.0725	0.1151	0.0192	0.0215	0.0264
	4	0.0429	0.0564	0.0888	0.0147	0.0162	0.0195
	5	0.0351	0.0456	0.0721	0.0119	0.0129	0.0153

	$F_\mu(\hat{k})$-level			$F_\tau(\hat{k})$-trend		
T	10%	5%	1%	10%	5%	1%
(c)						
100	4.133	4.929	6.730	4.162	4.972	6.873
500	3.935	4.651	6.281	3.928	4.669	6.315

functional form, the use of $k = 1$ and $k = 2$ can provide a better approximation. Notice that the appropriate test statistics using cumulative frequencies are dependent only on the particular frequencies employed in the testing regression since the individual frequency components are orthogonal to each other. In

Table Ib, we report the critical values of $\tau_\mu(n)$ or $\tau_\tau(n)$, for $n = 1, \ldots, 5$; we denote these critical values by $\tau_\mu(n)$ or $\tau_\tau(n)$ in order to differentiate them from the case where a single-frequency component is used. In Section 3, we present Monte Carlo evidence indicating that using more than one or two cumulative frequencies often leads to a significant loss of power. Thus, we suggest that more than one or two cumulative frequencies not be used. Unlike the case of determining the best-fitting single frequency, it is not feasible to test the optimal number of cumulative frequencies since adding more frequencies will necessarily reduce the SSR. Bierens (1997) notes a similar problem in his unit-root test using cumulative frequencies.

In most instances with highly persistent macroeconomic data, using the value $k = 1$ or $k = 2$ should be sufficient to capture the important breaks in the data. However, there are circumstances where the researcher may want to select some frequency other than $k = 1$ or $k = 2$. Hence, the third method we consider is to select k for using a completely data-driven method. We follow Davies (1987) by selecting the value of $k = \hat{k}$ that minimizes the SSR from regression (3a) or (3b). Specifically, for each integer value of k in the interval $1 \leqslant k \leqslant k^{\max}$, we estimate eqn (3a) or (3b) and select the value yielding the best fit.[5] As a practical matter, we set the maximum frequency at $k^{\max} = 5$ since low frequencies are associated with breaks. Now consider constructing eqn (4) by using the estimated frequency $\hat{k}$. We denote the test statistic utilizing $\hat{k}$ by:

$$\tau_i(\hat{k}) = \tau_i(k) \qquad \text{with } \hat{k}, i = \mu, \tau. \tag{6}$$

Determining the asymptotic distribution of $\hat{k}$ is complicated since we cannot assume that the data are stationary.[6] Moreover, the distribution of $\tau_i(\hat{k})$ depends on how accurately the selected trigonometric component mimics the essential features of the DGP. Thus, we resort to simulation methods to gauge the speed of convergence of the frequency estimate. The simulation results reported in section 3 show encouraging results that the unknown frequency k can be estimated quite accurately even in small samples. Since the simulations support the supposition that Davis' grid-search procedure yields a consistently estimated value of k when the true DGP is given by eqn (3a) or (3b), we suggest using the critical values in Table Ia for the estimated value of k. This data-dependent approach can be useful when no *ex ante* information about the degree of persistence or the number of breaks is available.

A related question of interest is whether any frequency components belong in the data. After all, the test for the null hypothesis $\sigma_u^2 = 0$ makes no specific assumption concerning the presence of breaks in the DGP. However, if the nonlinear trend is not present in the DGP, it is possible to obtain increased power by using the standard KPSS test. Thus, it becomes desirable to test for the absence of a nonlinear trend (i.e. $\gamma_1 = \gamma_2 = 0$). The usual F-test statistic for this hypothesis may be possibly calculated against the alternative of a nonlinear trend. One can consider the following F-test statistic that is calculated against the alternative nonlinear trend with a given frequency k

$$F_i(k) = \frac{(\mathrm{SSR}_0 - \mathrm{SSR}_1(k))/2}{\mathrm{SSR}_1(k)/(T-q)}, \qquad i = \mu, \tau, \tag{7}$$

where $\mathrm{SSR}_1(k)$ denotes the SSR from eqn (3a) or (3b), q is the number of regressors, and SSR_0 denotes the SSR from the regression without the trigonometric terms. When k is unknown, it is not possible to use eqn (7) directly. Instead, consider:

$$F_i(\hat{k}) = \max_k F(k), \quad i = \mu, \tau, \tag{8}$$

where $\hat{k} = \mathrm{argmax}_k F(k)$. It should be clear that eqn (8) is a supremum test since $\hat{k}$ minimizes the SSR for eqn (3a) or (3b); i.e. the F-statistic in eqn (8) is such that $\hat{k} = \mathrm{arginf}_k\, \mathrm{SSR}_1(k)$. If k is estimated, a test for the null hypothesis of linearity is complicated by the fact that k is unidentified under the null hypothesis. Although the distribution of the F-test is non-standard because of the presence of the nuisance parameter, the critical values of the F-test can be tabulated. As such, we simulated the critical values of $F_i(\hat{k})$ under the null of linearity and report them in Table Ic. However, it is important to note that the F-test can exhibit undue power if the data are non-stationary. Thus, the F-test has the limitation that rejection of the null of linearity does not necessarily mean the existence of the nonlinear trend. Therefore, the F-test can be used only when the null of stationarity is not rejected.[7]

One final issue concerns the effect of ignoring a break or nonlinear trend when using the standard KPSS test. Perron (1989) suggested that there is a bias against rejecting a false unit root if an existing structural break is ignored in the usual Dickey–Fuller (DF) test. The nature of the problem is different in stationarity tests since the null and alternative hypotheses are reversed. Lee *et al.* (1997) initially pointed out that stationarity tests exhibit size distortions under the null rather than a loss of power. We observe a similar problem when an existing nonlinear term is ignored. Formally, we can show Lemma 1.

LEMMA 1. *Suppose that a nonlinear term exists in the data such that the DGP implies eqn* (1) *with* $\gamma \neq 0$, *but the nonlinear term is ignored and usual KPSS tests are employed. Then,*

$$\tau_{\mathrm{KPSS}} = O_p(T)$$

The proof is deferred to the Appendix.

As such, the usual KPSS-type stationarity tests will diverge when nonlinear trends are ignored. This leads to over-rejections of the true stationarity null hypothesis in favour of the false unit-root hypothesis. Thus, it is important to control for nonlinear trends in stationary tests. In the Section 3, we examine how serious is the effect of ignoring a nonlinear trend.

3. SMALL SAMPLE PERFORMANCE

In this section, simulation evidence is presented to demonstrate the size and power properties of our test using the sample sizes often found in applied work. All of our Monte Carlo experiments use 20,000 replications of the DGPs given in eqn (3a) or (3b). We first examine the size of the tests using the grid-search method to estimate the frequency k. Thus, $\hat{k}$ is the value that minimizes the SSR from regression (3a) or (3b). Table II reports the 1%, 5% and 10% rejection frequencies for various values of T, k, γ_1 and γ_2 in the DGP. It is evident that the size is quite good when $\gamma_1 = \gamma_2 = 1.0$ or 0.5. It is encouraging to see how often the search procedure picks the correct trigonometric frequency. In these cases, the correct frequency is picked almost 100% of the time. Not surprisingly, the percentage of correct picks increases with sample size and with the importance of the trigonometric terms in the DGP. In Table II, we also report the percentages of underestimating ($\hat{k} < k$), overestimating ($\hat{k} > k$), and correctly estimating ($\hat{k} = k$) the unknown frequency k. An examination of the table shows that the unknown frequency is fairly well estimated. When the coefficients in γ and/or the sample size is small, there is a mild size distortion such that the procedure is too conservative. However, for moderate T, γ_1 and/or γ_2, the size is quite good since the chance of detecting the correct value of k increases accordingly.[8] In Figure 2, we depict the rate of convergence of the estimate of the frequency ($\hat{k}$) to its true value ($k = 1$) in the DGP, for increasing sample sizes of T. The graph shows the proportion of incorrect frequency estimates ($\hat{k} \neq 1$) for varying values of $\gamma = [\gamma_1, \gamma_2]$. It is clear that $\hat{k}$ converges to its correct value quickly when the Fourier coefficients (γ_1 and γ_2) are big. It is also seen that $\hat{k}$ approaches k even when γ is rather small. In particular, when $\gamma = 0.2$, the proportion of incorrect frequency estimates is negligible in sample sizes in excess of 500.

