Testing for a unit root with a nonlinear Fourier function

Walter Enders

and

Junsoo Lee*

April 18, 2004

Abstract

The paper develops a unit-root test that allows for an unknown number of structural breaks with unknown functional forms. The test is based on the fact that the behavior of such series can often be captured using a single frequency component of a Fourier approximation. 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 include in the estimating equation. Our proposed test does not exhibit any serious size distortions, and shows decent power. The appropriate use of the test is illustrated using real GDP and the interest rate differential.

Keywords: Structural breaks, nonlinear models, Fourier approximation.

JEL Classifications: C12, C22, E17

* Corresponding author

Authors: Walter Enders, Professor and Bidgood Chair of Economics and Finance, Department of Economics, Finance & Legal Studies, University of Alabama, Tuscaloosa, AL 35487-0224, Phone (205) 348-8972, Fax (205) 348-0590, Email wenders@cba.ua.edu; Junsoo Lee (corresponding author), Associate Professor, the same address, Phone (205) 348-8978, Email jlee@cba.ua.edu.

The authors wish to thank Ralf Becker, Mark Wohar, Kyung-so Im, Michael McCracken, and seminar participants at the Twelfth Annual Symposium of the Society for Nonlinear Dynamics and Econometrics at the Federal Reserve Bank in Atlanta, the University of Alabama, and the Georgia Tech University, for their helpful comments.

1. Introduction

Perron's (1989) seminal paper illustrates the problem of ignoring a structural break when testing the null hypothesis of a unit root. If the break date is known, Perron (1989) shows how to modify the standard Dickey-Fuller (DF) test by including dummy variables to capture the changes in the level and trend. Since the break date is often unknown, a number of papers allow the break date to be estimated along with the other parameters of the model[1]. Note that the existing unit-root literature assumes, *a priori,* the presence of one or two structural breaks in the level and/or the trend of the series in question. In principle, it is possible to have more than two breaks, but no such unit-root tests are readily available; it is cumbersome to obtain asymptotic distributions and critical values for different combinations of breaks when the number of breaks is more than two. Nevertheless, the performance of the existing unit-root tests will critically hinge on the accuracy of estimated break locations and the assumed number of breaks. The issue is important since, in applied work, both the break dates and the number of breaks are likely to be unknown.

Moreover, structural breaks are assumed to occur instantaneously or manifest themselves contemporaneously, taking abrupt jumps in the mean or immediate changes in the slope. This assumption may not be realistic in many cases. For instance, as discussed in section 5, it is clear that the full impact of the oil price shock on real macroeconomic variables did not occur immediately. As such, it would be desirable to consider tests for a unit root allowing for breaks such that the deterministic component of the model is a smooth transition process. In that regard, Leybourne, Newbold and Vougas (1998) examined the procedures to test for a unit root in the

[1] Recent works include Clemente, Montanes and Reyes (1998), Vogelsang and Perron (1998), Sen (1993), and Lee and Strazicich (2003), among others.

presence of a gradual structural change; see also Kepetanios, Shin and Snell (2003). In their paper, however, only one break can be allowed for and the break date needs to be specified.

The aim of this paper is to develop a unit-root test that allows for an unknown number of structural breaks with unknown functional forms. Specifically, we attempt to approximate unknown functional forms of nonlinearity by using a Fourier function, which is a linear combination of sine and cosine functions. Although our test can detect sharp breaks, it is designed to work best when breaks are gradual. If changes are instantaneous or abrupt, the traditional approach of using dummy variables to capture structural changes may be more appropriate. However, if structural breaks are smooth, our approximation can work better. One important feature of our approximation method is that we do not need to assume that the dates of structural changes and the number of structural changes are known *a priori*. Our goal is to control for the effect of unknown forms of nonlinear deterministic terms in testing for a unit root.

We illustrate that a series containing multiple structural breaks with unknown functional forms can often be captured using a single frequency component of a Fourier approximation. Hence, instead of selecting specific break dates, the number of breaks, and the form of the breaks, the specification problem is transformed into incorporating a frequency component into the estimating equation. We also show that our proposed unit-root test does not exhibit any serious size distortions, and shows decent power. Throughout the paper, "$\rightarrow$" indicates weak convergence as $T \rightarrow \infty$.

2. Approximating a nonlinear trend with a Fourier series

A simple modification of the Dickey-Fuller (DF) type test is to allow the intercept to be a time-dependent function denoted by $\alpha(t)$ so that

$$y_t = \alpha(t) + \beta y_{t-1} + \gamma \cdot t + \varepsilon_t \tag{1}$$

where ε_t is a stationary disturbance with variance σ^2 and $\alpha(t)$ is a deterministic function of t. The key feature of (1) is the form of $\alpha(t)$. In general, the functional form of $\alpha(t)$ is unknown so that it is not possible to estimate (1) directly. However, regardless of the actual form of $\alpha(t)$, for any desired level of accuracy, it is possible to write:

$$\alpha(t) = \alpha_0 + \sum_{k=1}^{n} \alpha_k \sin(2\pi kt/T) + \sum_{k=1}^{n} \beta_k \cos(2\pi kt/T); \quad n < T/2 \tag{2}$$

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

Note that equation (2) underlies the approach adopted by Bierens (1997) who suggests using Chebishev polynomials to approximate a nonlinear deterministic trend. Busetti and Harvey (2003) adopt a similar approach for seasonality tests. However, the use of many frequency components can lead to an over-fitting problem, and it seems difficult to obtain valid tests to determine the number of frequencies. To keep the problem tractable, we consider a Fourier approximation using a *single* frequency component, so that

$$\alpha(t) \cong \alpha_0 + a_1 \sin(2\pi kt/T) + a_2 \cos(2\pi kt/T) \tag{3}$$

where: k represents the single frequency selected for the approximation, and a_1 and a_2 measure the amplitude and displacement of the sinusoidal component of the deterministic term. As evidenced by papers such as Gallant (1984), Davies (1987), Gallant and Souza (1991), and Becker, Enders and Hurn (2004), a Fourier approximation using a single frequency component is shown to capture successfully the behavior of an unknown functional form.

In the absence of a nonlinear trend, $a_1 = a_2 = 0$ so that the standard Dickey-Fuller specification emerges as a special case. However, if there is a break or nonlinear trend, at least one Fourier frequency must be present in the data generating process. Thus, instead of positing the specific form of $\alpha(t)$, the issue is to select the proper frequency to include in (3). One key issue for our test is whether a single frequency can mimic the types of breaks typically seen in economic data. Although (3) is especially suitable to mimicking smooth breaks, the solid lines in Panels 1 to 6 of Figure 1 show the six series ($T = 60$) containing the sharp structural breaks used in Clements and Hendry (1999) and Becker, Enders and Hurn (2004). Panels 1 and 2 illustrate the effects of shifting a single break towards the end of the data set. Panels 3 and 4 allow for two breaks (or what might also be called a temporary break) and Panels 5 and 6 depict two distinct breaks. Of course, the essential features of all six series are invariant to inverting their magnitudes or to reordering the data from $t = 60$ to $t = 1$.

The dashed line (short dashes) in Panel 1, shows the time path of $\alpha(t)$ obtained by setting $k = 1$, $a_0 = 1.167$, $a_1 = -0.231$, and $a_2 = 0.150$. The values of a_0, a_1 and a_2 were selected by regressing y_t on $\alpha(t)$ 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 ($SSR = 1.05$). In contrast, if we use only an intercept term, $SSR = 3.33$. As noted by Davies (1987), the fit of a sinusoidal function can be sometimes be improved by using fractional frequencies. Towards this end, we preformed grid search for the best fitting frequency in the interval 1/512 to 5 by steps of 1/512. The frequency with the best fit ($k = 0.645$) resulted in $SSR = 0.585$. The other dashed line (long dashes) shows the time path of $\alpha(t)$ for $k = 0.645$, $a_0 = 1.25$, $a_1 = -0.291$, and $a_2 = -0.131$.

For our purposes, the precise parameter values for the other panels of the figure are not especially important. The key points illustrated by the six panels are:

1. A Fourier approximation of a structural break using a single frequency can often mimic the pattern in the data reasonably well. It is well known that breaks shift the spectral density function towards frequency zero. As such, the selection of the most appropriate frequency to mimic the breaks can occur at the low end of the spectrum.
2. The Fourier approximation does not require that the pattern of the break be symmetric. The fitted values shown in Panels 1 and 3 are identical. If we were to replace break 2 with $y_t = 1 \bullet (t \leq 45) + 1.5 \bullet (t > 45)$, the values of the *SSR* in Panels 2 and 4 would also be identical. This is especially interesting since some tests for breaks, such as the Bai and Perron (1998) test, do find such u-shaped breaks especially well.
3. Integer frequencies require that the starting and ending values of the series be equal. A shown by Davis (1987), a fractional frequency can often capture the effects of a break occurring near the beginning or end of the data.

Becker, Enders and Hurn (2004) show that a simple test based on the procedure outlined above can have reasonable power. Using a 5% significance level, the power of their test using a single Fourier frequency is 48.5%, 29.4%, 40.8%, 28.4%, 87.7% and 17.1% for the series shown in Panels 1 through 6, respectively. Moreover, they show that if the number of breaks is unknown, such a test can have better power than the Bai-Perron (1998) test.

Similarly, a Fourier approximation with a single frequency can capture some of the essential features of a series with a trend break. The solid lines in Figure 2 depict four different processes with trend breaks. For each of the series, we estimated the regression $y_t = a_0 + a_1 \sin(2\pi kt/T) + a_2 \cos(2\pi kt/T) + \gamma \cdot t + e_t$, using the integer frequencies in the interval $1 \leq k \leq 5$. The fitted values for the frequency resulting in the lowest sum of squared residuals are shown by the lines with the short dashes. For example, Panel 1 uses the values $k = 2$, $a_0 = 0.670$, $a_1 = -0.062$, $a_2 = -0.111$ and $\gamma = 0.016$. We repeated the exercise using each fractional frequency in the interval 1/512 to 5 in steps of 1/512. The results using the fractional frequencies are shown by the long dashes in the four panels of Figure 2. The four panels suggest that a Fourier approximation using a single frequency is capable of mimicking breaking trends quite well. As in the case of a break in the intercept, the frequency can be of a small order.[2] Also note that the integer frequencies seem to do as well as fractional frequencies.

3. Test statistics

Since a single frequency component can mimic a reasonable variety of breaks, it seems reasonable to consider the following data-generating process (DGP):

$$y_t = a_0 + \gamma \cdot t + a_1 \sin(2\pi kt/T) + a_2 \cos(2\pi kt/T) + e_t \tag{4}$$

$$e_t = \beta e_{t-1} + \varepsilon_t \tag{5}$$

Note that $\beta = 1$ under the null hypothesis of a unit root, and $\beta < 1$ under the alternative hypothesis.

By adopting the above DGP, the asymptotic distribution for the test of the null hypothesis $\beta = 1$ is invariant to the magnitudes of a_0, a_1, a_2 and γ. We consider two different testing procedures. The first is the Lagrange Multiplier (LM) testing procedure following Schmidt and Phillips (1992) and Amsler and Lee (1995). We examine and provide the asymptotic distributions of the LM-type tests. The second is based on the usual Dickey-Fuller testing

---

[2] Notice that we included a linear trend in the set of deterministic regressors. In principal, it is possible to mimic a continually increasing or decreasing series using a small value for the fractional frequency.

procedure. The specific expressions for the asymptotic distributions of the DF version of the tests differ from those of the LM type tests. However, since the asymptotic properties are not noticeably different from those of the LM type tests, they are omitted to save space.

For the LM test, we employ a two-step estimation methodology. In the first step, we employ the LM principle by imposing the null restriction and estimate the following regression in first-differences:

$$\Delta y_t = \delta_0 + \delta_1 \Delta\sin(2\pi kt/T) + \delta_2 \Delta\cos(2\pi kt/T) + u_t \tag{6}$$

We denote the estimated coefficients as $\tilde{\delta}_0$, $\tilde{\delta}_1$ and $\tilde{\delta}_2$ and construct a detrended series using these coefficients as:

$$\tilde{S}_t = y_t - \tilde{\psi} - \tilde{\delta}_0 t - \tilde{\delta}_1 \sin(2\pi kt/\mathrm{T}) - \tilde{\delta}_2 \cos(2\pi kt/T), \quad t=2, \ldots, T \tag{7}$$

where $\tilde{\psi} = y_1 - \tilde{\delta}_0 - \tilde{\delta}_1 \sin(2\pi k/T) - \tilde{\delta}_2 \cos(2\pi k/T)$, and y_1 is the first observation of y_t. We subtract $\tilde{\psi}$ from y_t to control for the effect of initial values, thus making $\tilde{S}_1 = 0$. The second-step regression equation is

$$\Delta y_t = \phi \tilde{S}_{t-1} + d_0 + d_1 \Delta\sin(2\pi kt/T) + d_2 \Delta\cos(2\pi kt/T) + \varepsilon_t. \tag{8}$$

If $\tilde{S}_{t-1}$ is not stationary, it must be the case that $\phi = 0$; hence, the LM test statistic is:

$$\tau_{LM} = t\text{-statistic for the null hypothesis } \phi = 0. \tag{9}$$

To allow for serially correlated as well as heterogeneously distributed innovations, we assume that the innovations ε_t in the DGP (5) satisfy the regularity conditions of Phillips and Perron (1988, p. 336). Specifically, we assume:

Assumption 1. (a) $E(\varepsilon_t) = 0$ for all t; (b) $\sup_t \mathrm{E}|\varepsilon_t|^{\delta+\omega} < \infty$ for some $\delta > 2$ and $\omega > 0$; (c) $\sigma^2 = \lim_{T\to\infty} T^{1} E(S_T^2)$ exists and $\sigma^2 > 0$, where $S_T = \sum_{t=1}^{T} \varepsilon_t$; (d) $\{\varepsilon_t\}_1^{\infty}$ is strong mixing with mixing numbers α_m that satisfy: $\sum_{1}^{\infty} \alpha_m^{1-2/\delta} < \infty$.

We also assume that the usual error variance exists:

$$\sigma_\varepsilon^2 = \lim_{T\to\infty} T^{1} E(\varepsilon_1^2 + ... + \varepsilon_T^2)$$

The innovation variance σ_ε^2 is estimated as the sum of squared residuals from regression (8). In the presence of serial correlation, we need to estimate the long-run variance σ^2 by choosing a truncation lag parameter l and a set of weights w_j, $j = 1, .., l$ such that: $\tilde{\sigma}^2 = \hat{\gamma}_0 + 2\Sigma w_j\hat{\gamma}_j$, where $\hat{\gamma}_j$ is the jth sample autocovariance of the residuals from (8). Then we modify the statistics accordingly with the correction factor $\tilde{\omega}^2 = \tilde{\sigma}^2 / \tilde{\sigma}_\varepsilon^2$ to correct for the effect of autocorrelated errors as in Schmidt and Phillips (1992). An alternative approach is to augment (8) with lagged values of $\Delta\tilde{S}_{t-j}$. In the examples below, we consider this second approach.