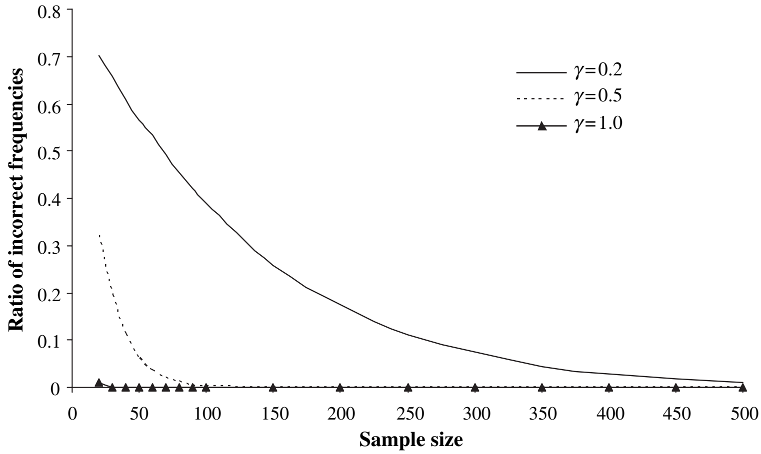
We next examine the power properties of our tests. A concern is that the inclusion of trigonometric functions masks the nonstationarity of a time series. Since a nonstationary time series can move in long swings, the inclusion of a low-frequency trigonometric component might result in a non-rejection of stationarity. Nevertheless, our Monte Carlo experiments demonstrate that this potential problem is rather mild. The DGP under consideration is eqn (1) with various values for σ_u^2 and $\varepsilon_t \sim N(0, 1)$. Thus, the magnitude of σ_u^2 will determine the strength of the departure from the null hypothesis. In this simulation, we set $\gamma_1 = \gamma_2 = 0.5$. Table III reports the simulation results using a known frequency of $k = 1$, 2, or 3. We also report results using the estimated frequency ($\hat{k}$) for each case where the true frequency is 1 and 2. For comparison purposes, the simulation results on the power of the usual KPSS tests are reported as well. As the KPSS test does not allow for a nonlinear trend, the DGP for KPSS results did not include any trigonometric frequencies. As shown in the table, the power properties of our $\tau_\mu(k)$ and$\tau_\tau(k)$ tests using 5% critical values are largely comparable with those of the KPSS tests. When σ_u^2 is small, say $\sigma_u^2 < 0.001$ or lower, the DGP is essentially the same as the case where $\sigma_u^2 \approx 0$. When σ_u^2 gets

TABLE II

REJECTION FREQUENCIES OF $\tau_\mu(\hat{k})$ AND $\tau_\tau(\hat{k})$ TESTS UNDER THE NULL

		$\tau_\mu(\hat{k})$-level						$\tau_\tau(\hat{k})$-trend					
T	k	$\hat{k} < k(\%)$	$\hat{k} = k(\%)$	$\hat{k} > k(\%)$	1%	5%	10%	$\hat{k} < k(\%)$	$\hat{k} = k(\%)$	$\hat{k} > k(\%)$	1%	5%	10%
$\gamma_1 = \gamma_2 = 1$													
100	1	0	100	0	0.010	0.049	0.102	0	100	0	0.010	0.051	0.101
	2	0	100	0	0.010	0.052	0.101	0	100	0	0.010	0.050	0.101
	3	0	100	0	0.010	0.047	0.097	0	100	0	0.011	0.051	0.010
500	1	0	100	0	0.010	0.049	0.100	0	100	0	0.010	0.050	0.100
	2	0	100	0	0.011	0.051	0.098	0	100	0	0.010	0.049	0.099
	3	0	100	0	0.011	0.050	0.099	0	100	0	0.009	0.050	0.101
$\gamma_1 = \gamma_2 = 0.5$													
100	1	0	100	0	0.008	0.047	0.098	0	98	2	0.011	0.052	0.102
	2	0	100	0	0.010	0.048	0.101	1	99	0	0.012	0.053	0.104
	3	0	100	0	0.009	0.047	0.095	0	100	0	0.009	0.049	0.096
500	1	0	100	0	0.009	0.048	0.098	0	100	0	0.011	0.051	0.100
	2	0	100	0	0.011	0.050	0.099	0	100	0	0.010	0.048	0.099
	3	0	100	0	0.008	0.048	0.097	0	100	0	0.011	0.053	0.102
$\gamma_1 = \gamma_2 = 0.2$													
100	1	0	61	49	0.007	0.048	0.097	0	49	51	0.005	0.033	0.071
	2	10	61	29	0.005	0.034	0.076	11	58	31	0.002	0.023	0.056
	3	20	61	19	0.005	0.031	0.070	20	60	20	0.002	0.021	0.053
500	1	0	99	1	0.012	0.054	0.103	0	94	6	0.013	0.055	0.105
	2	0	99	1	0.009	0.051	0.099	1	98	1	0.010	0.046	0.097
	3	0	99	1	0.010	0.049	0.098	1	98	1	0.009	0.045	0.095

Note: DGP implies eqn(1) where trig functions are present. The frequency is selected according to the SSR minimization criterion.



Proportion of incorrect frequency estimates as a function of the sample size and the parameter $\gamma(=\gamma 1=\gamma 2)$.

FIGURE 2. Convergence to correct frequency estimate $k = 1$.

bigger, the rejection frequencies increase. We note that the power of the test is quite good even when using the estimated frequency $\hat{k}$.

Lemma 1 shows that the usual KPSS-type stationarity tests will diverge when existing breaks are ignored. We now examine the magnitude of the size distortion in finite samples. In this experiment, the DGP implies eqn (3a) or (3b) such that nonlinearity exists in the data, but the usual KPSS tests are employed. The simulation results are shown in Table IV. While it is not surprising that the standard KPSS test is clearly oversized when the amplitude of the trigonometric terms is high, it is important to note the significant size distortion for very small trigonometric amplitudes or for a large sample size.

We now examine the performance of the three, previously discussed strategies for selecting k:

(1) Using a fixed value of $k = 1$ ($k1$).
(2) Using cumulative frequencies $k_1 = 1$ and $k_2 = 2$ ($k12$).
(3) Using the data-dependent frequency estimated by minimizing SSR (SSR).

We consider the DGP where the true frequency is $k = 0$, 1, 2 and 3. The case with $k = 0$ refers to the DGP where the nonlinear trend is absent. We report results on the size properties in Table V. The results in panels A, B and C correspond to each of the above three different testing strategies, $k1$, $k12$ and SSR, respectively. We consider different values of γ. When $k = 0$ or 1 in the DGP, it is evident that all strategies show decent size properties. However, $k1$ is over-sized when $k = 2$ or higher in the DGP. Similarly, $k12$ also leads to over-rejections of the null when $k = 3$ or higher. This result is rather obvious since these strategies use misspecified frequencies. On the other hand, the third

TABLE III

5% Rejection Frequencies of $\tau_\mu(k)$ and $\tau_\tau(k)$ Tests under the alternative

T	Case[a]	Value of σ_u^2 0.0001	0.001	0.01	0.1	1	100	10,000
$\tau_\mu(k)$-level								
100	1	0.053	0.097	0.364	0.785	0.962	0.988	0.987
	2	0.057	0.154	0.552	0.917	0.985	0.989	0.991
	3	0.063	0.163	0.582	0.929	0.990	0.995	0.995
	$\hat{k}$, $k = 1$	0.056	0.090	0.340	0.758	0.950	0.978	0.978
	$\hat{k}$, $k = 2$	0.059	0.155	0.499	0.817	0.953	0.981	0.981
	τ_{KPSS}	0.063	0.168	0.587	0.927	0.989	0.994	0.998
500	1	0.152	0.561	0.945	0.999	1.000	1.000	1.000
	2	0.289	0.767	0.981	0.999	1.000	1.000	1.000
	3	0.303	0.783	0.987	1.000	1.000	1.000	1.000
	$\hat{k}$, $k = 1$	0.157	0.563	0.931	0.999	1.000	1.000	1.000
	$\hat{k}$, $k = 2$	0.287	0.768	0.970	0.999	1.000	1.000	1.000
	τ_{KPSS}	0.307	0.788	0.997	1.000	1.000	1.000	1.000
$\tau_\tau(k)$-trend								
100	1	0.051	0.064	0.144	0.702	0.988	0.999	1.000
	2	0.058	0.080	0.311	0.788	0.966	0.987	0.987
	3	0.053	0.082	0.346	0.860	0.985	0.997	0.997
	$\hat{k}$, $k = 1$	0.055	0.064	0.139	0.611	0.955	0.988	0.987
	$\hat{k}$, $k = 2$	0.051	0.077	0.297	0.713	0.943	0.973	0.974
	τ_{KPSS}	0.054	0.084	0.352	0.878	0.993	0.999	0.999
500	1	0.072	0.329	0.954	1.000	1.000	1.000	1.000
	2	0.137	0.530	0.933	0.999	1.000	1.000	1.000
	3	0.137	0.607	0.976	1.000	1.000	1.000	1.000
	$\hat{k}$, $k = 1$	0.075	0.313	0.928	1.000	1.000	1.000	1.000
	$\hat{k}$, $k = 2$	0.127	0.521	0.931	0.998	1.000	1.000	1.000
	τ_{KPSS}	0.137	0.621	0.983	1.000	1.000	1.000	1.000

Note: [a] The case with $k = 1,2$, or 3 refers to the experiment where the true frequency is used in the estimation. The case with $[\hat{k}, k = 1]$ (or $[\hat{k}, k = 2]$) denotes the case where the DGP implies $k = 1$ (or 2) and the value of k is obtained by minimizing the SSR. The τ_{KPSS} case provides the results of the KPSS test. For these results, the DGP did not include a nonlinear trend.

TABLE IV

Rejection Frequencies for Standard KPSS Tests Ignoring Nonlinear Trends

		$\gamma_1 = \gamma_2 = 1.0$			$\gamma_1 = \gamma_2 = 0.5$			$\gamma_1 = \gamma_2 = 0.2$		
T	k	1%	5%	10%	1%	5%	10%	1%	5%	10%
Level KPSS										
100	1	1.000	1.000	1.000	0.784	0.951	0.986	0.089	0.255	0.386
	2	0.337	0.840	0.986	0.070	0.244	0.421	0.017	0.082	0.150
	3	0.027	0.180	0.360	0.019	0.092	0.175	0.012	0.056	0.117
500	1	1.000	1.000	1.000	1.000	1.000	1.000	0.740	0.918	0.965
	2	1.000	1.000	1.000	0.971	1.000	1.000	0.099	0.289	0.462
	3	1.000	1.000	1.000	0.334	0.753	0.951	0.034	0.126	0.227
Trend KPSS										
100	1	1.000	1.000	1.000	0.896	0.972	0.986	0.135	0.321	0.442
	2	1.000	1.000	1.000	0.297	0.762	0.894	0.021	0.114	0.221
	3	0.216	0.960	0.998	0.023	0.197	0.421	0.011	0.064	0.128
500	1	1.000	1.000	1.000	1.000	1.000	1.000	0.801	0.929	0.963
	2	1.000	1.000	1.000	1.000	1.000	1.000	0.331	0.713	0.853
	3	1.000	1.000	1.000	0.998	1.000	1.000	0.057	0.283	0.488

TABLE V

Size of the Tests using Different Testing Strategies

DGP[a]		Panel A[b] ($k = 1$)						Panel B[b] ($k_1 = 1$ and $k_2 = 2$)						Panel C[b] (SSR)					
		$T = 100$			$T = 500$			$T = 100$			$T = 500$			$T = 100$			$T = 500$		
Model	k	$\gamma =$ 0.2	$\gamma =$ 0.5	$\gamma =$ 1.0	$\gamma =$ 0.2	$\gamma =$ 0.5	$\gamma =$ 1.0	$\gamma =$ 0.2	$\gamma =$ 0.5	$\gamma =$ 1.0	$\gamma =$ 0.2	$\gamma =$ 0.5	$\gamma =$ 1.0	$\gamma =$ 0.2	$\gamma =$ 0.5	$\gamma =$ 1.0	$\gamma =$ 0.2	$\gamma =$ 0.5	$\gamma =$ 1.0
Level	0	0.048	0.051	0.051	0.049	0.054	0.051	0.049	0.050	0.051	0.050	0.054	0.048	0.017	0.017	0.018	0.015	0.018	0.017
	1	0.051	0.046	0.049	0.051	0.048	0.049	0.049	0.047	0.048	0.050	0.049	0.048	0.048	0.047	0.049	0.054	0.048	0.049
	2	0.171	0.823	1.00	0.800	1.00	1.00	0.050	0.051	0.052	0.049	0.051	0.051	0.033	0.048	0.052	0.051	0.050	0.051
	3	0.086	0.317	0.946	0.376	1.00	1.00	0.123	0.628	1.00	0.651	1.00	1.00	0.031	0.047	0.047	0.049	0.048	0.050
Trend	0	0.052	0.051	0.053	0.053	0.050	0.048	0.050	0.049	0.050	0.055	0.049	0.049	0.007	0.006	0.007	0.006	0.006	0.006
	1	0.050	0.049	0.051	0.050	0.051	0.050	0.051	0.049	0.052	0.051	0.050	0.051	0.033	0.052	0.051	0.055	0.051	0.050
	2	0.324	0.981	1.00	0.948	1.00	1.00	0.050	0.051	0.048	0.050	0.051	0.051	0.023	0.052	0.050	0.046	0.048	0.049
	3	0.160	0.891	1.00	0.831	1.00	1.00	0.323	0.982	1.00	0.950	1.00	1.00	0.021	0.049	0.051	0.045	0.053	0.050

Notes: [a]The DGP implies eqn (1) with the true frequency, $k = 0,1,2$ and 3. The case with $k = 0$ implies that the nonlinear trend is absent.
[b]We consider three testing strategies:
Panel A: The frequency $k = 1$ is used in the estimation regardless of different values of true frequencies in the DGP.
Panel B: The cumulative frequencies using both $k = 1$ and $k = 2$ are used in the estimation regardless of different values of true frequencies in the DGP.
Panel C: The frequency that minimizes the SSR is selected.

TABLE VI

POWER OF THE TESTS USING DIFFERENT TESTING STRATEGIES

DGP[a]		Panel A[b]($k = 1$)				Panel B[b]($k_1 = 1$ and $k_2 = 2$)				Panel C[b](SSR)			
		$T = 100$		$T = 500$		$T = 100$		$T = 500$		$T = 100$		$T = 500$	
Model	σ_u^2	$k = 0$	$k = 2$	$k = 0$	$k = 2$	$k = 0$	$k = 2$	$k = 0$	$k = 2$	$k = 0$	$k = 2$	$k = 0$	$k = 2$
Level	1e-4	0.053	0.817	0.158	1.000	0.054	0.052	0.108	0.115	0.022	0.059	0.116	0.2871
	1e-3	0.092	0.797	0.573	1.000	0.077	0.071	0.453	0.440	0.057	0.155	0.510	0.7675
	1e-2	0.355	0.741	0.936	0.994	0.256	0.260	0.862	0.865	0.320	0.499	0.920	0.9704
	1e-1	0.842	0.872	1.000	1.000	0.662	0.674	0.997	0.999	0.747	0.817	0.999	0.999
	1	0.985	0.972	1.000	1.000	0.912	0.909	1.000	1.000	0.948	0.953	1.000	1.000
	100	0.996	0.987	1.000	1.000	0.964	0.963	1.000	1.000	0.981	0.981	1.000	1.000
	10000	0.997	0.989	1.000	1.000	0.966	0.965	1.000	1.000	0.981	0.981	1.000	1.000
Trend	1e-4	0.109	0.980	0.067	1.000	0.052	0.056	0.060	0.063	0.007	0.051	0.017	0.127
	1e-3	0.110	0.980	0.335	1.000	0.051	0.050	0.194	0.202	0.011	0.077	0.180	0.521
	1e-2	0.241	0.960	0.959	0.998	0.094	0.098	0.892	0.881	0.060	0.297	0.870	0.931
	1e-1	0.784	0.915	1.000	1.000	0.519	0.505	1.000	1.000	0.549	0.713	0.997	0.998
	1	0.994	0.991	1.000	1.000	0.978	0.977	1.000	1.000	0.941	0.943	1.000	1.000
	100	1.000	0.999	1.000	1.000	0.999	1.000	1.000	1.000	0.973	0.973	1.000	1.000
	10000	1.000	0.999	1.000	1.000	0.999	0.999	1.000	1.000	0.974	0.974	1.000	1.000