To find the asymptotic distribution of the test statistic, we need to establish:

Lemma 1: *Suppose that y_t is generated by the DGP in* (4) *and* (5) *with* $\beta = 1$, *and one adopts the first step testing regression* (6). *Then,*

$$\sqrt{T}\,(\tilde{\delta}_0 - \delta_0) \to \sigma W(1)$$

$$\frac{1}{\sqrt{T}}(\tilde{\delta}_1 - \delta_1) \to \sigma\,[(2\pi k)\int_0^1 \cos^2(2\pi kr)dr]^{-1}[W(1)+(2\pi k)\int_0^1 \sin(2\pi kr)W(r)dr]$$

$$\frac{1}{\sqrt{T}}(\tilde{\delta}_2 - \delta_2) \to \sigma\,[\int_0^1 \sin^2(2\pi kr)dr]^{-1}\,[\int_0^1 \cos(2\pi kr)W(r)dr] \tag{10}$$

Proof. See the Appendix.

Utilizing the results in Lemma 1, the asymptotic distribution of τ_{LM} is such that:

Theorem 1: *Suppose that y_t is generated by the DGP in* (4) *and* (5) *with* $\beta = 1$, *and one adopts the testing regressions* (6) - (8). *Then, under the null hypothesis:*

$$\tau_{LM} \rightarrow -\frac{1}{2}(\sigma_\varepsilon/\sigma)\left[\int_0^1 \underline{V}(r)^2 dr\right]^{-1/2} \qquad (11)$$

where $\underline{V}(r)$ *is the projection of the process* $V(r)$ *on the orthogonal complement of the space spanned by the trigonometric function* $dz = (1, d\sin(2\pi kr), d\cos(2\pi kr))'$*; and where* $V(r) =$ $W(r) - rW(1) - [(2\pi k)\int_0^1 \cos^2(2\pi kr)dr]^{-1}[W(1)+(2\pi k)\int_0^1 \sin(2\pi kr)W(r)dr]\cdot\sin(2\pi kr) -$ $[\int_0^1 \sin^2(2\pi kr)dr]^{-1}[\int_0^1 \cos(2\pi kr)W(r)dr]\cdot\cos(2\pi kr)$*, with* $W(r)$ *being a Wiener process on* $r \in [0, 1]$.

Proof. See the Appendix.

The above results show that the asymptotic distribution of τ_{LM} depends on the frequency k, but is invariant to all other coefficients in the DGP. In particular, the asymptotic distribution of the test for a unit root is invariant to coefficients a_1 and a_2. As such, the test for a unit root does not depend of the nature of the nonlinear trend.[3]

It is important to note that the test statistics will not be invariant to the magnitudes of a_1 and a_2 unless *both* the sine and cosine functions in the estimating equations (6) and (8). If either the sine or cosine function is excluded from (8), the test diverges when the coefficient a_1 or a_2 in the DGP (4) is not zero.[4] Therefore, unlike Bierens (1997), we do not need to assume $a_1 = a_2 = 0$

[3] At first, it would seem that the LM test statistic does not depend on the frequency k. Saikkonen and Lutkephol (2002) show that LM-type tests with a general nonlinear shift function will not depend on the nuisance parameters in the nonlinear function. They basically extend the finding of Amsler and Lee (1995) who initially show that the LM-type test will not depend on the nuisance parameter indicating the location of level shifts. This result occurs, since the first step regression is based on first differences. However, this invariance does not hold in the LM tests with the trigonometric functions. Intuitively, the first difference of the sine (cosine) function becomes another trigonometric cosine (sine) function. The dependency of the test on the frequency does not pose a problem, however, as seen from the simulation results in Section 4.

[4] We examine this issue more closely in Section 4.

under the null when testing for $\beta = 1$, and the nonlinear trend can exist both under the null and alternative hypotheses.

To obtain critical values via simulations, we employ the DGP in (4) and (5) with $\beta = 1$ and $a_1 = a_2 = 0$. Since our tests are invariant a_1 or a_2 under the null, using any non-zero values of these coefficients will lead to the same critical values. Pseudo-*iid N(0,1)* random numbers were generated using the Gauss procedure RNDNS 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 of τ_{LM} are reported in the right-hand side of Table 1 for the sample sizes $T = 100$ and 500. The critical values were calculated using 100,000 replications for different frequency values of $k = 1,..,10$.

As suggested by Figures 1 and 2, the choice of $k = 1$ (or possibly $k = 2$) can mimic many types of structural breaks. In most circumstances, the researcher can impose this pre-specified value to in order to approximate a wide variety of actual data-generating processes. In this way, (6) can be estimated directly and the estimated coefficients can be used to construct (7). The *t*-statistic for the null hypothesis $\phi = 0$ in (8) can be compared to the critical values reported in the right-hand side of Table 1. In essence, the method filters out a low frequency component (such as a break or other form of nonlinearity) that might interfere with the unit-root test for $\beta = 1$.

One problem with this procedure is that the critical values of τ_{LM} are further from zero than those for a linear model. Thus, as we explore in detail below, it is possible to increase the power of the unit-root test by pre-testing for a nonlinear trend. In order to develop such a test, we use the following *F*-statistic against the alternative nonlinear trend with a given frequency *k*:

$$F(k) = \frac{(SSR_0 - SSR_1(k))/2}{SSR_1(k)/(T-q)} \qquad (12)$$

Here, $SSR_1(k)$ denotes the SSR from equation (8), q is the number of regressors, and SSR_0 denotes the SSR from the regression without the trigonometric terms. The distribution of the F-statistic is non-standard when the unit-root null is imposed on the DGP. Specifically, the asymptotic distribution of $F(k)$ obtained under the unit-root null hypothesis, is given as follows:

Lemma 2: *Suppose that y_t is generated by the DGP* (4) and (5) *with $\beta = 1$, and $a_1 = a_2 = 0$ such that the null implies the absence of the nonlinear functions. Then,*

$$F(k) \rightarrow \underset{k}{Max}\ \frac{1}{8}(\sigma_\varepsilon^2/\sigma^2)[(\int_0^1 \underline{V}(r)^2 dr)^{-1} - (\int_0^1 \underline{V}_0(r)^2 dr)^{-1}] \tag{13}$$

where $\underline{V}_0(r)$ the demeaned Brownian bridge, and $\underline{V}(r)$ is defined in Theorem 1.

Proof. See the Appendix.

In Table 2, we report the simulated critical values of $F(k)$ against the alternative with a nonlinear trend at k frequency. If the sample value of $F(k)$ is sufficiently large (so that the null hypothesis $a_1 = a_2 = 0$ is rejected), employ our τ_{LM} statistics using the nonlinear trend Fourier function. If the null is not rejected, it is possible to gain power by using the usual LM statistics without a nonlinear trend.

One remaining question concerns the effect on the usual LM unit-root tests if a nonlinear trend exists but it is ignored. Perron (1989) earlier suggested that there will be a bias against rejecting a false unit root if an existing structural break is ignored in the usual DF test. We examine the asymptotic property of the LM tests under this situation with a nonlinear trend.

Lemma 3: *Suppose that a nonlinear trend occurs in the data, and the DGP implies* (4) *and* (5) *with $\beta < 1$, but the nonlinear trend is ignored and usual LM tests with a linear trend are employed. Then, the resulting OLS estimate follows*:

$$\hat{\phi} \rightarrow \frac{\sigma_{\varepsilon}^{2}(\beta - 1)}{H(k,r)} \tag{14}$$

with

$$H(k,r) = \sigma_{\varepsilon}^{2} + (1/3)(\varepsilon_{\infty}^{2}+\varepsilon_{1}\varepsilon_{\infty}+\varepsilon_{1}^{2}) + a_{1}^{2}\int_{0}^{1}\sin^{2}(2\pi kr)dr + a_{2}^{2}\int_{0}^{1}\cos^{2}(2\pi kr)dr$$

$$+ a_{1}^{2} - 2\sigma W(1)[a_{1}\int_{0}^{1} r\cdot\sin(2\pi kr)dr + a_{2}\int_{0}^{1} r\cdot\cos(2\pi kr)dr];$$

where β *is the true parameter value in the DGP* (*5*).

Proof. See the Appendix.

First, looking at the numerator, we note that $\hat{\phi} \rightarrow 0$ as $\beta \rightarrow 1$. The obvious conclusion is that the unit-root null will not be rejected for β near unity. The power of the test will increase as $\beta \rightarrow 0$, which is also obvious. Second, the denominator gets larger as the magnitude of the coefficients a_1 and a_2 increase. Then, we will observe that $\hat{\phi} \rightarrow 0$ and this leads to non-rejections of the null. This result implies that there will be loss of power under the alternative, if the existing non-trend is ignored. Thus, the loss of power depends on the magnitude of the coefficients a_1 and a_2 under the alternative. The loss of power is understood in line with Perron's (1989) finding that unit-root tests will fail to reject a false unit root if an existing structural break is ignored. This finding applies to the case of a nonlinear trend, and illustrates the importance of controlling for a nonlinear trend.

A Data-Driven Method of Selecting *k*

A completely agnostic approach to the problem of detecting breaks is to select *k* using purely statistical means. We refrain from using the expression 'estimate *k*' since the trigonometric terms are employed to approximate, not actually identify, the potential breaks. We follow Davis (1987) by using a grid-search method such that the value $k = \hat{k}$ minimizes the sum of squared residuals (*SSR*) from (8). Specifically, for each integer value of k in the interval $1 \leq k \leq kmax$, we estimate (8) select k from the regression yielding the best fit. Although the paper reports results for *kmax* as large as 10, we suggest using the integer values 1 through 5 since low frequencies are associated with breaks. In contrast, as established in Becker, Enders and Hurn (2004), high frequency components could be due to various forms of stochastic parameter instability. Let $\hat{k}$ denote the value of k that yields the smallest sum of squared residuals. We denote the fact that the critical values of τ_{LM} are a function of $\hat{k}$ by:

$$\tau_{LM} = \tau_{LM}(\hat{k}) \tag{15}$$

Practically speaking, the distribution of the statistic depends on how accurately the Fourier approximation in (4) mimics the actual DGP. As discussed in Section 4, our simulations show encouraging results that the frequency is well-estimated. As such, we conjecture that the grid-search procedure yields a consistent estimate of k when the true DGP is given by (4). Thus, we suggest using the critical values in Table 1 for the estimated value of k. Moreover, since k is an unknown nuisance parameter, we can consider the following modification of the F-test given by (12):

$$F(\hat{k}) = \underset{k}{Max}\, F(k), \tag{16}$$

where $\hat{k} = \underset{k}{argmax}\, F(k)$. It is obvious that the value of k giving a minimum SSR value will maximize the F-statistic in (12) such that $\hat{k} = \underset{k}{arginf}\, SSR_1(k)$. We also report the critical values of $F(\hat{k})$ in the Panel 2 of Table 2. To utilize our tests, we first obtain $\hat{k}$ from (12) by minimizing the *SSR* and applying the F-test with $F(\hat{k})$ to examine whether a nonlinear trend exists. If the null of absence of a nonlinear trend is rejected, then we employ our suggested statistics with a nonlinear trend Fourier function. If the null is not rejected, we utilize the usual LM statistics without a nonlinear trend.

A Dickey-Fuller Version of the Test

We adopt a second version of the test based on the usual Dickey-Fuller specification. We introduce the DF-type test since it exhibits similar asymptotic properties and can be readily employed in empirical applications when it requires only the use of OLS. We do not provide asymptotic results for the DF version of the test, since their expressions are more complicated and their asymptotic properties are not different from the LM version of the test. Below, we will compare the performance of the DF-type test with the LM-type test via simulations.

In the DF version of the test, we nest both the null and alternative models and obtain the following testing regression:

$$\Delta y_t = \rho\, y_{t-1} + c_1 + c_2\, t + c_3 \sin(2\pi kt/T) + c_4 \cos(2\pi kt/T) + e_t \tag{17}$$

The nesting equation could have additionally included $\Delta\sin(2\pi kt/T)$ and $\Delta\cos(2\pi kt/T)$, but we omit these terms to avoid collinearity since $\Delta\sin(2\pi kt/T) = (2\pi k/T)\cos(2\pi kt/T)$, and $\Delta\cos(2\pi kt/T) = -(2\pi k/T)\sin(2\pi kt/T)$. As in the LM-version of the test, it is important to notice that we include both the sine and cosine functions in the testing regression. Intuitively, both terms are included by nesting the null and the alternative models. By including both terms, the resulting unit-root test statistics will be invariant to all coefficient parameters in the DGP (4) under the null. As shown in Section 4, if the sine or the cosine function is excluded, the resulting test statistic will not be invariant to the coefficients in the DGP (4), and the tests will exhibit potentially spurious rejections when the coefficients are non-zero; see Lee and Strazicich (2003) for a similar problem with linear structural changes.

The DF type test statistic using (17) will depend only on the frequency k and the sample size T. If the researcher is willing to specify $k = 1$, the test can be conducted directly. If the value of k is estimated, the test for a break can be performed as follows:

Step 1: Estimate (17) for all integer values of k such that $1 \leq k \leq 5$. The regression with the smallest SSR yields $\hat{k}$. If the residuals exhibit serial correlation, augment (17) with lagged values of Δy_t.

Step 2: Perform the F-test for the null hypothesis $c_3 = c_4 = 0$. The critical values for sample sizes of 100 and 500 are shown in the lower left-hand portion of Table 2. For example, with a sample size of 100, the critical value at the 5% level is 9.408. This is slightly larger than the associated value of 9.010 for the LM version of the test. If the frequency k is pre-specified, one can use the upper portion of the table. If k is estimated, the supremum values listed Panel 2 can be used. In either case, if the sample value of F is less than the critical value reported in Table 2, the null hypothesis of a linear trend is not rejected. At this circumstance, we recommend performing the usual linear Dickey-Fuller test.

Step 3: Let τ_{DF} denote the t-statistic for the null hypothesis $\rho = 0$ in (17). Critical values of τ_{DF} are reported on the left-hand side of Table 1 for each possible estimated value of k. As in the LM-version of the test, we suggest using these critical values even if k is estimated.

The Absence of a Time Trend: In some circumstances there is no need to include a deterministic time trend in (17). We refer to this test as τ_{DF_C}. For these situations, we obtained the appropriate critical values by excluding the trend function t from the estimating equation. The critical values obtained from our Monte Carlo simulations are shown in Table 3. For example, for $T = 100$, and $k = 1$, the 5% critical value for the null hypothesis $\rho = 0$ is –3.816. Notice that the corresponding value in Table 1 is –4.347. Hence, it is possible to increase the power of the test by excluding an unnecessary time trend from the estimating equation. The right-hand side of Table 3 shows the critical values for the F-test for the null hypothesis $c_3 = c_4 = 0$ for each value of k. When k is treated as an unknown, it is necessary to use the reported supremum $F(\hat{k})$ values.

4. The Monte Carlo experiments

In this section, we conduct a number of simulation experiments to evaluate performance of the unit-root tests combined with a Fourier representation of a nonlinear trend. All simulations are performed using Monte Carlo 20,000 replications. To save space, we report only the results where k is estimated from the data. Clearly, the test will have better performance is the actual value of k in the DGP is used. Our approach is a two-step procedure in that we first determine if a nonlinear trend exists or not. For the LM version of the test, we estimate (8) and test the null hypothesis $\delta_1 = \delta_2 = 0$ using the values reported in the lower portion of Table 2. Similarly, for the DF-type test, we estimate an equation in the form of (17) and test the null hypothesis $c_3 = c_4 = 0$. Hence, for each test, we use a supremum test $F(\hat{k})$ where $\hat{k} = arginf\ SSR(k)$. If the null hypothesis of a liner trend, $a_1 = a_2 = 0$ (or $\delta_1 = \delta_2 = 0$) cannot be rejected, we apply usual linear Dickey-Fuller or LM unit-root tests. Instead, if the null hypothesis is rejected, we use the DF or LM test statistics reported as τ_{DF} or τ_{LM} in Table 1. The lag determination is done jointly along the lines suggested by Ng and Perron (1995). Starting from a maximum of $p = 8$ lagged terms, the procedure looks for significance of the last augmented term. We use the 10% asymptotic normal value of 1.645 on the t-statistic of the last first-differenced lagged term.

Size of the test for $\beta = 1$: Table 4 reports the size of the LM version of our test for various values of a_1 and a_2, $\beta = 1.0$ and 0.9, values of $k = 1$ through 5 for a sample size of 100. Consider Panel (*a*) of the table for the case $k = 1$, $a_1 = 0$, and $a_2 = 5$. At the five-percent significance level, the test rejects the null hypothesis of a unit root (i.e., $\beta = 1$) in exactly five percent of the 20,000 Monte Carlo replications. At the ten-percent significance level, we reject the null hypothesis in 8.9 percent of the replications. Since the asymptotic critical values are invariant to the magnitudes of a_1 and a_2, it is not surprising to find that that the empirical rejection rates are similar across all non-zero values of these two parameters. Also observe that these results are insensitive to the actual value of k used in the DGP. Notice that when $a_1 = a_2 = 0$, the data-generating process (DGP) is actually linear. Nevertheless, at the five-percent and ten-percent significant levels, the null hypothesis $\beta = 1$ is rejected in seven percent and 12.9% of the replications, respectively.

We also performed the experiment for $T = 500$. The results, available from us on request, indicate that increasing the frequency from $k = 1$ to $k = 2$ or 3 induces a slight reduction in the empirical size of the test.

Power of the test: As in most unit-root tests, when the sample size is small, the power of the test is low. Consider Panel (*b*) of Table 4 for the case $k = 1$, $a_1 = 0$, and $a_2 = 5$. All simulations were conducted using the value of $\beta = 0.9$. At the five-percent and ten percent significance levels, the test correctly rejects the null hypothesis of a unit root (i.e., $\beta = 1$) in only 10.8 percent and 16.9 percent of the Monte Carlo replications, respectively. Notice that the test does better when $a_1 = a_2 = 0$ since the model is actually linear. Increasing the frequency also increases the size of the test. In a second experiment (available from us on request) shows that increasing the sample size to $T = 500$ greatly improves the power of the test.

Estimates of the frequency k: One attractive feature of the test is that it yields an approximation to the form of any break(s) present in the DGP. As such, it is important to document that the estimated frequency seems to be quite close to the actual frequency present in the DGP. The right-hand-portion of Table 4 reports the proportions of the frequencies estimated for a variety series. The test is most likely to select the correct frequency when a_1 and a_2 are large (since these increase the importance of the trigonometric components), when β is small (since a unit root is a 'low frequency' event) and when T is large. As such, we recommend using the estimated value of k in performing the test for $\beta = 1$.

Table 5 performs a similar Monte Carlo experiment using the DF version of the test. Notice that this version of the test also has very good size properties. As shown in Panels (*b*) and (*d*) of the table, the power is low with a sample size of 100 but is almost 100% in every case with a sample size of 500. For the same values of a_1 and a_2, the LM version of the test usually has better size and power properties and the estimated frequency is usually closer to the true frequency. Nevertheless, the differences are small enough that the applied researcher might want to use the DF version of the test as well.

Next, we conducted another Monte Carlo experiment to determine the effects of ignoring nonlinearities in the trend due to trigonometric components. Lemma 3 above indicates that ignoring a *nonlinear* trend affect the performance of the usual unit-root tests both under the null and alternative hypotheses. Tables 6 and 6a indicate the magnitudes of the size distortions and loss of power for the DF-type test and the LM-type test, respectively. When $\beta = 1$, the linear DF and LM tests exhibit serious size distortions. This is true regardless of the magnitudes of a_1 and a_2, the sample size, and the frequency present in the DGP. Under the alternative when $\beta = 0.9$, the power of the DF and the LM test is extremely low. In most instances, the power is near zero.

On the other hand, the simulation results in Tables 4 and 5 show that when the nonlinear trend does not exist in the DGP (i.e., $a_1 = a_2 = 0$), using the tests with the nonlinear trend does not seriously distort the performance of the test. The estimation of a nonlinear trend when the trend is actually linear does not distort the size of the test and the power is comparable to the usual unit-root tests.

Earlier, we noted that it is important to include both sine and cosine functions in our testing regressions to make the tests invariant to the coefficient parameters in the DGP. We now examine this issue using a Monte Carlo experiment. We assume that the DGP in (4) and (5) includes only the cosine function and does not include the sine function. Thus, we have:

$$y_t = a_0 + \gamma \cdot t + a_2 \cos(2\pi kt/T) + e_t \tag{4$'$}$$

$$e_t = \beta e_{t-1} + \varepsilon_t \tag{5$'$}$$

Then, the null and alternative model can be specified as:

Null: $\Delta y_t = \mu_0 + a_2 \Delta\cos(2\pi kt/T) + v_{1t}$

Alternative : $y_t = \mu_1 + \phi_0 \cdot t + \phi_2 \cos(2\pi kt/T) + v_{2t}$

Nevertheless, the DF and/or LM testing regressions should include both $\sin(2\pi kt/T)$ and $\cos(2\pi kt/T)$ to make the resulting unit-root statistics invariant to the coefficient a_2. The term $\sin(2\pi kt/T)$ is to be added to the testing regression which nests the null and alternative models, since $\Delta\cos(2\pi kt/T) = -(2\pi k/T)\sin(2\pi kt/T)$. To save space, we report only the results of a DF-type testing regression that omits $\sin(2\pi kt/T)$.

$$\Delta y_t = c_0 + c_2 t + c_4 \cos(2\pi kt/T) + \rho y_{t-1} + e_t \tag{17$'$}$$

Table 7 reports the results of 20,000 replications of (4)′ and (5)′ estimated using (17)′. However, as shown in Panel (*a*) of Table 7, the size of the test is correct only when $a_2 = 0$ so that the nonlinearity disappears. When $\beta = 0.9$, Panel (b) of Table 7 shows that the null hypothesis if a unit root is always rejected unless the actual DGP is linear (i.e., $a_2 = 0$). Thus, the null hypothesis is rejected too often. Although this misspecification might result in a test with good power, it also means that the test produces spurious rejections since it has the wrong size. The point is that the invariance of our test to a_1 and a_2 is obtained by nesting both the null and alternative models in the estimating equation.

5. Examples of the test

In this section, we apply our test to two well-studied examples. Both are designed to illustrate the ability of the test to mimic an unknown number of breaks of unknown functional form occurring at unknown break dates.

Real U.S. GDP: In his own work, Perron (1989) used the oil price shock of 1973 as the break date when examining quarterly values of real U.S. GNP. However, the actual dating of the oil price shock may not be so obvious. In fact, during 1973, OPEC increased posted prices by 5.7% on April 1, 11.9% on June 1, 17% on October 16, and declared an export embargo on October 20. Moreover, structural breaks need not occur instantaneously or manifest themselves contemporaneously. Consider the effects of the October 1973 oil price shock on real U.S. GDP measured in billions of 1996 dollars:

	1973:1	1973:2	1973:3	1973:4	1974:1	1974:2	1974:3	1974:4	1975:1
GDP	4092.3	4133.3	4117.0	4151.1	4119.3	4130.4	4084.5	4062.0	4010.0

Although Perron (1989) used 1973:1 as his break date, real GDP actually rose in 1973:2 and 1973:4, fell in 1974:1, rose in 1974:2 and then fell steadily for the next three quarters. Such behavior cannot be adequately captured by a single break in the trend. Instead, it is more likely that the main effect of the oil price shock on real GDP was delayed at least one quarter and that the effect of the shock was smooth and sustained. As such, it is seems plausible to allow for some form of smooth break.

In order to perform the LM version of out test, we obtained data on real U.S. GDP for the period 1947:1 through 2003:2 from the website of the Federal Reserve Bank of St. Louis (FRED). We let $\{y_t\}$ denote the natural logarithm of real U.S. GDP and estimated a regression in the form of (8) using integer values of k ranging from 1 to 10. Since the regression equation using $k = 1$ had the lowest residual sum of squares, we used this value as the consistent estimate of the actual frequency. Imposing $k = 1$, the estimated equation, with t-statistics in parentheses, becomes:

$$\underset{}{\Delta y_t} = \underset{(24.33)}{0.00868} - \underset{(-10.28)}{0.18569}\,\Delta\sin(2\pi t/T) + \underset{(3.41)}{0.06151}\,\Delta\cos(2\pi t/T) \qquad (6)'$$

Next, we used these three estimated coefficients to construct $\tilde{S}_t$ and $\tilde{\psi}$ as in (7) and estimated an equation in the form of (8). Since we wanted to control for serial correlation in the residuals, we augmented (8) with lagged changes of $\{\tilde{S}_t\}$. The tenth augmented lag had a t-statistic of 2.77 so we used ten lags of $\Delta\tilde{S}_{t-i}$ to obtain:

$$\Delta y_t = -0.0128\tilde{S}_{t-1} + 0.0094 - 0.1859\Delta\sin(2\pi t/T) + 0.0602\Delta\cos(2\pi t/T) + \sum_{i=1}^{10}\hat{\beta}_i\Delta\tilde{S}_{t-i} \qquad (8)'$$

The sample value of the F-statistic for the null hypothesis that the coefficients on the trigonometric components jointly equal zero is 128.19. If we compare this value to that reported in Table 2, it is clear that we can reject the hypothesis of linearity. However, the coefficient for $\tilde{S}_{t-1}$ has a t-statistic of –2.602. As shown in Table 1, at the 5% significance level, with $k = 1$, the critical value of τ_{LM} is –4.110. Hence, we cannot reject the null hypothesis of a unit root.

A key feature of the test is that it allows us to mimic the form of the nonlinearity present in the data. Since we are now most interested in the nature of the nonlinearity, we re-estimated equation (8)′ allowing for the possibility of fractional frequencies. Since fractional frequencies can best capture the nature of the nonlinearity, we performed a grid search using the frequencies in the interval $0 < k \leq 2$ we found that $k = 1.45$ resulted in the lowest sum of squared residuals. The top panel of Figure 3 plots the actual value of the $\{y_t\}$ series along with the fitted value of the trend for $k = 1.45$. You can clearly see that the nonlinear trend does capture the upward drift of the series. Panel *b* of Figure 3 shows just the time varying intercept from:

$$\Delta y_t = \underset{(30.66)}{0.01004} + \underset{(3.88)}{0.04307} \sin[2\pi(1.45)t/T] + \underset{(14.33)}{0.16120} \cos[2\pi(1.45)t/T] \qquad (8)''$$

Instead of a sharp break, Panel *b* indicates that the drift in U.S. GDP growth began to turn around sometime near the end of 1958. Trend growth continued to increase (reaching a maximum of 0.0174% per quarter near 1976. Thereafter, trend growth exhibited a sustained decline until 1997. In contrast, the type of break used by Perron (1989) assumes a constant long-run growth rate that displayed a one-time fall in 1973:1.

The Term-Structure of Interest Rates: A number of papers, including Enders and Granger (1998), Shin and Lee (2001), Lanne and Saikkonen (2002), and Seo (2003) suggest that interest rate spreads should be stationary such that the adjustment towards the long-run equilibrium follows a threshold process. To explore this possibility we obtained monthly values of the 3-month *T*-bill rate and the 1-year (R1) and 3-year (R3) rates on U.S. government securities over the 1990:1 through 2003:11 period. The time paths of the three rates are shown in Figure 4. It is clear that the 3-year rate was substantially above the two shorter-term interest rates throughout the early 1990s. However, in the 1995 - 2000 period, the spreads between R3 and the shorter-term rates declined substantially. In mid-2001, R3 did not decline as rapidly as the other two rates so that the spread became large.

Although the type of behavior shown in Figure 4 might be the result of a threshold process, it is plausible to argue that persistence of the magnitudes of the gaps are due to several structural breaks in the equilibrium level of the spread. Towards this end we estimated each possible interest rate pair as a linear process and as a nonlinear Fourier process in the DF-form:

$$\Delta y_t = \rho y_{t-1} + c_1 + c_3 \sin(2\pi kt/T) + c_4 \cos(2\pi kt/T) + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + \varepsilon_t \tag{18}$$

where y_t is the gap between the long-term and the short-term interest rate. Alternatively, y_t is the difference between the 1-year rate and the *T*-bill rate, the 3-year rate and the *T*-bill rate, and 3-year rate and the 1-year rate.

Our selected specification excluded *time* as a regressor since there is no theoretical reason to believe that there is a deterministic trend in interest rate spreads. Since this specification is nested within our more general model, the test statistics reported in Tables 1 and 2 are appropriately sized. However, it is possible to increase the power of our test by using the critical values reported in Table 3. Recall that these critical values were generated without using *time* as a regressor. For the DF version of the test, we estimated equation (18) using integer frequencies in the range $1 \leq k \leq 10$ and selected the values of k resulting in the smallest residual sum of squares. The resulting estimations can be summarized as follows:

R-long	R-short	*F*-value	τ_{DF}	*k*	*p*	DF	TAR
R1	*T*-bill	8.54	−5.51	1	12	−3.52	0.243
R3	*T*-bill	13.40	−5.97	1	11	−3.27	0.917
R3	R1	11.33	−5.39	1	11	−2.72	0.702

Notice that when y_t is the difference between the 1-year rate and the T-bill rate, the sample value of F for the null $c_3 = c_4 = 0$ is 8.54. Table 3 indicates that the critical values for T = 100 are 6.591, 7.783 and 10.627 at the 10%, 5% and 1% significance levels, respectively. Hence, if wanted to use the 1% significance level, it would be possible to conclude that the intercept for the spread between these two rates does not have a break. As such, it would be appropriate to perform a standard Dickey-Fuller test to determine if the spread is stationary. The linear Dickey-Fuller test (see the next-to-last column of the table) indicates that the t-statistic for the null hypothesis $\rho = 0$ is −3.52. Using the standard Dickey-Fuller distribution, we can just reject the null hypothesis at the 1% significance level (the critical value is −3.51).

The situation is a bit different for the other two regression equations. In both instances, the sample value of the F-statistic for the null hypothesis $c_3 = c_4 = 0$ can be rejected at the 1% level for $(\text{R3} - T\text{-bill})_t$ and for $(\text{R3} - \text{R1})_t$. Given our finding of the nonlinearity in the data, we test for a unit root using the critical values of τ_{DF_C} that are reported in Table 3. It should be clear that both of the sample values are sufficiently negative that we can reject the null hypothesis of a unit root at the 1% significance level (the critical value is –4.433). By way of comparison, the t-values (−3.27 and –2.72) for the standard Dickey-Fuller test imply that the spreads for $(\text{R3} - T\text{-bill})_t$ and $(\text{R3} - \text{R1})_t$ are unit-root processes.