Notes: [a]The DGP implies eqn(1) with corresponding values of σ_u^2 and the true frequency, $k = 0$ and 2 ($\gamma_1 = \gamma_2 = 0.5$). The case with $k = 0$ implies that the nonlinear trend is absent.
[b]We consider three testing strategies:
Panel A: The frequency $k = 1$ is used in the estimation regardless of different values of true frequencies in the DGP.
Panel B: The cumulative frequencies using both $k = 1$ and $k = 2$ are used in the estimation regardless of different values of true frequencies in the DGP.
Panel C: The frequency that minimizes the SSR is selected.

strategy SSR shows reasonable size in all cases. This is a result of the strategy searching for all frequencies up to 5. If this search region was reduced to $k^{max} = 2$, size distortions similar to those in the $k12$ strategy could be found. For small values of γ and T, there are moderate size distortions when the SSR criterion is used. This stems from the fact that frequency will not be precisely estimated in these cases. Moreover, we note that there is a size distortion in the case of $k = 0$. A by-product of the search procedure over the parameters defining the nonlinear trend is that the comparable test is conservative when the DGP under the null hypothesis does not have a nonlinear trend. The frequency that best fits any variations in the sample data will be selected and hence tilt the results towards accepting the null hypothesis.[9]

For our last simulation, we examine the power properties of these testing strategies for various values of σ_u^2. The results are reported in Table VI. For the case where no trigonometric term is present in the DGP ($k = 0$), the $k1$ strategy shows the largest power followed by the SSR and then the $k12$ strategy. This is a plausible result given that the test regression in the $k12$ strategy has the most parameters (in this case unnecessary) to be estimated. The SSR strategy has slightly worse power than the $k1$ strategy, as it involves a search procedure that will bias the procedure towards not rejecting the null hypothesis. The trade-off between the $k12$ strategy (or more generally the $k1n$ strategy using n cumulative frequencies) and the SSR strategy is as follows. The former will guarantee a correctly sized testing procedure even when no break is present in the data at all, whereas the latter will display a conservative size in that case, but will display higher power. A researcher will have to decide on which side of this trade-off he/she wants to be. As mentioned above, for the case $k = 2$ ($\gamma_1 = \gamma_2 = 0.5$) the $k1$ strategy does not provide inference with correct size. When comparing the $k12$ and SSR strategy it again transpires that the SSR procedure is preferable.[10] The reason for this is again the smaller number of parameters that need to be estimated. It was argued above that the $k12$ strategy could be easily extended to have a larger maximum frequency. From these results, it is clear that the price one will pay for this is a reduced power, whereas no such effect will arise when applying the SSR procedure. Overall, we recommend using the SSR strategy when there is no *a priori* information about the true frequency of the data.

4. REAL EXCHANGE RATES AND SMOOTH STRUCTURAL BREAKS

As judged by the number of papers on the topic, one of the more contentious issues in the economics literature concerns the validity of purchasing power parity (ppp) during the post-Bretton Woods period. At the time of writing this, Econlit contains 453 papers with the term 'purchasing power parity' in the title and 896 papers with 'ppp' in the title and/or the abstract. Part of the reason for the unresolved ambiguity is that real exchange rates are highly persistent and unit-root tests have low power over relatively short time spans. Another possibility is that real rates are subject to long swings, as in Engel and Hamilton (1990), or to regime switches. Moreover, Perron and Vogelsang (1992) added another dimension to the problem by considering the possibility that real exchange rates contain structural breaks. As such, it seems natural to test ppp under the null of stationarity allowing for unknown structural breaks.

In order to illustrate our test, we obtained quarterly values of the Canadian, Japanese and UK nominal exchange rates against the US dollar from the CD-ROM version of the *International Financial Statistics* over the 1973–2003 period. We also collected producer and consumer price indices, and for each country formed the real exchange rate as:

$$r_t = e_t + p_t - p_t^*$$

where r_t is the real exchange rate against the US dollar, p_t the US price level, p_t^* the foreign price level, e_t the foreign currency price of the US dollar, and all variables are expressed in natural logarithms such that $r_{1996:1} = 0$.

The time paths of the three real exchange rates constructed using the producer price indices are shown as the solid lines in the left-hand column of Figure 3. The

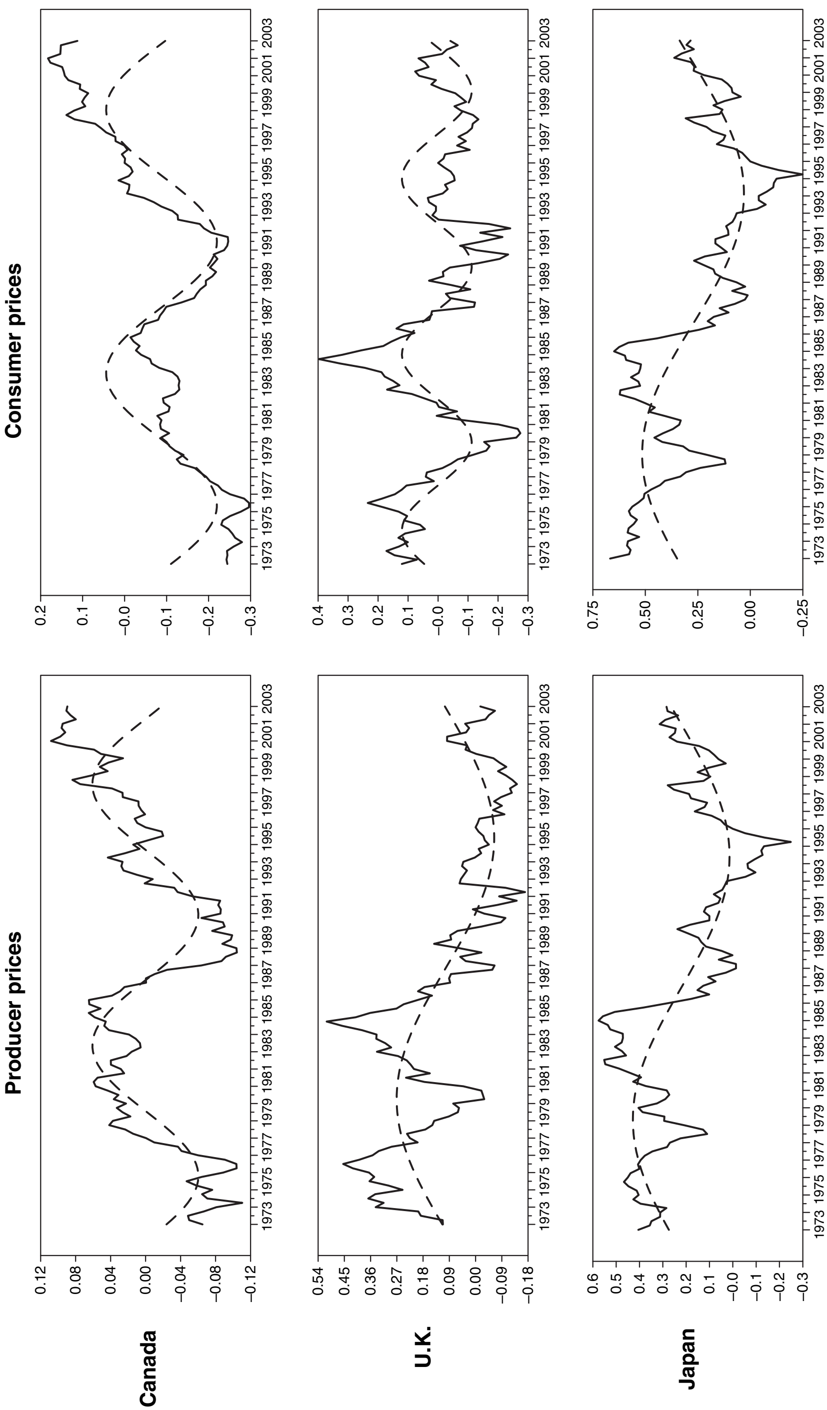


FIGURE 3. Logs of real exchange rates and fitted nonlinearities. (log 1996 = 0.0).

three exchange rates constructed using the consumer price indices (CPI) are shown as the solid lines in the right-hand column of Figure 3. Although all the real rates experienced a prolonged downward movement beginning 1985, it is not clear whether this is the result of a break or a unit root in the DGP. Moreover, 1976 might be the beginning of a break for the Canadian rates and 1979 might serve as the beginning of a break for the UK rates.