For our purposes, the key issue is how the spreads adjust over time. In particular, we wanted to examine whether the interest rate spreads follow a threshold process. When we apply Hansen's (1997) test for threshold behavior, we are unable to reject the null hypothesis of no threshold any conventional significance level. We used 1000 Monte Carlo replications to obtain the appropriate critical values for the test. The *prob*-values (listed under TAR in the table above) for the three spreads are 0.243, 0.917 and 0.702, respectively. Hence, it is unlikely that our Fourier approximation detects any type of threshold behavior.

To obtain a better understanding of the nature of the Fourier adjustment process, the solid line in Figure 5 shows the time path of the difference between the 3-year rate and 1-year rate. One can clearly see that the spread was much higher in the early 1990s and in 2002 than in the intervening years. The estimated model (with t-statistics in parentheses) is:

$$\Delta y_t = -0.223\, y_{t-1} + 0.160 + 0.080 \sin(2\pi kt/T) + 0.069 \cos(2\pi kt/T) + \sum_{i=1}^{11} \hat{\beta}_i \Delta y_{t-i} \qquad (18)'$$
$$\qquad (-5.38) \qquad (5.29) \quad (3.97) \qquad\qquad (4.16)$$

where: $k = 1$.

The time-varying mean of the $\{y_t\}$ sequence can be obtained by dividing the Fourier intercept [i.e., 0.160 + 0.080 sin($2\pi kt/T$) + 0.069 cos($2\pi kt/T$)] by 0.232 (*i.e.*, one minus the sum of the autoregressive coefficients). You can clearly see by the smooth line in Figure 5, the time-varying mean mimics the fact that the spread was much higher in the early 1990s and in the later part of the sample than in the intervening years. Given that the spreads are stationary, we also examined the breaks that are identified by the Bai-Perron procedure. Allowing for a maximum of five breaks with a minimum break-size of 12 months, the BIC selected three breaks. The first occurs at 1994:11, the second at 2000:12, and the third at 2002:6. The dashed line in Figure 5 shows the breakpoints and the four sub-period means of the (R3 – R1) spread. Notice that these sharp breaks are similar to the Fourier intercept. The difference, of course, is that the Fourier intercept is smooth and that the Bai-Perron procedure does not embody a unit-root test.

6. Summary and conclusion

The paper develops a unit-root test that allows for an unknown number of structural breaks with unknown functional forms. The test is based on the fact that a single frequency component of a Fourier approximation can capture a process of gradual change in a time-varying intercept. Nevertheless, as shown in Section 2, the test can often capture sharp breaks and/or nonlinear trends. Hence, instead of selecting specific break dates or the form of the breaks, the specification problem is transformed into selecting the proper frequency component to include in the estimating equation.

In Section 3, we show that our test is invariant to all parameters in the DGP except for the frequency. As such, we are able to develop a procedure to determine whether a Fourier frequency component belongs in the estimating equation. This *F*-test is a supremum-type test in that the selected frequency provides the smallest residual sum of squares. In Section 4, we show that our proposed unit-root test does not exhibit any serious size distortions, and shows good power. In Section 5, the appropriate use of the test is illustrated using real GDP and the interest rate differential. We show that real U.S. GDP can be characterized as a unit-root process with a nonlinear trend. We reaffirm the well-known result that interest rate spreads are stationary although the equilibrium value of the spread has been time-varying.

References

Amsler, C. and J. Lee, 1995, An LM test for a unit root in the presence of a structural change, *Econometric Theory* 11, 359 – 368.

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

Becker, R., W. Enders, and S. Hurn, 2004, A general test for time dependence in parameters, *Journal of Applied Econometrics*, forthcoming.

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 A. Harvey, 2003, Seasonality tests, *Journal of Business and Economic Statistics*, 21, 420 – 436.

Clemente, J., A. Montanes, and M. Reyes, 1998, Testing for a unit root in variables with a double change in the mean, *Economics Letters* 59, 175-182.

Clements M. P. and D. F. Hendry, 1999, On winning forecasting competitions, *Spanish Economic Review*, 1, 123 – 160.

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

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

Gallant, R., 1984, The Fourier flexible form, *American Journal of Agricultural Economics* 66, 204 – 208.

Gallant, R. and G. Souza, 1991, On the Asymptotic normality of Fourier flexible form estimates, *Journal of Econometrics* 50, 329 – 353.

Hansen, B., 1997, Inference in TAR models, *Studies in Nonlinear Dynamics and Econometrics* 2, Article 1.

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

Lanne, M. and P. Saikkonen, 2002, Threshold autoregressions for strongly autocorrelated time series, *Journal of Business and Economic Statistics* 20, 282 – 289.

Lee, J., and M. Strazicich, 2003, Minimum LM unit root tests with two structural breaks, *Review of Economics and Statistics* 85, 082-1089.

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

Ng, S. and P. Perron, 1995, Unit Root Tests in ARMA Models with data-dependent methods for the selection of the truncation lag, *Journal of the American Statistical Association* 90, 269-281.

Perron, P., 1989, The great crash, the oil price shock, and the unit root hypothesis, *Econometrica* 57, 1361 – 1401.

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

Saikkonen, P. and H. Lutkepohl, 2002, Testing for a Unit root in a time series with a level shift at unknown time, *Econometric Theory* 18, 313-348.

Schmidt, P. and P. Phillips, 1992, LM Tests for a unit root in the presence of deterministic Trends, *Oxford Bulletin of Economics and Statistics* 54, 257-287.

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.

Seo, B. (2003), Nonlinear mean reversion in the term structure of interest rates, *Journal of Economic Dynamics and Control* 27, 2243 – 65.

Shin, D. and O. Lee, 2001, Tests for asymmetry in possibly nonstationary time series data, *Journal of Business and Economic Statistics* 19, 233 – 44.

Vogelsang, T. and P. Perron, 1998, Additional tests for a unit root allowing for a break in the trend function at an unknown time, *International Economic Review* 39 (4), 1073-1100.

Table 1: Critical Values of τ_{DF} and τ_{LM}

		τ_{DF}			τ_{LM}		
T	*k*	1%	5%	10%	1%	5%	10%
100	1	-4.954	-4.347	-4.050	-4.687	-4.110	-3.820
	2	-4.700	-4.039	-3.704	-4.235	-3.565	-3.220
	3	-4.461	-3.770	-3.424	-3.977	-3.301	-2.961
	4	-4.294	-3.626	-3.294	-3.842	-3.179	-2.856
	5	-4.199	-3.551	-3.222	-3.765	-3.117	-2.806
	10	-4.031	-3.425	-3.124	-3.606	-3.019	-2.733
	Linear*	-4.044	-3.450	-3.146	-3.632	-3.054	-2.766
500	1	-4.835	-4.278	-4.006	-4.585	-4.041	-3.780
	2	-4.578	-3.985	-3.676	-4.152	-3.550	-3.222
	3	-4.371	-3.750	-3.426	-3.914	-3.299	-2.977
	4	-4.252	-3.627	-3.304	-3.804	-3.184	-2.881
	5	-4.163	-3.560	-3.247	-3.740	-3.135	-2.834
	10	-4.027	-3.447	-3.155	-3.603	-3.047	-2.769
	Linear*	-3.977	-3.423	-3.134	-3.575	-3.033	-2.754

* These are the critical values of the usual DF or LM statistics with a linear trend.

Table 2: Critical Values of *F(k)* and $F(\hat{k})$

F(k)

		τ_{DF}			τ_{LM}		
T	*k*	10%	5%	1%	10%	5%	1%
100	1	7.219	8.700	12.000	7.182	8.575	11.629
	2	4.622	5.985	9.200	3.771	4.963	7.746
	3	3.329	4.414	7.027	2.918	3.844	6.133
	4	2.930	3.853	5.811	2.627	3.447	5.546
	5	2.681	3.532	5.497	2.479	3.274	5.144
	10	2.338	3.046	4.780	2.304	3.027	4.708
500	1	6.925	8.287	11.166	6.859	8.157	10.85
	2	4.549	5.843	8.597	3.738	4.882	7.520
	3	3.388	4.460	6.826	2.921	3.844	5.966
	4	2.868	3.732	5.719	2.652	3.452	5.378
	5	2.711	3.520	5.368	2.514	3.281	5.117
	10	2.420	3.133	4.711	2.352	3.087	4.756

Panel 2: $F(\hat{k}) = Max\ F(k)$

	τ_{DF}			τ_{LM}		
T	10%	5%	1%	10%	5%	1%
100	8.052	9.408	12.469	7.679	9.010	11.983
500	7.659	8.852	11.523	7.344	8.532	11.084

Table 3: Critical Values of τ_{DF_C} without a linear trend and $F(k)$

		τ_{DF_C}			$F(k)$		
T	*k*	1%	5%	10%	1%	5%	10%
100	1	-4.433	-3.816	-3.495	10.193	7.137	5.756
	2	-3.975	-3.270	-2.900	6.736	4.256	3.207
	3	-3.733	-3.059	-2.710	5.471	3.539	2.680
	4	-3.618	-2.968	-2.640	5.111	3.302	2.494
	5	-3.543	-2.910	-2.597	4.916	3.139	2.396
	Linear*	-3.525	-2.902	-2.583			
					$F(\hat{k}) = Max\ F(k)$		
					10.627	7.783	6.591
500	1	-4.362	-3.762	-3.456	9.566	6.837	5.580
	2	-3.886	-3.239	-2.892	6.404	4.170	3.190
	3	-3.702	-3.060	-2.727	5.537	3.521	2.679
	4	-3.583	-2.970	-2.646	5.100	3.267	2.510
	5	-3.541	-2.938	-2.619	4.909	3.155	2.444
	Linear*	-3.435	-2.870	-2.572			
					$F(\hat{k}) = Max\ F(k)$		
					9.952	7.448	6.360

Table 4: Finite Sample Performance of τ_{LM}

(a) $T = 100$, $\beta = 1$

DGP			5%	10%	Relative frequency of the selected values of k						
k	a_1	a_2	Rej. rate	Rej. rate	Linear*	1	2	3	4	5	6 ≥
1	0	0	0.070	0.129	0.95	0.04	0.00	0.00	0.00	0.00	0.00
	0	5	0.050	0.089	0.69	0.31	0.00	0.00	0.00	0.00	0.00
	3	0	0.056	0.097	0.87	0.13	0.00	0.00	0.00	0.00	0.00
	3	5	0.051	0.093	0.57	0.43	0.00	0.00	0.00	0.00	0.00
2	0	0	0.071	0.130	0.95	0.04	0.00	0.00	0.00	0.00	0.00
	0	5	0.051	0.098	0.23	0.00	0.77	0.00	0.00	0.00	0.00
	3	0	0.037	0.066	0.80	0.00	0.20	0.00	0.00	0.00	0.00
	3	5	0.049	0.098	0.09	0.00	0.91	0.00	0.00	0.00	0.00
3	0	0	0.075	0.131	0.95	0.04	0.00	0.00	0.00	0.00	0.00
	0	5	0.049	0.099	0.00	0.00	0.00	1.00	0.00	0.00	0.00
	3	0	0.044	0.083	0.46	0.00	0.00	0.54	0.00	0.00	0.00
	3	5	0.050	0.101	0.00	0.00	0.00	1.00	0.00	0.00	0.00
4	0	0	0.070	0.130	0.95	0.04	0.00	0.00	0.00	0.00	0.00
	0	5	0.051	0.100	0.00	0.00	0.00	0.00	1.00	0.00	0.00
	3	0	0.049	0.095	0.11	0.00	0.00	0.00	0.89	0.00	0.00
	3	5	0.051	0.101	0.00	0.00	0.00	0.00	1.00	0.00	0.00
5	0	0	0.075	0.131	0.95	0.04	0.01	0.00	0.00	0.00	0.00
	0	5	0.048	0.097	0.00	0.00	0.00	0.00	0.00	1.00	0.00
	3	0	0.050	0.102	0.01	0.00	0.00	0.00	0.00	0.99	0.00
	3	5	0.051	0.101	0.00	0.00	0.00	0.00	0.00	1.00	0.00

(b) $T = 100$, $\beta = 0.9$

DGP			5%	10%	Relative frequency of the selected values of k						
k	a_1	a_2	Rej. rate	Rej. rate	Linear*	1	2	3	4	5	6 ≥
1	0	0	0.268	0.437	0.99	0.01	0.00	0.00	0.00	0.00	0.00
	0	5	0.108	0.169	0.72	0.28	0.00	0.00	0.00	0.00	0.00
	3	0	0.097	0.151	0.92	0.08	0.00	0.00	0.00	0.00	0.00
	3	5	0.113	0.193	0.57	0.43	0.00	0.00	0.00	0.00	0.00
2	0	0	0.275	0.438	0.99	0.01	0.00	0.00	0.00	0.00	0.00
	0	5	0.210	0.350	0.11	0.00	0.89	0.00	0.00	0.00	0.00
	3	0	0.130	0.196	0.74	0.00	0.26	0.00	0.00	0.00	0.00
	3	5	0.205	0.360	0.02	0.00	0.98	0.00	0.00	0.00	0.00
3	0	0	0.271	0.435	0.98	0.01	0.00	0.00	0.00	0.00	0.00
	0	5	0.242	0.407	0.00	0.00	0.00	1.00	0.00	0.00	0.00
	3	0	0.198	0.315	0.38	0.00	0.00	0.62	0.00	0.00	0.00
	3	5	0.247	0.408	0.00	0.00	0.00	1.00	0.00	0.00	0.00
4	0	0	0.277	0.441	0.98	0.01	0.00	0.00	0.00	0.00	0.00
	0	5	0.253	0.416	0.00	0.00	0.00	0.00	1.00	0.00	0.00
	3	0	0.245	0.407	0.09	0.00	0.00	0.00	0.91	0.00	0.00
	3	5	0.252	0.412	0.00	0.00	0.00	0.00	1.00	0.00	0.00
5	0	0	0.271	0.434	0.98	0.01	0.00	0.00	0.00	0.00	0.00
	0	5	0.257	0.416	0.00	0.00	0.00	0.00	0.00	1.00	0.00
	3	0	0.253	0.419	0.01	0.00	0.00	0.00	0.00	0.99	0.00
	3	5	0.254	0.419	0.00	0.00	0.00	0.00	0.00	1.00	0.00