To compare our test with the standard linear unit-root tests, we first estimated an augmented DF equation for each real exchange rate series. Consider

$$\Delta y_t = \alpha_0 + \rho y_{t-1} + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + \varepsilon_t$$

We excluded 'time' as a regressor since any type of trend component is inconsistent with ppp. The estimated value of ρ, the lag length (p), and the t-statistic for the null hypothesis $\rho = 0$ are reported in the third to the fifth columns of Table VII. At the 5% significance level, it is possible to reject the null hypothesis of a unit root only for the UK/US real exchange rate constructed using CPI.

Next, we applied a standard KPSS test to each series. Since there is a substantial amount of persistence in each series, we experimented with various truncation lags when estimating the long-run variances. When we used an eight-quarter truncation lag for all series, the KPSS statistics reported in Table VII are such that the null hypothesis of stationarity is maintained for the Canadian/US rate constructed with producer prices and the UK/US real exchange rate constructed using CPI. The results are virtually identical when using the optimal truncation lag, as in Andrew's (1991), for each series, individually.

TABLE VII

Tests for Unit Roots and Stationarity in the Real Exchange Rate Series

	DF test				Nonlinear stationarity test		
	Lags	ρ	τ	KPSS (τ_{KPSS})	$\hat{k}$	$F_\mu(\hat{k})$	$\tau_\mu(\hat{k})$
Producer prices							
Canada	3	−0.052	−2.11	0.365*	2	70.01	0.196*
UK	7	−0.068	−2.20	0.921	1	61.55	0.127*
Japan	3	−0.054	−2.12	0.700	1	95.94	0.048*
Consumer prices							
Canada	3	−0.020	−1.64	0.935	2	66.52	0.958
UK	7	−0.138	−3.22*	0.281*	3	41.76	0.245*
Japan	3	−0.050	−2.21	0.837	1	90.59	0.073*

Note: 'Lags' is the number of lags selected for the augmented Dickey–Fuller test, ρ is the coefficient of interest in the DF test, τ is the DF t-statistic for the null hypothesis $\rho = 0$, and τ_{KPSS} is the value of the standard KPSS test for the null hypothesis of stationarity. At the 5% significance level, the critical value for τ is −2.89 and the critical value for KPSS is 0.463. Note that '*' denotes that purchasing power parity is supported at the 5% significance level.

The situation is quite different when we applied our nonlinear version of the stationarity test to the data. We report results using the grid-search method since we have no *a priori* notion concerning the number of nature of the breaks in the data. First, as in eqn (3a), we regressed each real rate on a constant, $\sin(2\pi kt/T)$ and $\cos(2\pi kt/T)$ for each integer frequency in the interval (1, 5). The frequencies resulting in the best fit $(\hat{k})$ are reported in Table VII. It is interesting to note that the sole value of $\hat{k}$ greater than 2 is that for the real UK/US rate constructed using (CPIs). However, this is the single case found to be stationary by the DF and the standard KPSS test.

The key issue, however, is to determine whether the various series are stationary. As such, we calculated $\tau_{\mu}(\hat{k})$ as in eqn (6) using the estimated frequencies. We used various methods to select the truncation lag for constructing the long-run variances. A common value of 8 worked best for the series as a group and the results are not very sensitive for the individual lags selected by Andrew's (1991) automatic bandwidth selection method. At the 5% significant level, Table Ia indicates that the appropriate critical values are 0.1720, 0.4152 and 0.4480 for the frequencies $k = 1$, 2, and 3, respectively. Comparing these values with the calculated sample values reported in Table VII indicates that the null hypothesis of stationarity can be rejected only for the Canadian/US rate constructed using consumer prices, but the null is not rejected for the other five series at the 5% significance level.

The F-statistics for the null hypothesis $\gamma_1 = \gamma_2 = 0$ are shown in the table. Since these coefficients depend on the nuisance parameter k, it is necessary to perform the test using the critical values reported in Table Ic rather than a traditional F table. Nevertheless, it is clear that the null hypothesis of linearity is rejected. The estimated time paths of the time-varying intercepts are shown by the dashed lines in Figure 3. An examination of the figure indicates that the all Fourier approximations seem reasonable and support the notion of long swings in real exchange rates.

5. SUMMARY AND CONCLUDING REMARKS

In many circumstances, it is natural to test the validity of a theory such that the theory itself forms the basis of the null hypothesis and its refutation forms the alternative. Standard unit-root tests, such as the DF test, have the null hypothesis of a unit root and the alternative hypothesis of stationarity. For this reason, KPSS-type tests have become popular for testing theories, such as purchasing power parity, that imply that a particular variable is stationary. In the recent time-series literature, several papers have extended the KPSS-type tests to allow for one sharp or smooth break. Nevertheless, many macroeconomic time-series variables are subject to an unknown number of breaks with unknown functional forms. We develop a stationarity test for such circumstances that relies on the fact that a single-frequency component of a

Fourier approximation can mimic a wide variety of breaks and other types of nonlinearities. Instead of selecting specific break dates, the number of breaks, and the form of the breaks, we suggest including a single-frequency component to include in the estimating equation. Since breaks shift the spectral density function towards zero, we suggest controlling for breaks using a frequency of 1 or 2 for data with high persistence.

When the most appropriate frequency component is unknown, we use a grid-search method over the low frequencies to find a consistent estimate of the frequency. The frequency component itself is treated as a nuisance parameter since our ultimate aim is to develop a test for stationarity. The asymptotic distribution for our test is derived and the empirical performance is such that the test does not exhibit any serious size distortions, and shows reasonable power. The appropriate use of the test is illustrated using real exchange rates in the post-Bretton Woods period. The theory of purchasing power parity implies that such rates should be stationary. However, a standard DF test allows us to reject a unit root for only one of the six rates considered. Instead, our KPSS-type test allowing for multiple breaks supports ppp for five of the six real exchange rates.

APPENDIX

Proof of Proposition 1

$$(a) \quad \frac{1}{T}\sum_{t=1}^{T}\sin\left(\frac{2\pi kj}{T}\right) \to \int_0^1 \sin(2\pi ka)\mathrm{d}a = \frac{1}{2\pi k}(1-\cos(2\pi k))$$

$$(b) \quad \frac{1}{T}\sum_{t=1}^{T}\cos\left(\frac{2\pi kj}{T}\right) \to \int_0^1 \cos(2\pi ka)\mathrm{d}a = \frac{\sin(2\pi k)}{2\pi k}$$

$$(c) \quad \frac{1}{T}\sum_{j=1}^{[rT]}\sin\left(\frac{2\pi kj}{T}\right) = \frac{r}{rT}\sum_{j=1}^{[rT]}\sin\left(\frac{2\pi kj}{T}\right) \to r\int_0^r \sin(2\pi ka)\mathrm{d}a = \frac{r}{2\pi k}(1-\cos(2\pi kr))$$

$$(d) \quad \frac{1}{T}\sum_{j=1}^{[rT]}\cos\left(\frac{2\pi kj}{T}\right) = \frac{r}{rT}\sum_{j=1}^{[rT]}\cos\left(\frac{2\pi kj}{T}\right) \to r\int_0^r \cos(2\pi ka)\mathrm{d}a = r\frac{\sin(2\pi kr)}{2\pi k}$$