Table 5: Finite Sample Performance of τ_{DF}

(a) $T = 100$, $\beta = 1$

DGP			5%	10%	Relative frequency of the selected values of k						
k	a_1	a_2	Rej. rate	Rej. rate	Linear*	1	2	3	4	5	6 ≥
1	0	0	0.073	0.132	0.95	0.04	0.01	0.00	0.00	0.00	0.00
	0	5	0.046	0.080	0.85	0.15	0.00	0.00	0.00	0.00	0.00
	3	0	0.070	0.125	0.88	0.11	0.01	0.00	0.00	0.00	0.00
	3	5	0.052	0.092	0.71	0.29	0.00	0.00	0.00	0.00	0.00
2	0	0	0.072	0.132	0.95	0.04	0.01	0.00	0.00	0.00	0.00
	0	5	0.051	0.099	0.31	0.01	0.69	0.00	0.00	0.00	0.00
	3	0	0.046	0.084	0.76	0.01	0.23	0.00	0.00	0.00	0.00
	3	5	0.050	0.099	0.12	0.00	0.88	0.00	0.00	0.00	0.00
3	0	0	0.074	0.133	0.95	0.04	0.01	0.00	0.00	0.00	0.00
	0	5	0.050	0.104	0.01	0.00	0.00	0.99	0.00	0.00	0.00
	3	0	0.049	0.091	0.44	0.00	0.00	0.56	0.00	0.00	0.00
	3	5	0.052	0.102	0.00	0.00	0.00	1.00	0.00	0.00	0.00
4	0	0	0.076	0.135	0.95	0.04	0.01	0.00	0.00	0.00	0.00
	0	5	0.051	0.100	0.00	0.00	0.00	0.00	1.00	0.00	0.00
	3	0	0.052	0.099	0.11	0.00	0.00	0.00	0.89	0.00	0.00
	3	5	0.052	0.102	0.00	0.00	0.00	0.00	1.00	0.00	0.00
5	0	0	0.074	0.132	0.95	0.04	0.01	0.00	0.00	0.00	0.00
	0	5	0.050	0.098	0.00	0.00	0.00	0.00	0.00	1.00	0.00
	3	0	0.050	0.101	0.01	0.00	0.00	0.00	0.00	0.99	0.00
	3	5	0.050	0.098	0.00	0.00	0.00	0.00	0.00	1.00	0.00

(b) $T = 100$, $\beta = 0.9$

DGP			5%	10%	Relative frequency of the selected values of k						
k	a_1	a_2	Rej. rate	Rej. rate	Linear*	1	2	3	4	5	6 ≥
1	0	0	0.208	0.345	0.98	0.01	0.01	0.00	0.00	0.00	0.00
	0	5	0.086	0.122	0.84	0.16	0.00	0.00	0.00	0.00	0.00
	3	0	0.111	0.189	0.93	0.06	0.00	0.00	0.00	0.00	0.00
	3	5	0.106	0.175	0.67	0.33	0.00	0.00	0.00	0.00	0.00
2	0	0	0.207	0.348	0.98	0.01	0.01	0.00	0.00	0.00	0.00
	0	5	0.159	0.276	0.17	0.00	0.83	0.00	0.00	0.00	0.00
	3	0	0.121	0.191	0.70	0.00	0.30	0.00	0.00	0.00	0.00
	3	5	0.159	0.285	0.03	0.00	0.97	0.00	0.00	0.00	0.00
3	0	0	0.207	0.345	0.98	0.01	0.01	0.00	0.00	0.00	0.00
	0	5	0.184	0.321	0.00	0.00	0.00	1.00	0.00	0.00	0.00
	3	0	0.164	0.266	0.35	0.00	0.00	0.65	0.00	0.00	0.00
	3	5	0.188	0.326	0.00	0.00	0.00	1.00	0.00	0.00	0.00
4	0	0	0.207	0.348	0.98	0.01	0.01	0.00	0.00	0.00	0.00
	0	5	0.188	0.325	0.00	0.00	0.00	0.00	1.00	0.00	0.00
	3	0	0.191	0.322	0.08	0.00	0.00	0.00	0.92	0.00	0.00
	3	5	0.192	0.327	0.00	0.00	0.00	0.00	1.00	0.00	0.00
5	0	0	0.205	0.348	0.98	0.01	0.01	0.00	0.00	0.00	0.00
	0	5	0.190	0.331	0.00	0.00	0.00	0.00	0.00	1.00	0.00
	3	0	0.189	0.327	0.01	0.00	0.00	0.00	0.00	0.99	0.00
	3	5	0.188	0.326	0.00	0.00	0.00	0.00	0.00	1.00	0.00

(c) $T = 500$, $\beta = 1$

DGP			5%	10%	Relative frequency of the selected values of k						
k	a_1	a_2	Rej. rate	Rej. rate	Linear*	1	2	3	4	5	6 ≥
1	0	0	0.072	0.131	0.95	0.04	0.01	0.00	0.00	0.00	0.00
	0	5	0.062	0.110	0.93	0.06	0.01	0.00	0.00	0.00	0.00
	3	0	0.072	0.129	0.94	0.05	0.01	0.00	0.00	0.00	0.00
	3	5	0.062	0.106	0.90	0.09	0.01	0.00	0.00	0.00	0.00
2	0	0	0.074	0.128	0.95	0.04	0.01	0.00	0.00	0.00	0.00
	0	5	0.040	0.074	0.89	0.01	0.10	0.00	0.00	0.00	0.00
	3	0	0.062	0.111	0.93	0.03	0.04	0.00	0.00	0.00	0.00
	3	5	0.042	0.073	0.83	0.01	0.17	0.00	0.00	0.00	0.00
3	0	0	0.076	0.132	0.95	0.04	0.01	0.00	0.00	0.00	0.00
	0	5	0.040	0.072	0.75	0.01	0.00	0.24	0.00	0.00	0.00
	3	0	0.048	0.089	0.93	0.02	0.01	0.05	0.00	0.00	0.00
	3	5	0.041	0.076	0.59	0.00	0.00	0.40	0.00	0.00	0.00
4	0	0	0.075	0.133	0.95	0.04	0.01	0.00	0.00	0.00	0.00
	0	5	0.044	0.087	0.46	0.01	0.00	0.00	0.53	0.00	0.00
	3	0	0.041	0.076	0.89	0.02	0.00	0.00	0.09	0.00	0.00
	3	5	0.048	0.093	0.23	0.00	0.00	0.00	0.77	0.00	0.00
5	0	0	0.073	0.130	0.95	0.04	0.01	0.00	0.00	0.00	0.00
	0	5	0.048	0.096	0.16	0.00	0.00	0.00	0.00	0.84	0.00
	3	0	0.036	0.070	0.80	0.01	0.00	0.00	0.00	0.19	0.00
	3	5	0.048	0.099	0.04	0.00	0.00	0.00	0.00	0.96	0.00

(d) $T = 500$, $\beta = 0.9$

DGP			5%	10%	Relative frequency of the selected values of k						
k	a_1	a_2	Rej. rate	Rej. rate	Linear*	1	2	3	4	5	6 ≥
1	0	0	1.000	1.000	0.99	0.00	0.00	0.00	0.00	0.00	0.00
	0	5	0.861	0.880	0.15	0.86	0.00	0.00	0.00	0.00	0.00
	3	0	0.992	0.999	0.90	0.10	0.00	0.00	0.00	0.00	0.00
	3	5	0.940	0.941	0.06	0.94	0.00	0.00	0.00	0.00	0.00
2	0	0	1.000	1.000	0.99	0.00	0.00	0.00	0.00	0.00	0.00
	0	5	0.943	0.953	0.06	0.00	0.94	0.00	0.00	0.00	0.00
	3	0	0.953	0.993	0.60	0.00	0.40	0.00	0.00	0.00	0.00
	3	5	0.990	0.990	0.01	0.00	0.99	0.00	0.00	0.00	0.00
3	0	0	1.000	1.000	1.00	0.00	0.00	0.00	0.00	0.00	0.00
	0	5	0.992	0.994	0.01	0.00	0.00	0.99	0.00	0.00	0.00
	3	0	0.966	0.994	0.44	0.00	0.00	0.56	0.00	0.00	0.00
	3	5	0.999	1.000	0.00	0.00	0.00	1.00	0.00	0.00	0.00
4	0	0	1.000	1.000	0.99	0.00	0.00	0.00	0.00	0.00	0.00
	0	5	1.000	1.000	0.00	0.00	0.00	0.00	1.00	0.00	0.00
	3	0	0.982	0.998	0.31	0.00	0.00	0.00	0.69	0.00	0.00
	3	5	1.000	1.000	0.00	0.00	0.00	0.00	1.00	0.00	0.00
5	0	0	1.000	1.000	0.99	0.00	0.00	0.00	0.00	0.00	0.00
	0	5	1.000	1.000	0.00	0.00	0.00	0.00	0.00	1.00	0.00
	3	0	0.994	0.999	0.18	0.00	0.00	0.00	0.00	0.82	0.00
	3	5	1.000	1.000	0.00	0.00	0.00	0.00	0.00	1.00	0.00

Table 6: Effects of Ignoring Nonlinear Trends (DF Tests)

(a) $T = 100$

	DGP		$\beta = 1.0$				$\beta = 0.9$			
k	a_1	a_2	5% rej. rate	10% rej. rate	5% crit. value	10% crit. value	5% rej. rate	10% rej. rate	5% crit. value	10% crit. value
1	3	3	0.013	0.025	-2.83	-2.46	0.002	0.004	-2.24	-1.98
	3	5	0.002	0.004	-2.10	-1.74	0.000	0.000	-1.41	-1.23
	5	3	0.006	0.013	-2.50	-2.09	0.000	0.000	-1.69	-1.45
	5	5	0.001	0.002	-1.69	-1.29	0.000	0.000	-0.96	-0.79
2	3	3	0.000	0.001	-2.05	-1.84	0.000	0.001	-2.03	-1.86
	3	5	0.000	0.000	-1.69	-1.52	0.000	0.000	-1.58	-1.46
	5	3	0.000	0.000	-1.57	-1.43	0.000	0.000	-1.48	-1.38
	5	5	0.000	0.000	-1.33	-1.20	0.000	0.000	-1.22	-1.14
3	3	3	0.000	0.000	-2.04	-1.88	0.000	0.000	-2.05	-1.93
	3	5	0.000	0.000	-1.85	-1.68	0.000	0.000	-1.75	-1.65
	5	3	0.000	0.000	-1.63	-1.52	0.000	0.000	-1.61	-1.53
	5	5	0.000	0.000	-1.55	-1.44	0.000	0.000	-1.48	-1.41
4	3	3	0.000	0.000	-2.17	-2.01	0.000	0.000	-2.16	-2.06
	3	5	0.000	0.000	-2.02	-1.87	0.000	0.000	-1.93	-1.85
	5	3	0.000	0.000	-1.82	-1.73	0.000	0.000	-1.79	-1.74
	5	5	0.000	0.000	-1.79	-1.68	0.000	0.000	-1.72	-1.66
5	3	3	0.000	0.000	-2.31	-2.17	0.000	0.000	-2.32	-2.24
	3	5	0.000	0.000	-2.21	-2.08	0.000	0.000	-2.14	-2.06
	5	3	0.000	0.000	-2.05	-1.95	0.000	0.000	-2.02	-1.97
	5	5	0.000	0.000	-2.04	-1.93	0.000	0.000	-1.97	-1.92

Table 6a: Effects of Ignoring Nonlinear Trends (LM Tests)

(a) $T = 100$

	DGP		$\beta = 1.0$				$\beta = 0.9$			
k	a_1	a_2	5% rej. rate	10% rej. rate	5% crit. value	10% crit. value	5% rej. rate	10% rej. rate	5% crit. value	10% crit. value
1	3	3	0.010	0.023	-2.43	-2.10	0.003	0.009	-2.24	-2.03
	3	5	0.002	0.005	-1.91	-1.64	0.000	0.000	-1.61	-1.49
	5	3	0.003	0.006	-1.91	-1.65	0.000	0.001	-1.69	-1.54
	5	5	0.001	0.002	-1.56	-1.36	0.000	0.000	-1.35	-1.26
2	3	3	0.000	0.001	-1.86	-1.70	0.000	0.002	-2.03	-1.89
	3	5	0.000	0.000	-1.46	-1.37	0.000	0.000	-1.55	-1.47
	5	3	0.000	0.000	-1.47	-1.37	0.000	0.000	-1.56	-1.49
	5	5	0.000	0.000	-1.27	-1.20	0.000	0.000	-1.32	-1.27
3	3	3	0.000	0.000	-1.80	-1.70	0.000	0.000	-1.99	-1.89
	3	5	0.000	0.000	-1.51	-1.45	0.000	0.000	-1.61	-1.56
	5	3	0.000	0.000	-1.50	-1.45	0.000	0.000	-1.62	-1.56
	5	5	0.000	0.000	-1.36	-1.32	0.000	0.000	-1.44	-1.40
4	3	3	0.000	0.000	-1.88	-1.81	0.000	0.000	-2.06	-1.99
	3	5	0.000	0.000	-1.65	-1.61	0.000	0.000	-1.76	-1.72
	5	3	0.000	0.000	-1.66	-1.62	0.000	0.000	-1.76	-1.72
	5	5	0.000	0.000	-1.55	-1.52	0.000	0.000	-1.62	-1.59
5	3	3	0.000	0.000	-2.04	-1.97	0.000	0.000	-2.19	-2.13
	3	5	0.000	0.000	-1.86	-1.83	0.000	0.000	-1.95	-1.92
	5	3	0.000	0.000	-1.86	-1.83	0.000	0.000	-1.95	-1.92
	5	5	0.000	0.000	-1.78	-1.76	0.000	0.000	-1.85	-1.82