$$(e)\quad \frac{1}{T^2}\sum_{t=1}^{T} t\cdot\sin\left(\frac{2\pi kj}{T}\right) \to \int_0^1 r\,\sin(2\pi kr)\mathrm{d}r = \frac{1}{(2\pi k)^2}\sin(2\pi k) - \frac{1}{2\pi k}\cos(2\pi k)$$

$$(f)\quad \frac{1}{T^2}\sum_{t=1}^{T} t\cdot\cos\left(\frac{2\pi kj}{T}\right) \to \int_0^1 r\cos(2\pi kr)\mathrm{d}r = \frac{1}{(2\pi k)^2}[\cos(2\pi k) + 2\pi k\,\sin(2\pi k) - 1]$$

$$(g)\quad \frac{1}{T}\sum_{t=1}^{T}\sin^2\left(\frac{2\pi kt}{T}\right) \to \int_0^1 \sin^2(2\pi kr)\mathrm{d}r = \frac{1}{2}\int_0^1(1-\cos(4\pi kr))\mathrm{d}r = \frac{1}{2} - \frac{\sin(4\pi k)}{4\pi k}$$

$$(h)\quad \frac{1}{T}\sum_{t=1}^{T}\cos^2\left(\frac{2\pi kt}{T}\right) \to \int_0^1 \cos^2(2\pi kr)\mathrm{d}r = \int_0^1(1-\sin^2(2\pi kr))\mathrm{d}r = \frac{1}{2} + \frac{\sin(4\pi k)}{4\pi k}.$$

The results for (i) and (j) are standard. For the proofs for (k) and (l), we employ the result in Bierens (1994, Lemma 9.6.3):

$$\sum_{t=2}^{T} F\left(\frac{t}{T}\right)u_t = F(1)S_T(1) - \int_0^1 f(r)S_T(r)\mathrm{d}r,$$

where $f(r)$ is $F'(r)$. To obtain the result for (k), we use $F(x) = \cos(2\pi kt/T)$. Then, it can be shown that

$$F(1)S_T(1) - \int_0^1 f(r)S_T(r)\mathrm{d}r = \sigma[W(1) + (2\pi k)\int_0^1 \sin(2\pi kr)W(r)\mathrm{d}r].$$

For the result in eqn (l), we choose $F(x) = \sin(2\pi kt/T)$, and the desired result is obtained. □

Proof of Theorem 1. We first consider the level stationarity test with a Fourier function. As in the text, we consider $\tau_\mu(k)$, where $\tilde{e}_t$ are the OLS residuals from the regression (3a) with $X_t = [1,\ \sin(2\pi kt/T),\ \cos(2\pi kt/T)]'$. We examine

$$\begin{aligned}\frac{1}{\sqrt{T}}\tilde{S}_{[rT]} &= \frac{1}{\sqrt{T}}\sum_{j=1}^{[rT]}\tilde{e}_j = \frac{1}{\sqrt{T}}\sum_{j=1}^{[rT]}[e_j - X_j'D_T[D_TX'XD_T]^{-1}D_TX'e] \\ &= \frac{1}{\sqrt{T}}S_{[rT]} - \frac{1}{\sqrt{T}}\sum_{j=1}^{[rT]}X_j'D_T\cdot[D_TX'XD_T]^{-1}\cdot D_TX'e \qquad (A.1)\end{aligned}$$

where

$$X = (X_1,\ldots,X_T)',\quad e = (e_1,\ldots,e_T)',\quad D_T = \mathrm{diag}\left[\frac{1}{\sqrt{T}},\frac{1}{\sqrt{T}},\frac{1}{\sqrt{T}}\right]\quad\text{and}\quad S_{[rT]} = \sum_{j=1}^{[rT]}e_j.$$

We can show

$$\frac{1}{\sqrt{T}}\sum_{j=1}^{[rT]} X_j' D_T = \left[r, \frac{1}{T}\sum_{j=1}^{[rT]} \sin\left(\frac{2\pi kj}{T}\right), \frac{1}{T}\sum_{j=1}^{[rT]} \cos\left(\frac{2\pi kj}{T}\right)\right]' \to [r, s_r, c_r]'.$$

$$[D_T X' X D_T]^{-1} = \begin{bmatrix} T/T & T^{-1}\sum\sin(2\pi kt/T) & T^{-1}\sum\cos(2\pi kt/T) \\ & T^{-1}\sum\sin^2(2\pi kt/T) & 0 \\ & & T^{-1}\sum\cos^2(2\pi kt/T) \end{bmatrix}^{-1}$$

$$\to \begin{bmatrix} 1 & s_0 & c_0 \\ & s_2 & 0 \\ & & c_2 \end{bmatrix}^{-1}$$

$$D_T X' e = \left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T} e_t, \frac{1}{\sqrt{T}}\sum_{t=1}^{T} e_t \sin\left(\frac{2\pi kt}{T}\right), \frac{1}{\sqrt{T}}\sum_{t=1}^{T} e_t \cos\left(\frac{2\pi kt}{T}\right)\right]'$$
$$\to \sigma[f_1, f_3, f_4]'.$$

Then, after a tedious algebra, we can show that the second term in (A.1) follows

$$\begin{aligned} &\frac{1}{\sqrt{T}}\sum_{j=1}^{[rT]} X_j' D_T \cdot [D_T X' X D_T]^{-1} \cdot D_T X' e \\ &\to [-c_0 c_r s_0 f_3 + r c_2 s_0 f_3 + c_r s_0^2 f_1 + c_0 c_r s_2 f_1 - c_r s_2 f_4 - r c_2 s_2 f_1 + r c_0 s_2 f_4 + c_0^2 s_r f_3 \\ &\quad - c_2 s_r f_3 + c_2 s_0 s_r f_1 - c_0 s_0 s_r f_4]/[c_2 s_0^2 + c_0^2 s_2 - c_2 s_2] \\ &\equiv \sigma E_\mu(k, r). \end{aligned}$$

Therefore, the expression in eqn (A.1) follows

$$\frac{1}{\sqrt{T}}\tilde{S}_{[rT]} \to \sigma[W(r) - E_\mu(k, r)] \equiv \underline{V}_\mu(r). \tag{A.2}$$

Then, it is easy to show that the numerator of $\tau_\mu(k)$ follows

$$\frac{1}{T^2}\sum_{t=1}^{T}\tilde{S}_{[rT]}^2 \to \sigma^2 \int_0^1 \underline{V}_\mu(r)^2 \mathrm{d}r.$$

As the denominator of $\tau_\mu(k)$ is a consistent estimator of σ^2, the desired result is obtained.

Next, we consider the trend stationarity test with a Fourier function, $\tau_\tau(k)$, where $(\tilde{e}_t)$ are the OLS residuals from the regression (3b) with $X_t = [1,\ t,\ \sin(2\pi kt/T),\ \cos(2\pi kt/T)]'$. We again examine the similar expression as in eqn (A.1) with $D_T = \mathrm{diag}[1/\sqrt{T}, 1/T^{1.5}, 1/\sqrt{T}, 1/\sqrt{T}]$

$$\frac{1}{\sqrt{T}}\tilde{S}_{[rT]} = \frac{1}{\sqrt{T}}S_{[rT]} - \frac{1}{\sqrt{T}}\sum_{j=1}^{[rT]} X_j' D_T \cdot [D_T X' X D_T]^{-1} \cdot D_T X' e. \tag{A.3}$$