Table 7: Effects of Non-nesting Testing Regressions

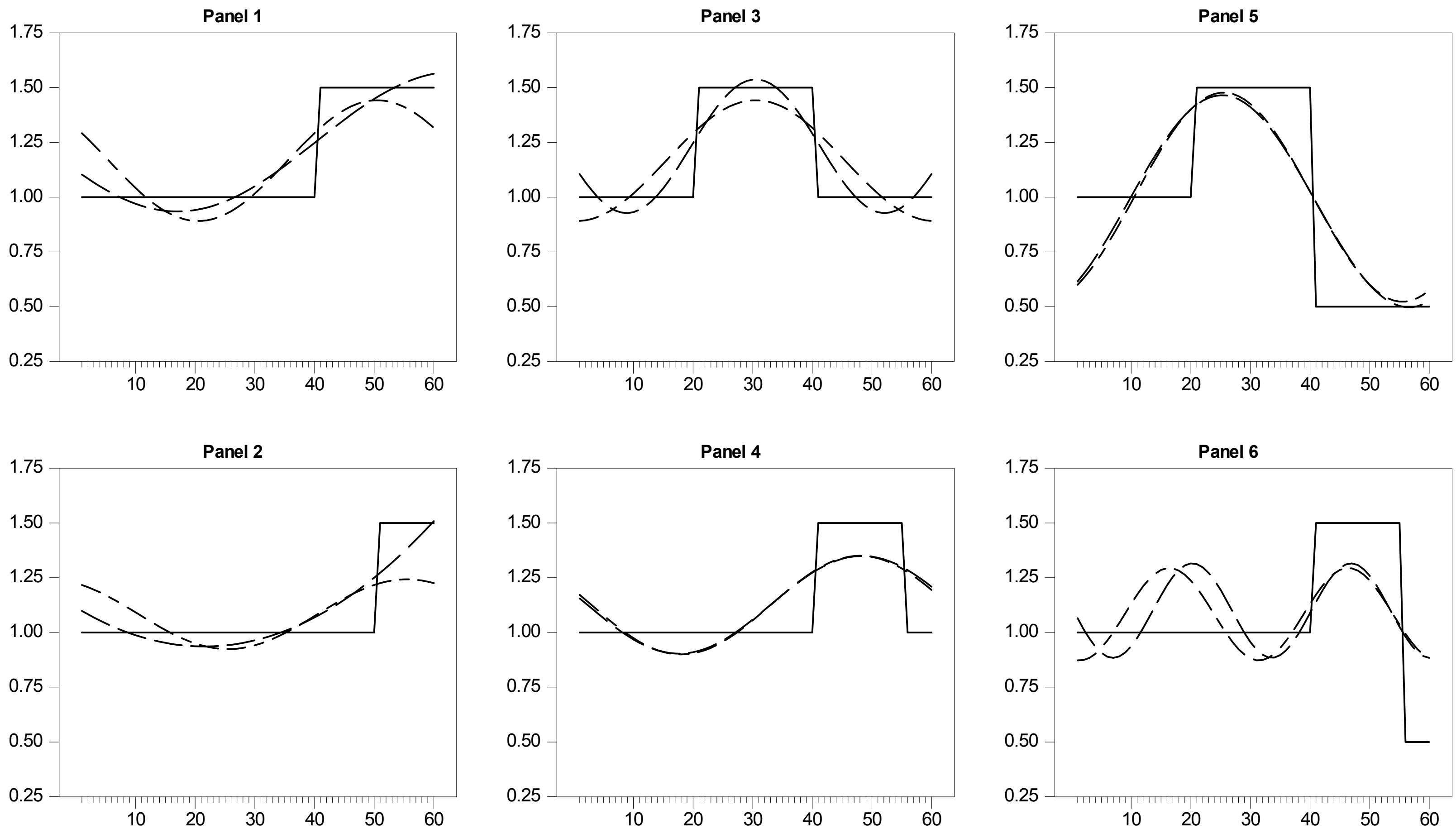
(a) $T = 100$, $\beta = 1$

DGP			5%	10%	Relative frequency of the selected values of k						
k	a_1	a_2	Rej. rate	Rej. rate	Linear*	1	2	3	4	5	6 ≥
1	0	0	0.056	0.104	0.99	0.01	0.00	0.00	0.00	0.00	0.00
	0	3	0.041	0.068	0.98	0.02	0.00	0.00	0.00	0.00	0.00
	0	5	0.042	0.059	0.95	0.05	0.00	0.00	0.00	0.00	0.00
2	0	0	0.059	0.111	0.99	0.01	0.00	0.00	0.00	0.00	0.00
	0	3	0.055	0.066	0.95	0.00	0.05	0.00	0.00	0.00	0.00
	0	5	0.183	0.187	0.81	0.00	0.19	0.00	0.00	0.00	0.00
3	0	0	0.062	0.114	0.99	0.01	0.00	0.00	0.00	0.00	0.00
	0	3	0.086	0.095	0.92	0.00	0.00	0.08	0.00	0.00	0.00
	0	5	0.253	0.255	0.75	0.00	0.00	0.25	0.00	0.00	0.00
4	0	0	0.059	0.111	0.99	0.01	0.00	0.00	0.00	0.00	0.00
	0	3	0.132	0.138	0.87	0.00	0.00	0.00	0.13	0.00	0.00
	0	5	0.392	0.394	0.61	0.00	0.00	0.00	0.39	0.00	0.00
5	0	0	0.053	0.103	0.99	0.01	0.00	0.00	0.00	0.00	0.00
	0	3	0.173	0.181	0.83	0.00	0.00	0.00	0.00	0.17	0.00
	0	5	0.573	0.573	0.43	0.00	0.00	0.00	0.00	0.57	0.00

(b) $T = 100$, $\beta = 0.9$

DGP			5%	10%	Relative frequency of the selected values of k						
k	a_1	a_2	Rej. rate	Rej. rate	Linear*	1	2	3	4	5	6 ≥
1	0	0	0.197	0.339	1.00	0.00	0.00	0.00	0.00	0.00	0.00
	0	3	0.039	0.062	0.98	0.02	0.00	0.00	0.00	0.00	0.00
	0	5	0.062	0.066	0.94	0.07	0.00	0.00	0.00	0.00	0.00
2	0	0	0.197	0.335	1.00	0.00	0.00	0.00	0.00	0.00	0.00
	0	3	0.105	0.116	0.90	0.00	0.10	0.00	0.00	0.00	0.00
	0	5	0.323	0.323	0.68	0.00	0.32	0.00	0.00	0.00	0.00
3	0	0	0.186	0.326	1.00	0.00	0.00	0.00	0.00	0.00	0.00
	0	3	0.181	0.184	0.82	0.00	0.00	0.18	0.00	0.00	0.00
	0	5	0.489	0.489	0.51	0.00	0.00	0.49	0.00	0.00	0.00
4	0	0	0.180	0.316	1.00	0.00	0.00	0.00	0.00	0.00	0.00
	0	3	0.263	0.266	0.74	0.00	0.00	0.00	0.26	0.00	0.00
	0	5	0.663	0.664	0.34	0.00	0.00	0.00	0.66	0.00	0.00
5	0	0	0.194	0.342	1.00	0.00	0.00	0.00	0.00	0.00	0.00
	0	3	0.390	0.393	0.61	0.00	0.00	0.00	0.00	0.39	0.00
	0	5	0.829	0.829	0.17	0.00	0.00	0.00	0.00	0.83	0.00

Figure 1: Level Breaks and a Fourier Approximation



The figures were generated using:

Panel 1: $y_t = 1\bullet(t \leq 40) + 1.5\bullet(t > 40)$;

Panel 2: $y_t = 1\bullet(t \leq 50) + 1.5\bullet(t > 50)$;

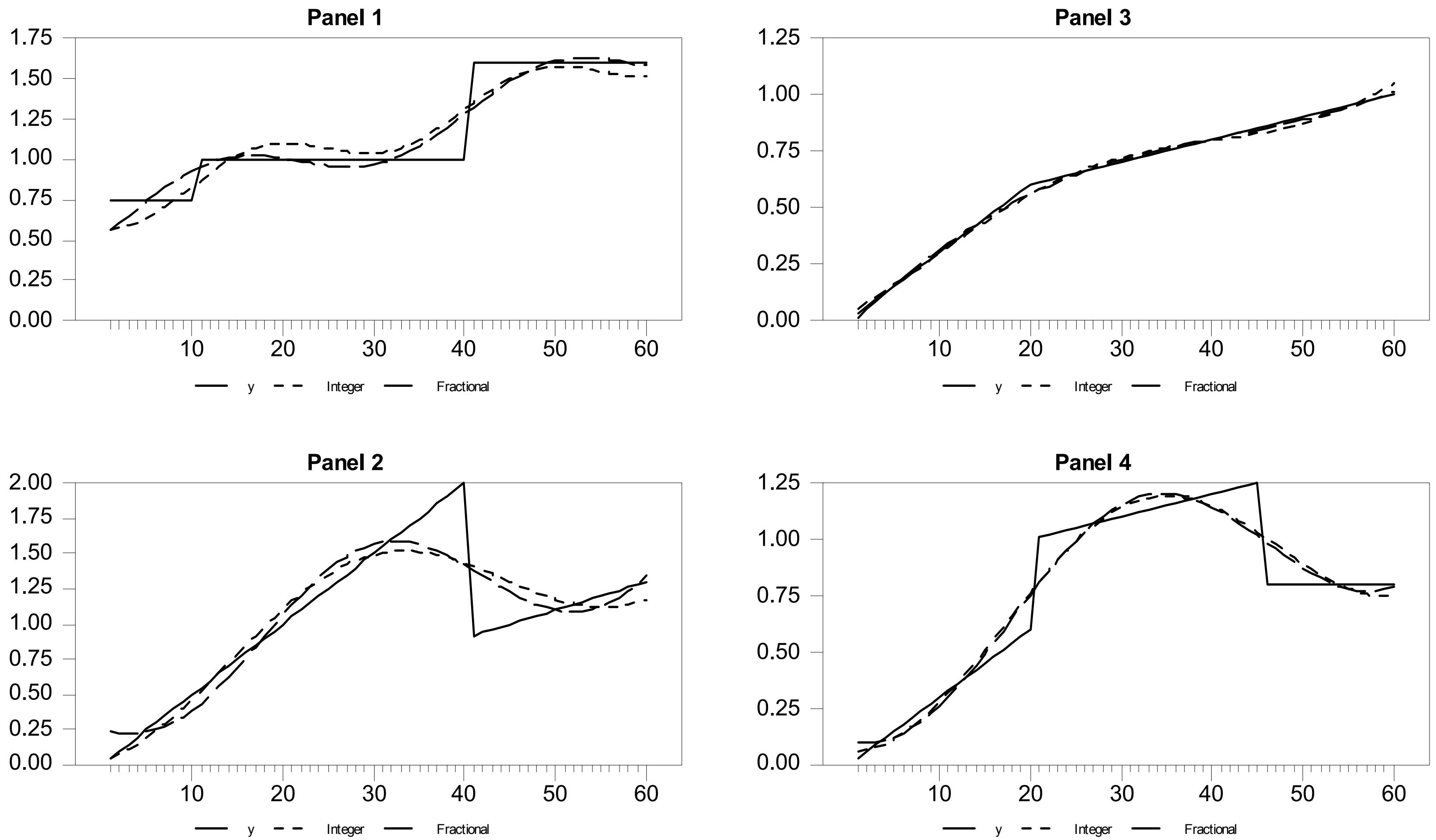
Panel 3: $y_t = 1\bullet(20 \leq t \text{ or } t > 40) + 1.5\bullet(20 < t \leq 40)$;

Panel 4: $y_t = 1\bullet(40 \leq t \text{ or } t > 55) + 1.5\bullet(40 < t \leq 55)$;

Panel 5: $y_t = 1\bullet(t \leq 20) + 1.5\bullet(20 < t \leq 40) + 0.5\bullet(t > 40)$;

Panel 6: $y_t = 1\bullet(t \leq 40) + 1.5\bullet(40 < t \leq 55) + 0.5\bullet(t > 55)$

Figure 2: Trend Breaks and a Fourier Approximation

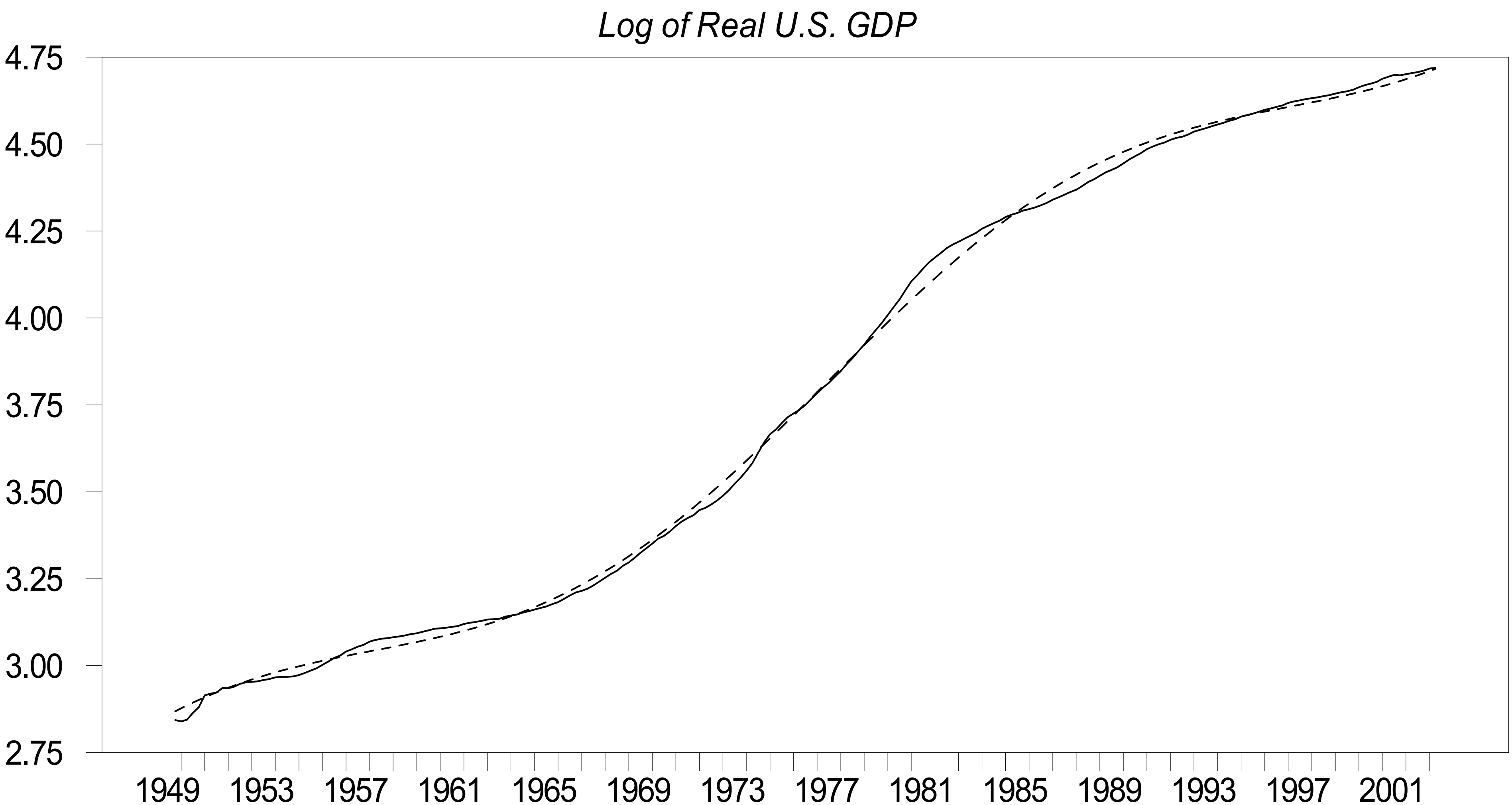


The figures were generated using:

Panel 1: $y_t = 0.75 \bullet (t \leq 10) + 1.0 \bullet (10 < t < 40) + 1.6 \bullet (t \geq 40)$; Panel 2: $y_t = 0.5t \bullet (t \leq 40) + (1.0 + 0.2t) \bullet (t > 40)$

Panel 3: $y_t = 0.03t \bullet (t \leq 20) + (0.4 + 0.1t) \bullet (t > 20)$; Panel 4: $y_t = 0.3t \bullet (t \leq 20) + (0.8 + 0.1t) \bullet (20 < t < 45) + 0.8 \bullet (t \leq 45)$

Figure 3: A Structural Break in U.S. GDP Growth



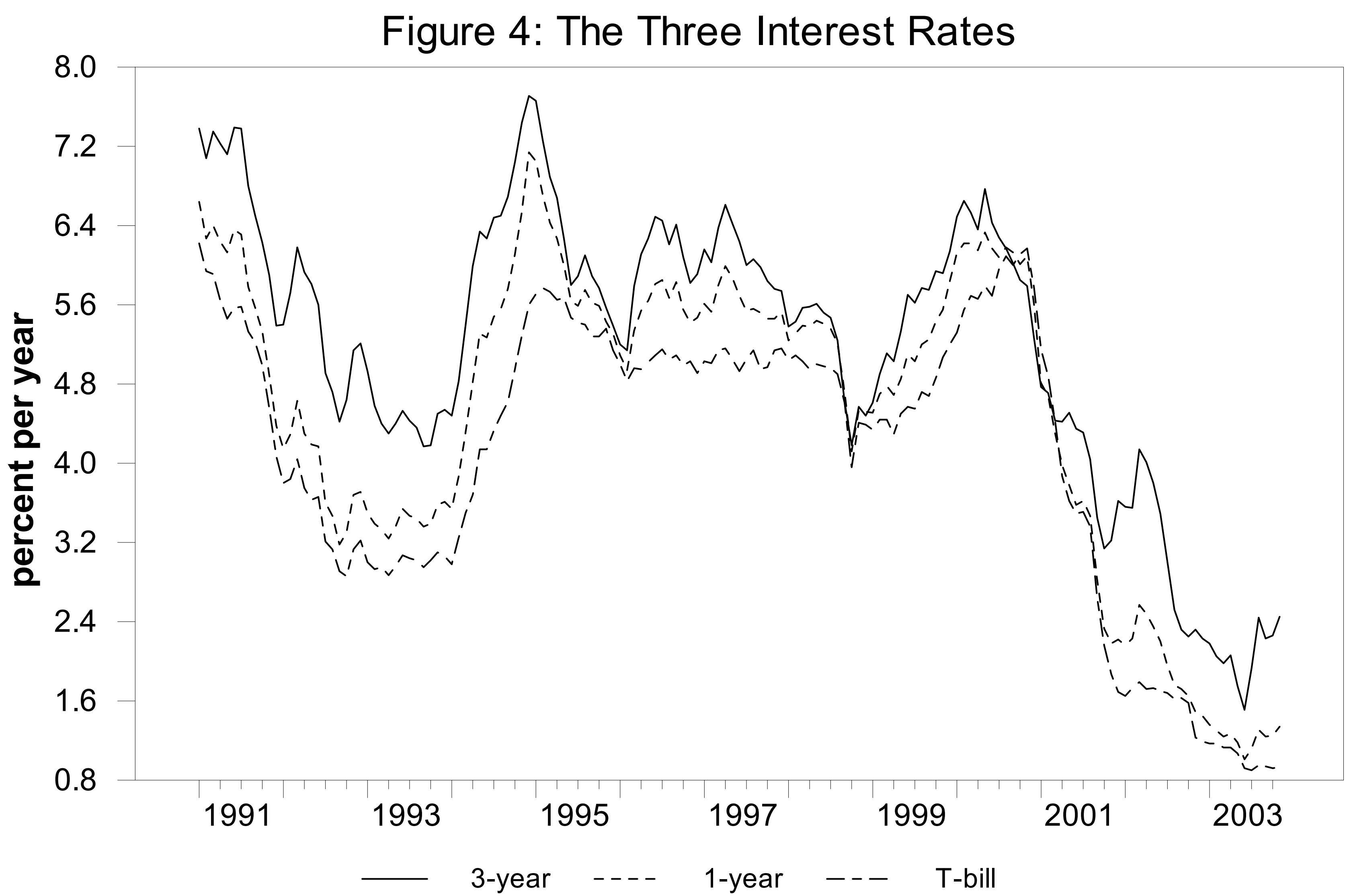
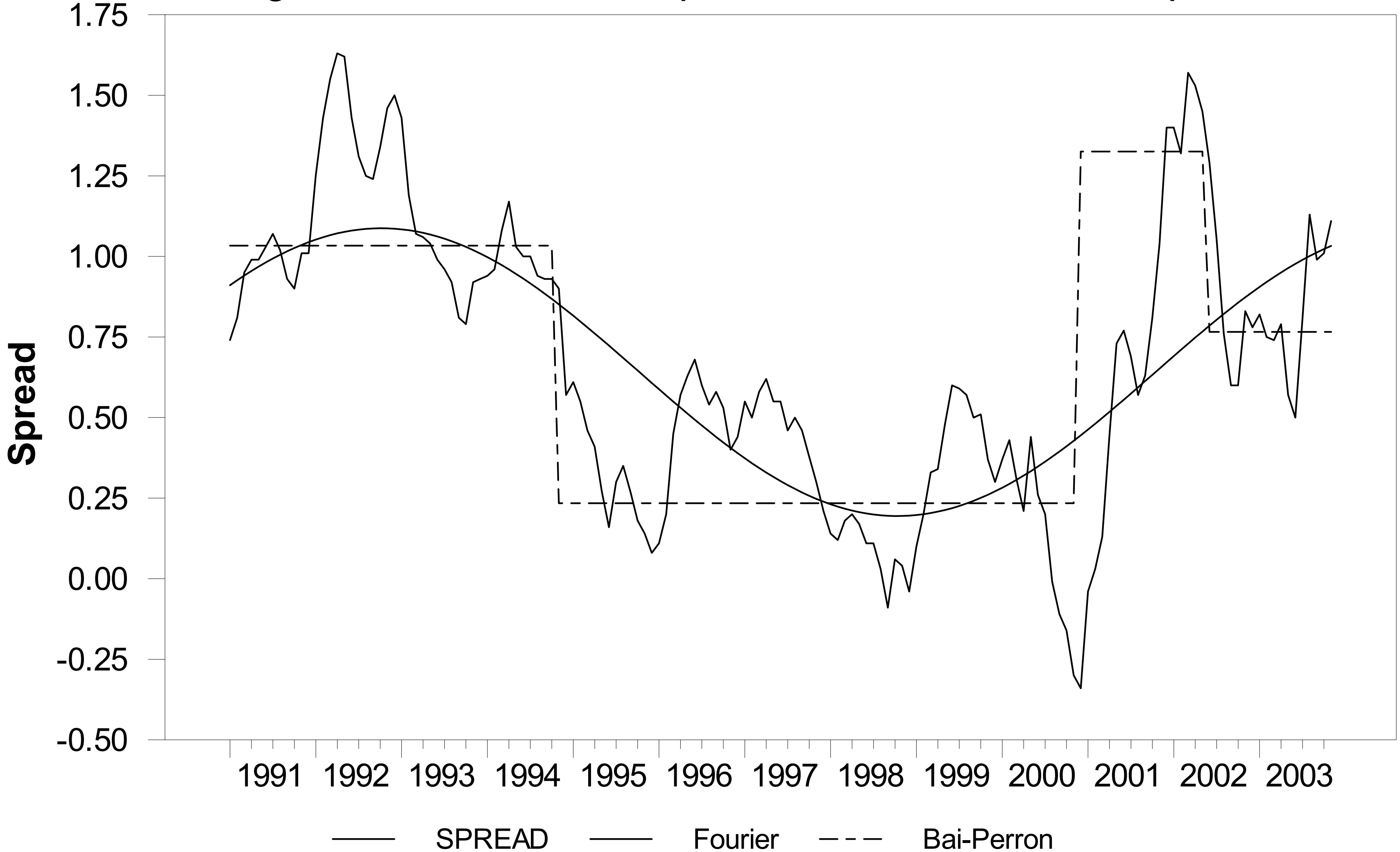


Figure 4: The Three Interest Rates
percent per year
8.0
7.2
6.4
5.6
4.8
4.0
3.2
2.4
1.6
0.8
1991
1993
1995
1997
1999
2001
2003
3-year
1-year
T-bill

Figure 5: The R3 - R1 Spread and the Two Intercepts



APPENDIX

Proof of Lemma 1

We start by examining the distribution of the estimated coefficients from the first step regression in (6). Denoting $\delta = (\delta_0, \delta_1, \delta_2)'$, and $\Delta Z_t = [1, \Delta\sin(2\pi kt/T), \Delta\cos(2\pi kt/T)]'$, and $D_T = diag[\sqrt{T}, 1/\sqrt{T}, 1/\sqrt{T}]$, we can have

$$D_T(\tilde{\delta} - \delta) = D_T(\Delta Z'\Delta Z)^{-1}\Delta Z'u = [D_T^{-1}(\Delta Z'\Delta Z)D_T^{-1}]^{-1}\cdot D_T^{-1}\Delta Z'u \qquad \textbf{(A.1)}$$

where $\Delta Z = (\Delta Z_2, .., \Delta Z_T)'$ and $u = (u_2, .., u_T)'$.

First, it is easy to show that

$$D_T^{-1}(\Delta Z'\Delta Z)D_T^{-1} = diag\Big[\frac{T\text{-}1}{T},\ T\sum_{t=2}^{T}\Delta\sin^2(2\pi kt/T),\ T\sum_{t=2}^{T}\Delta\cos^2(2\pi kt/T)\Big], \qquad \textbf{(A.2)}$$

for which all off-diagonal terms are zero due to the orthogonality property that $\sum_{t=2}^{T}\Delta\sin(2\pi kt/T)\Delta\cos(2\pi kt/T) = 0$, and $\sum_{t=2}^{T}\Delta\sin(2\pi kt/T) = \sum_{t=2}^{T}\Delta\cos(2\pi kt/T) = 0$. Since $\Delta\sin(2\pi kt/T) = (2\pi k/T)\cos(2\pi kt/T)$, and $\Delta\cos(2\pi kt/T) = -(2\pi k/T)\sin(2\pi kt/T)$, we can show

$$T\sum_{t=2}^{T}\Delta\sin^2(2\pi kt/T) \rightarrow (2\pi k)^2\int_0^1 \cos^2(2\pi kr)dr \qquad \textbf{(A.3)}$$

$$T\sum_{t=2}^{T}\Delta\cos^2(2\pi kt/T) \rightarrow (2\pi k)^2\int_0^1 \sin^2(2\pi kr)dr \qquad \textbf{(A.4)}$$

Second, we have

$$D_T^{-1}\Delta Z'u = \Big[\frac{1}{\sqrt{T}}\sum_{t=2}^{T}u_t,\ \sqrt{T}\sum_{t=2}^{T}u_t\Delta\sin(2\pi kt/T),\ \sqrt{T}\sum_{t=2}^{T}u_t\Delta\cos(2\pi kt/T)\Big]'.$$

We note that $\frac{1}{\sqrt{T}}\sum_{t=2}^{T}u_t \rightarrow \sigma W(1)$, which is a standard result. For the second and third terms, we need to utilize the following asymptotics in Proposition 1. Then, combining the above results in (A.2) ~ (A.4), we can obtain the results in Lemma 1. □

Proposition 1

$$\sqrt{T}\sum_{t=2}^{T}u_t\Delta\sin(2\pi kt/T) \rightarrow \sigma(2\pi k)[W(1) + (2\pi k)\int_0^1 \sin(2\pi kr)W(r)dr] \qquad \textbf{(A.5)}$$

$$\sqrt{T}\sum_{t=2}^{T}u_t\Delta\cos(2\pi kt/T) \rightarrow \sigma(2\pi k)^2[\int_0^1 \cos(2\pi kr)W(r)dr] \qquad \textbf{(A.6)}$$

Proof: We employ the result in Bierens (1994, Lemma 9.6.3): $\sum_{t=2}^{T}F(t/T)u_t = F(1)S_T(1) -$

$\int_0^1 f(r)S_T(r)dr$, where $f(r)$ is $F'(r)$. For (A.5), we choose $F(x) = \cos(2\pi kt/T)$. Then, we can show $F(1)S_T(1) - \int_0^1 f(r)S_T(r)dr = \sigma[W(1) + (2\pi k)\int_0^1 \sin(2\pi kr)W(r)dr]$. For (A.6), we choose $F(x) = \sin(2\pi kt/T)$, and follow the similar procedure to obtain the desired result. □

Proof of Theorem 1

We let $S_t = \sum_{j=2}^{t} \varepsilon_j$ and $[rT]$ be the integer part of rT, $r \in [0,1]$. Then, it is easy to show that the expression $\tilde{S}_t$ in (8) can be given as follows:

$$\frac{1}{\sqrt{T}}\tilde{S}_{[rT]} = \frac{1}{\sqrt{T}}S_{[rT]} - \frac{1}{\sqrt{T}}(\tilde{\delta}_0 - \delta_0)\, rT - \frac{1}{\sqrt{T}}(\tilde{\delta}_1 - \delta_1)\sin(2\pi krT/T) - \frac{1}{\sqrt{T}}(\tilde{\delta}_2 - \delta_2)\cos(2\pi krT/T)$$

$$\rightarrow \sigma\, V(r) = \sigma\Big\{W(r) - rW(1) - [(2\pi k)\int_0^1 \cos^2(2\pi kr)dr]^{-1}[W(1) + (2\pi k)\int_0^1 \sin(2\pi kr)W(r)dr]\cdot\sin(2\pi kr) - [\int_0^1 \sin^2(2\pi kr)dr]^{-1} [\int_0^1 \cos(2\pi kr)W(r)dr]\cdot\cos(2\pi kr)\Big\} \qquad \textbf{(A.7)}$$

Now, from the second step regression (8), we obtain:

$$\tilde{\phi} = (\tilde{S}_1{}' M_{\Delta Z}\, \tilde{S}_1)^{-1}\,(\tilde{S}_1{}' M_{\Delta Z}\, \Delta y), \qquad \textbf{(A.8)}$$

where $\tilde{S}_1 = (\tilde{S}_1, .., \tilde{S}_{T-1})'$, $\Delta Z = (\Delta Z_2, .., \Delta Z_T)'$, $\Delta y = (\Delta y_2, .., \Delta y_T)'$, and $M_{\Delta Z} = I - \Delta Z(\Delta Z'\Delta Z)^{-1}\Delta Z'$.