We can show

$$\frac{1}{\sqrt{T}}\sum_{j=1}^{[rT]} X_j' D_T = \left[r, \frac{1}{T^2}\frac{1}{2}(rT)(rT-1), \frac{1}{T}\sum_{j=1}^{[rT]}\sin\left(\frac{2\pi kj}{T}\right), \frac{1}{T}\sum_{j=1}^{[rT]}\cos\left(\frac{2\pi kj}{T}\right)\right]'.$$

$$\rightarrow \left[r, \frac{1}{2}r^2, s_r, c_r\right]'.$$

$$[D_T X' X D_T]^{-1} = \begin{bmatrix} T/T & T^{-1}\sum t & T^{-1}\sum \sin(2\pi kt/T) & T^{-1}\sum \cos(2\pi kt/T) \\ & T^{-3}\sum t^2 & T^{-2}\sum t\sin(2\pi kt/T) & T^{-2}\sum t\cos(2\pi kt/T) \\ & & T^{-1}\sum \sin^2(2\pi kt/T) & 0 \\ & & & T^{-1}\sum \cos^2(2\pi kt/T) \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & 1/2 & s_0 & c_0 \\ & 1/3 & s_1 & c_1 \\ & & s_2 & 0 \\ & & & c_2 \end{bmatrix}^{-1}$$

$$D_T X' e = \left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T} e_t, \frac{1}{T^{1.5}}\sum_{t=1}^{T} te_t, \frac{1}{\sqrt{T}}\sum_{t=1}^{T} e_t \, \sin\left(\frac{2\pi kt}{T}\right), \frac{1}{\sqrt{T}}\sum_{t=1}^{T} e_t \, \cos\left(\frac{2\pi kt}{T}\right)\right]'$$

$$\rightarrow \sigma[f_1, f_2, f_3, f_4]'.$$

Then, after a tedious algebra, we can show that the second term in eqn (A.3) follows

$$\frac{1}{\sqrt{T}}\sum_{j=1}^{[rT]} X_j' D_T \cdot [D_T X' X D_T]^{-1} \cdot D_T X' e$$
$$\rightarrow [-12c_1c_rs_1f_3 - 6rc_2s_1f_3 + 6r^2c_2s_1f_3 + 12c_rs_1^2f_4 + 12rc_2s_1^2f_1 - 6c_1c_rs_2f_1 + 12c_1c_rs_2f_2$$
$$- c_rs_2f_4 + 12rc_1^2s_2f_1 - 4rc_2s_2f_1 + 6rc_2s_2f_2 - 6rc_1s_2f_4 + 3r^2c_2s_2f_1 + 6r^2c_2s_2f_2 + 6r^2c_1s_2f_4$$
$$+ 12c_1^2s_rf_3 - c_2s_rf_3 - 6c_1s_1s_rf_1 + 12c_2s_1s_rf_2 - 12c_1s_1s_rf_4]/[12c_2s_1^2 + 12c_1^2s_2 - c_2s_2]$$
$$\equiv \sigma E_\tau(k, r).$$

We note in passing that we utilized the fact that $s_0 = c_0 = 0$ for integer k, to simplify the expression above. For integer values of k, the above expression can be further simplified by noting that

$$[D_T X' X D_T]^{-1} \rightarrow \begin{pmatrix} 1 & 1/2 & 0 & 0 \\ & 1/3 & -1/(2\pi k) & 0 \\ & & 1/2 & 0 \\ & & & 1/2 \end{pmatrix}^{-1}.$$

Therefore, the expression in eqn (A.3) follows

$$\frac{1}{\sqrt{T}}\tilde{S}_{[rT]} \rightarrow \sigma[W(r) - E_\tau(k, r)] \equiv \underline{V}_\tau(r). \tag{A.4}$$

Then, similarly,

$$\frac{1}{T^2}\sum_{t=1}^{T}\tilde{S}_{[rT]}^2 \rightarrow \sigma^2\int_0^1 \underline{V}_\tau(r)^2 \mathrm{d}r.$$

Accordingly, the desired result follows. □

Proof of Lemma 1. We begin with the case of the level stationarity test. The DGP implies eqn (1) with $\sigma_u^2 = 0$, X_t=[1], and $Z_t = [\sin(2\pi kt/T), \ \cos(2\pi kt/T)]'$ such that a Fourier nonlinear term is present, but Z_t is ignored in the testing regressions where the usual KPSS regression is used.

$$y_t = \alpha + e_t. \tag{A.5}$$

We let $\hat{e}_t$ be the residuals from this regression. Then, it can be shown that

$$\hat{e}_t = e_t - X_t'(X'X)^{-1}X'e + Z_t'\gamma - X_t'(X'X)^{-1}X'Z\gamma$$

where $Z = (Z_1, \ldots, Z_T)'$. Then, we have

$$\frac{1}{\sqrt{T}}\hat{S}_{[rT]} = \frac{1}{\sqrt{T}}\sum_{t=1}^{[rT]}[e_t - X_t'(X'X)^{-1}X'e] + \frac{1}{\sqrt{T}}\sum_{t=1}^{[rT]}Z_t'\gamma - \frac{1}{\sqrt{T}}\sum_{t=1}^{[rT]}X_t'(X'X)^{-1}X'Z\gamma. \tag{A.6}$$

The first term follows: $\sigma[W(r) - rW(1)]$, which is shown in Kwaitowski *et al*. (1992). The second term follows from Proposition 1c,

$$\sqrt{T}\frac{1}{T}\sum_{t=1}^{[rT]}\left[\gamma_1 \sin\left(\frac{2\pi kt}{T}\right) + \gamma_2 \cos\left(\frac{2\pi kt}{T}\right)\right] = \sqrt{T}\frac{r}{2\pi k}(\gamma_1(1-\cos(2\pi kr)) + \gamma_2 \ \sin(2\pi kr))$$
$$= O_p(\sqrt{T}).$$

The third term in eqn (A.6) is shown to follow

$$-\sqrt{T}r\frac{1}{T}\sum_{t=1}^{T}\left(\gamma_1 \sin\left(\frac{2\pi kt}{T}\right) + \gamma_2 \cos\left(\frac{2\pi kt}{T}\right)\right) = -\sqrt{T}r(\gamma_1 s_0 + \gamma_2 c_0) = O_p(\sqrt{T}).$$

The third term disappears for integer frequencies k, but the second remains $O_p(\sqrt{T})$. Thus, $(1/\sqrt{T})\hat{S}_{[rT]} = O_p(\sqrt{T})$. Then, the numerator of the KPSS is shown to diverge

$$\frac{1}{T^2}\sum_{t=1}^{T}\hat{S}_t^2 = O_p(T).$$

Therefore, the KPSS-level statistic ignoring existing unattended nonlinearity will also diverge as the sample size increases.

Next, we examine the case of the trend stationarity test. The DGP implies eqn (1) with $X_t = [1, \ t]$, and $Z_t = [\sin(2\pi kt/T), \ \cos(2\pi kt/T)]'$ such that a Fourier nonlinear term is present, but Z_t is ignored in the testing regressions where the usual KPSS regression is used.

$$y_t = \alpha + \xi t + e_t. \tag{A.7}$$

Letting $\hat{e}_t$ be the residuals from this regression, we have

$$\frac{1}{\sqrt{T}}\hat{S}_{[rT]} = \frac{1}{\sqrt{T}}\sum_{t=1}^{[rT]}[e_t - X_t'(X'X)^{-1}X'e] + \frac{1}{\sqrt{T}}\sum_{t=1}^{[rT]}Z_t'\gamma - \frac{1}{\sqrt{T}}\sum_{t=1}^{[rT]}X_t'R(RX'XR)^{-1}RX'Z\gamma \tag{A.8}$$

where the scaling matrix is given as $R = \mathrm{diag}[1/\sqrt{T}, 1/T^{1.5}]$. The first term approaches the second-level standard Brownian bridge, which is shown in Kwaitowski *et al*. (1992). The second term is $O_p(\sqrt{T})$, as in the case with the level stationary test. For the third term in eqn (A.8), we note

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{[rT]} X_t' R \to \left[r, \frac{1}{2}r^2\right] \quad \text{and} \quad (RX'XR)^{-1} \to \begin{bmatrix} 4 & -6 \\ -6 & 12 \end{bmatrix}.$$

And, we divide the last product term by $\sqrt{T}$ to obtain

$$\frac{1}{\sqrt{T}} RX'\mathrm{Z}\gamma = \begin{pmatrix} \frac{1}{T}\sum_{t=1}^{T}(\gamma_1 \ \sin(2\pi kt/T) + \gamma_2 \ \cos(2\pi kt/T)) \\ \frac{1}{T^2}\sum_{t=1}^{T}(\gamma_1 t \cdot \sin(2\pi kt/T) + \gamma_2 \ \cos(2\pi kt/T)) \end{pmatrix} \to \begin{bmatrix} (\gamma_1 s_0 + \gamma_2 c_0) \\ (\gamma_1 s_1 + \gamma_2 c_1) \end{bmatrix}.$$

Therefore, the third term in eqn (A.8) is $O_p(\sqrt{T})$. Thus, $(\frac{1}{\sqrt{T}})\hat{S}_{[rT]} = O_p(\sqrt{T})$. Then, the numerator of the KPSS diverges, and accordingly, the KPSS trend statistic ignoring existing unattended nonlinearity will also diverge as the sample size increases. □

NOTES

1. We use the term 'unit-root test' to refer to any test with the null hypothesis of a unit root and an alternative of stationarity. Similarly, we use 'stationarity test' to describe a test with the null hypothesis of stationarity and an alternative of a unit root.
2. In fact, it is difficult to distinguish between a structural break and certain types of nonlinearities. Clearly, a series with a break can be viewed as a special case of a process that is nonlinear in its parameters. As such, our approach can be viewed as an attempt to provide a general procedure to approximate unknown nonlinear components. Nevertheless, the discussion in the text focuses on structural breaks.
3. Throughout the paper, '$\to$' indicates weak convergence as T approaches ∞.
4. Critical values with $T = 5000$ were generated but show only negligible variation to those at $T = 1000$.
5. Readers familiar with Davies (1987) will recognize that he develops a test for the unobserved frequency component $Z_t = \gamma_1 \sin(2\pi kt/T) + \gamma_2 \cos(2\pi kt/T)$ with i.i.d. errors.
6. A similar issue arises in Bai and Perron (1998). The asymptotic distribution of the break-point estimates can be calculated only on the assumption that the data are stationary. In particular, they showed that the break-point estimates are consistent, depending on the magnitude of the breaks. However, no theory is readily available for their break-point estimates when the order of integration is unknown.
7. We are grateful to an anonymous referee who correctly suggested this procedure.

8. We have conducted additional simulations on the performance of the F-statistic testing for linearity. Simulation results show that the null of linearity is rejected almost 100% of the time when $\gamma_1 = \gamma_2 = 1.0$ or 0.5, when the true DGP is stationary. As expected, when the magnitude of the nonlinear trend is mild ($\gamma_1 = \gamma_2 = 0.2$), the null of linearity is rejected less frequently for about 68% of replications for the level test $\tau_\mu(k)$ and 61% of replications for the trend test $\tau_\tau(k)$. Thus, the F-test performs reasonably well. When the DGP implies that the data are nonstationary (as we examine the power of the stationarity test), the null of linearity is rejected almost 100% of the time in replications when $\gamma_1 = \gamma_2 = 0.2$. As previously noted, the F-test will exhibit non-trivial rejections of the null of linearity. Finally, note that is possible to bootstrap the F-statistic for any particular sample. Details of the methodology and the performance of the bootstrapped statistic are available from the authors upon request.
9. The same effect leads to the over-rejection of the null hypothesis of no structural break when the maximum Chow test for a one-time structural break (the Chow test evaluated at that break point that maximizes the test statistic) is evaluated. In that case the nonlinear function is part of the alternative hypothesis and hence the search effect tilts the results towards the alternative hypothesis, whereas in this paper the nonlinear function is part of the null hypothesis and hence the effect tilts towards the null hypothesis.
10. The ($k = 2$) results in panel C are identical to those in Table III for [$\hat{k}$, $k = 2$].

Corresponding authors: Walter Enders is Professor and Bidgood Chair of Economics and Finance, Department of Economics, Finance & Legal Studies, University of Alabama, Tuscaloosa, AL 35487-0224, USA. Tel.: (205) 348-8972; Fax: (205) 348-0590; E-mail: wenders@cba.ua.edu

Ralf Becker is Lecturer, Economics, School of Social Sciences, University of Manchester, Manchester, UK. E-mail: ralf.becker@manchester.ac.uk

Junsoo Lee is Associate Professor, Department of Economics, Finance & Legal Studies, University of Alabama Tuscaloosa, AL, USA. E-mail: jlee@cba.ua.edu

REFERENCES

Andrews, D. W. K. (1991) Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 59, 817–58.

Bai, J. and Perron, P. (1998) Estimating and testing linear models with multiple structural changes. *Econometrica* 66, 47–78.

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

Bierens, H. (1994) *Topics in Advanced Econometrics: Estimation, Testing, and Specification of Cross-section and Time Series Models*. New York: Cambridge University Press.

Bierens, H. (1997) Testing for a unit root with drift hypothesis against nonlinear trend stationarity, with an application to the US price level and interest rate. *Journal of Econometrics* 81, 29–64.

Busetti, F. and Harvey, A. (2001) Testing for the presence of a random walk in series with structural breaks. *Journal of Time Series Analysis* 22, 127–50.

Busetti, F. and Taylor, R. (2003) Testing against stochastic trend and seasonality in the presence of unattended breaks and unit roots. *Journal of Econometrics* 117, 21–53.

Clemente, J., Montanes, A. and Reyes, M. (1998) Testing for a unit root in variables with a double change in the mean. *Economics Letters* 59, 175–82.

Davies, R. B. (1987) Hypothesis testing when a nuisance parameter is only identified under the alternative. *Biometrika* 47, 33–43.

Enders, W. and Lee, J. (2005) *Testing for a Unit-root with a Nonlinear Fourier Function*. Mimeo, University of Alabama.

Engel, C. and Hamilton, J. (1990) Long swings in the dollar: are they in the data and do markets know it? *American Economic Review* 80, 689–713.

Gallant, R. (1984) The Fourier flexible form. *American Journal of Agricultural Economics* 66, 204–8.

Gallant, R. and Souza, G. (1991) On the asymptotic normality of Fourier flexible form estimates. *Journal of Econometrics* 50, 329–53.

Harvey, D. and Mills, T (2003) A note on Busetti-Harvey tests for stationarity in series with structural breaks. *Journal of Time Series Analysis* 24, 159–64.

Harvey, D. and Mills, T. (2004) Tests for stationarity in series with endogenously determined structural change. *Oxford Bulletin of Economics and Statistics* 66, 863–94.

Kapetanios, G., Shin, Y., and Snell, A. (2003) Testing for a unit root in the nonlinear STAR framework. *Journal of Econometrics* 112, 359–79.

Kurozumi, E. (2002) Testing for stationarity with a break. *Journal of Econometrics* 108, 63–99.

Kwaitowski, D., Phillips, P. C. B., Schmidt, P. and Shin, Y. (1992) Testing the null hypothesis of stationarity against the null hypothesis of a unit root. *Journal of Econometrics* 54, 159–78.

Lee, J. and Strazicich, M. (2003), Minimum LM unit root tests with two structural breaks, Review of Economics and Statistics 85, 1082–9.

Lee, J., Huang, C. and Shin, Y. (1997) On stationarity tests in the presence of structural breaks. *Economics Letters* 55, 165–72.

Leybourne, S., Newbold, P. and Vougas, D. (1998) Unit roots and smooth transitions. *Journal of Time Series Analysis* 19, 83–97.

Luukkonen, R., Saikkonen, P. and Teresvirta, T. (1988) Testing linearity against smooth transition autoregressive models. *Biometrika* 75, 491–9.

Perron, P. (1989) The Great Crash, the Oil Price Shock, and the unit root hypothesis. *Econometrica* 57, 1361–401.

Perron, P. and Vogelsang, T. (1992) Nonstationarity and level shifts with an application to Purchasing power parity. *Journal of Business and Economics Statistics* 10, 301–20.

Presno, M. and Lopez, A. (2003) Response surface estimates of stationarity tests with a structural break. *Economic Letters* 78, 395–9.

Sen, A. (2003) On unit-root tests when the alternative is a trend-break stationary process. *Journal of Business and Economic Statistics* 21, 174–84.

Vogelsang, T. and Perron, P. (1998) Additional tests for a unit root allowing for a break in the trend function at an unknown time. *International Economic Review* 39, 1073–100.