From the results in (A.7), we have:

$$T^{-2}\, \tilde{S}_1{}' M_{\Delta Z}\, \tilde{S}_1 \rightarrow \sigma^2 \int_0^1 \underline{V}(r)^2 dr, \qquad \textbf{(A.9)}$$

where $\underline{V}(r)$ is the projection of the process $V(r)$ on the orthogonal complement of the space spanned by $dz = (1,\, d\sin(2\pi kr),\, d\cos(2\pi kr))'$ where $r \in [0, 1]$. That is,

$$\underline{V}(r) = V(r) - \tilde{\delta}'dz,$$

with

$$\tilde{\delta} = \underset{\delta}{argmin} \int_0^1 (V(r) - \delta' dz)^2 dr \ .$$

Following SP, we can similarly show that for the second term in (A.8):

$$\frac{1}{T}\tilde{S}_1{'}M_{\Delta Z}\Delta y = \frac{1}{T}\tilde{S}_1{'}M_{\Delta Z}\,\varepsilon = \frac{1}{T}\tilde{S}_1{'}\,\underline{\varepsilon} \;\rightarrow\; -0.5\sigma_\varepsilon^2\;, \qquad \textbf{(A.10)}$$

where $\underline{\varepsilon} = M_{\Delta Z}\,\varepsilon$. Theorem 1 is thus proved by combining the results in (A.9) - (A.10). □

Proof of Lemma 2

We now obtain the asymptotic distribution of the F-test in (13) that is based on the testing regression for the LM type statistic. First, we examine SSR_0 which is obtained from the (restricted) regression with a linear trend for the usual LM statistic.

$$\Delta y_t = \hat{\phi}\,\hat{S}_{t-1} + \hat{d}_0 + \hat{u}_t,$$

where $\hat{S}_{t-1} = (y_t - y_1) - \hat{\gamma}(t\text{-}1)$ and $\hat{\gamma} = (1/T)\sum_{t=2}^{T}\Delta y_t$, and where $\hat{\phi}$ and $\hat{d}_0$ are the OLS estimates in this regression. Then, when a unit root assumption is imposed, we get:

$$SSR_0 = \sum_{t=2}^{T}\hat{u}_t^2 = \sum_{t=2}^{T}(\varepsilon_t - \hat{d}_0 - \hat{\phi}\hat{S}_{t-1})^2$$

which can be expressed as

$$\sum_{t=2}^{T}[(\varepsilon_t - \bar{\varepsilon}) - \hat{\phi}(\hat{S}_{t-1} - \bar{S}_1)]^2$$
$$= \sum_{t=2}^{T}(\varepsilon_t - \bar{\varepsilon})^2 + \hat{\phi}^2\sum_{t=2}^{T}(\hat{S}_{t-1} - \bar{S}_1)^2 - 2\hat{\phi}\sum_{t=2}^{T}(\hat{S}_{t-1} - \bar{S}_1)(\varepsilon_t - \bar{\varepsilon}) \qquad \textbf{(A.11)}$$

where $\bar{\varepsilon} = \frac{1}{T}\sum_{t=1}^{T}\varepsilon_t$, and $\bar{S}_1 = \frac{1}{T\text{-}1}\sum_{t=2}^{T}\hat{S}_{t-1}$. The first term in (A.11) will be cancelled with the same term that appears in SSR_1 under the null of linearity. The second term in the above expression can be written as

$$(T\hat{\phi})^2T^{-2}\sum_{t=2}^{T}(\hat{S}_{t-1} - \bar{S}_1)^2 \rightarrow [-\frac{1}{2}(\sigma_\varepsilon^2/\sigma^2)\,(\int_0^1 \underline{V}_0(r)^2dr)^{-1}]^2(\sigma^2\int_0^1 \underline{V}_0(r)^2dr)$$
$$= \frac{1}{4}(\sigma_\varepsilon^4/\sigma^2)\,(\int_0^1 \underline{V}_0(r)^2dr)^{-1} \qquad \textbf{(A.12)}$$

where $\underline{V}_0(r)$ is the demeaned Brownian bridge. The third term in (A.11) is shown to follow

$$-2(T\hat{\phi})\cdot(1/T)\sum_{t=2}^{T}(\hat{S}_{t-1} - \bar{S}_1)(\varepsilon_t - \bar{\varepsilon}) \rightarrow -2\,[-\frac{1}{2}(\sigma_\varepsilon^2/\sigma^2)\,(\int_0^1 \underline{V}_0(r)^2dr)^{-1}](-\frac{1}{2}\sigma_\varepsilon^2)$$
$$= -\frac{1}{2}(\sigma_\varepsilon^4/\sigma^2)\,(\int_0^1 \underline{V}_0(r)^2dr)^{-1} \qquad \textbf{(A.13)}$$

Next, SSR_1 is similarly obtained from the unrestricted regression (8).

$$\Delta y_t = \tilde{\phi}\,\tilde{S}_{t-1} + \tilde{d}_0 + \tilde{d}_1\,\Delta\sin(2\pi kt/T) + \tilde{d}_2\,\Delta\cos(2\pi kt/T) + \tilde{u}_t.$$

where $\tilde{S}_{t-1}$ is defined in (7), and where $\tilde{d}_0$, $\tilde{d}_1$, and $\tilde{d}_2$ are the OLS estimates in this regression.

Then, we get under the null of a unit root:

$$SSR_1 = \sum_{t=2}^{T} \tilde{u}_t^2 = \sum_{t=2}^{T} (\varepsilon_t - \tilde{d}_0 - \tilde{d}_1 \Delta\sin(2\pi kt/T) - \tilde{d}_2 \Delta\cos(2\pi kt/T) - \tilde{\phi}\, \tilde{S}_{t-1})^2$$

which can be expressed as

$$\sum_{t=2}^{T} (\underline{\varepsilon}_t - \tilde{\phi}\, \underline{\tilde{S}}_{t-1})^2$$

$$= \sum_{t=2}^{T} \underline{\varepsilon}_t^{\,2} + \tilde{\phi}^{\,2} \sum_{t=2}^{T} \underline{\tilde{S}}_{t-1}^{\,2} - 2\tilde{\phi} \sum_{t=2}^{T} \underline{\tilde{S}}_{t-1} \underline{\varepsilon}_t \qquad \textbf{(A.14)}$$

where $\underline{\varepsilon}_t$ is the element of $\underline{\varepsilon} = M_{\Delta Z}\, \varepsilon$, which is given in (A.10), and $\underline{\tilde{S}}_{t-1}$ is the element of $\underline{\tilde{S}}_1 = M_{\Delta Z}\, \tilde{S}_1$ with $\tilde{S}_1 = (\tilde{S}_1, .., \tilde{S}_{T-1})'$. The first term in (A.14) is cancelled with the similar term in SSR_0 under the null of the absence of the nonlinear terms. The second and third terms in (A.14) follow the same asymptotics as in (A.12) and (A.13), except that $\underline{V}_0(r)$ is replaced with $\underline{V}(r)$, which is defined in (A.9).

Finally, the denominator of the F-statistic is given as

$$\frac{1}{T\text{-}q} \sum_{t=2}^{T} \tilde{u}_t^2 + \frac{1}{T\text{-}q}\, O_p(1) \rightarrow \sigma_\varepsilon^2 \qquad \textbf{(A.15)}$$

where the $O_p(1)$ terms of the above expression are the same terms in (A.14). The asymptotic distribution of the F-statistic is given by collecting terms in (A.12) through (A.17).

$$F(k) \rightarrow \frac{1}{8}(\sigma_\varepsilon^2/\sigma^2)[(\int_0^1 \underline{V}(r)^2 dr)^{-1} - (\int_0^1 \underline{V}_0(r)^2 dr)^{-1}] \qquad \square$$

Proof of Lemma 3

The DGP implies (5), which includes the nonlinear trigonometric terms, but they are ignored in the testing regressions. Thus, the first step regression is

$$\Delta y_t = \delta_0 + u_t$$

Then, we can show that the OLS estimate of δ_0 follows:

$$\hat{\delta}_0 = \text{mean of } \Delta y = \delta_0 + \frac{1}{T}\sum_{t=2}^{T} \delta_1 \Delta\sin(2\pi kt/T) + \frac{1}{T}\sum_{t=2}^{T} \delta_2 \Delta\cos(2\pi kt/T) + \frac{1}{T}\sum_{t=2}^{T} \Delta e_t$$

$$= \delta_0 + 0 + 0 + \overline{\Delta e} \qquad \textbf{(A.16)}$$

We construct a detrended series using $\hat{\delta}_0$ as:

$$\hat{S}_t = y_t - y_1 - \tilde{\delta}_0 (t\text{-}1) = y_t - y_1 - \delta_0 (t\text{-}1) - (\tilde{\delta}_0 - \delta_0)(t\text{-}1)$$

$$= (e_t - e_1) + a_1 \sin(2\pi kt/T) + a_2 \cos(2\pi kt/T) - a_1 \sin(2\pi k/T) - a_2 \cos(2\pi k/T)$$
$$- (t-1)\,\overline{\Delta e}\,.$$

The second step regression involves

$$\Delta y_t = \phi\,\hat{S}_{t-1} + c + u_t$$

For simplicity, we ignore the constant term, which is 0 in the population. Then, for the denominator of $\hat{\phi}$, we have:

$$\frac{1}{T}\sum_{t=2}^{T} \hat{S}_{t-1}^{\,2} \approx \frac{1}{T}\sum_{t=2}^{T} \hat{S}_t^{\,2}$$
$$= \frac{1}{T}\sum_{t=2}^{T} \Big[(e_t - e_1) + a_1\sin(2\pi kt/T) + a_2 \cos(2\pi kt/T) - a_1\sin(2\pi k/T) - a_2 \cos(2\pi k/T)$$
$$- (t-1)\,\overline{\Delta e}\,\Big]^2$$

This can be expressed as:

$$\frac{1}{T}\sum_{t=2}^{T} [(e_t - e_1)- (t-1)\,\overline{\Delta e}]^2 + \frac{1}{T}\sum_{t=2}^{T} [a_1 \sin(2\pi kt/T) + a_2 \cos(2\pi kt/T)]^2$$
$$+ \frac{1}{T}\sum_{t=2}^{T} [- a_1 \sin(2\pi k/T) - a_2 \cos(2\pi k/T)]^2 + 2\frac{1}{T}\sum_{t=2}^{T} [(e_t - e_1)- (t-1)\,\overline{\Delta e}]\,[a_1 \sin(2\pi kt/T)$$
$$+ a_2 \cos(2\pi kt/T)] + 2\frac{1}{T}\sum_{t=2}^{T} [(e_t - e_1) - (t-1)\,\overline{\Delta e}]\,[-a_1 \sin(2\pi k/T) - a_2 \cos(2\pi k/T)]$$
$$+ 2\frac{1}{T}\sum_{t=2}^{T} [\, a_1 \sin(2\pi kt/T) + a_2 \cos(2\pi kt/T)][-a_1 \sin(2\pi k/T) - a_2 \cos(2\pi k/T)] \qquad \textbf{(A.17)}$$

For each term, we can show:

$$\frac{1}{T}\sum_{t=2}^{T} [(e_t - e_1)- (t-1)\,\overline{\Delta e}]^2 \to \sigma_\varepsilon^2 + (1/3)(\varepsilon_\infty^2+\varepsilon_1\varepsilon_\infty+\varepsilon_1^2) \qquad \textbf{(A.18a)}$$

$$\frac{1}{T}\sum_{t=2}^{T} [a_1 \sin(2\pi kt/T) + a_2 \cos(2\pi kt/T)]^2$$
$$\to a_1^2 \int_0^1 \sin^2(2\pi kr)dr + a_2^2 \int_0^1 \cos^2(2\pi kr)dr \qquad \textbf{(A.18b)}$$

$$\frac{1}{T}\sum_{t=2}^{T} [- a_1 \sin(2\pi k/T) - a_2 \cos(2\pi k/T)]^2 \to a_2^2 \qquad \textbf{(A.18c)}$$

$$2\frac{1}{T}\sum_{t=2}^{T} [(e_t - e_1)- (t-1)\,\overline{\Delta e}]\,[a_1 \sin(2\pi kt/T) + a_2 \cos(2\pi kt/T)]$$
$$= 2\frac{1}{T}\sum_{t=2}^{T} (e_t - e_1)a_1 \sin(2\pi kt/T) + 2\frac{1}{T}\sum_{t=2}^{T} (e_t - e_1)\, a_2 \cos(2\pi kt/T)$$
$$- 2\frac{1}{T}\sum_{t=2}^{T} (t-1)\overline{\Delta e}\cdot a_1 \sin(2\pi kt/T) - 2\frac{1}{T}\sum_{t=2}^{T} (t-1)\,\overline{\Delta e}\cdot a_2 \cos(2\pi kt/T)$$
$$\to 0 + 0 - 2\sigma W(1)a_1 \int_0^1 r\cdot\sin(2\pi kr)dr - 2\sigma W(1)a_2 \int_0^1 r\cdot\cos(2\pi kr)dr \qquad \textbf{(A.18d)}$$

In the above, the first and the second terms in (A.20d) are degenerate since $\frac{1}{\sqrt{T}}\sum_{t=2}^{T} e_t \sin(2\pi kt/T)$ $\rightarrow \sigma\int_0^1 \cos(2\pi kr)W(r)dr,$ and $\frac{1}{\sqrt{T}}\sum_{t=2}^{T} e_t \cos(2\pi kt/T) \rightarrow \sigma W(1) + \sigma(2\pi k)\int_0^1 \sin(2\pi kr)W(r)dr.$ The results for the third and fourth terms follow as given in the above since $T\overline{\Delta e} \rightarrow \sigma W(1)$, $\frac{1}{T}\sum_{t=2}^{T}$ $(t/T)\sin(2\pi kt/T) \rightarrow \sigma\int_0^1 r\sin(2\pi kr)dr,$ and $\frac{1}{T}\sum_{t=2}^{T}(t/T)\cos(2\pi kt/T) \rightarrow \sigma\int_0^1 r\cos(2\pi kr)dr.$ Thus, the asymptotic distribution of the numerator of $\hat{\phi}$ is given by collecting the terms in (A.18a) - (A.18d).

Next, we examine the numerator of $\hat{\phi}$. Since $\Delta\hat{S}_t = \Delta e_t - \overline{\Delta e} + a_1\Delta \sin(2\pi kt/T) + a_2\Delta\cos(2\pi kt/T)$, we have:

$$\frac{1}{T}\sum_{t=2}^{T}\Delta\hat{S}_t\hat{S}_{t-1} = \frac{1}{T}\sum_{t=2}^{T}\big[(\Delta e_t - \overline{\Delta e} + a_1\Delta\sin(2\pi kt/T) + a_2\Delta\cos(2\pi kt/T)\big]\big[(e_t - e_1)$$
$$+ a_1\sin(2\pi kt/T) + a_2\cos(2\pi kt/T) - a_1\sin(2\pi k/T) - a_2\cos(2\pi k/T) - (t\text{-}1)\,\overline{\Delta e}\big]$$
$$= \frac{1}{T}\sum_{t=2}^{T}\Delta e_t - \overline{\Delta e})\,\big[(e_{t-1} - e_1) - (t\text{-}2)\,\overline{\Delta e}\big]$$
$$+ \frac{1}{T}\sum_{t=2}^{T}(\Delta e_t - \overline{\Delta e})\,\big[a_1\sin(2\pi kt/T) + a_2\cos(2\pi kt/T) - a_1\sin(2\pi k/T) - a_2\cos(2\pi k/T)\big]$$
$$+ \frac{1}{T}\sum_{t=2}^{T}\big[a_1\Delta\sin(2\pi kt/T) + a_2\Delta\cos(2\pi kt/T)\big]\big[a_1\sin(2\pi kt/T) + a_2\cos(2\pi kt/T)$$
$$- a_1\sin(2\pi k/T) - a_2\cos(2\pi k/T)\big]$$
$$+ \frac{1}{T}\sum_{t=2}^{T}\big[a_1\Delta\sin(2\pi kt/T) + a_2\Delta\cos(2\pi kt/T)\big]\big[(e_{t-1} - e_1) - (t\text{-}2)\,\overline{\Delta e}\big] \qquad \textbf{(A.19)}$$

The first term in the last equation in (A.19) follows:

$$\frac{1}{T}\sum_{t=2}^{T}(\Delta e_t - \overline{\Delta e})\,\big[(e_{t-1} - e_1) - (t\text{-}2)\,\overline{\Delta e}\big] = \frac{1}{T}\sum_{t=2}^{T}\Delta e_t(e_{t-1} - e_1) - \frac{1}{T}\sum_{t=2}^{T}\overline{\Delta e}\,(e_{t-1} - e_1)$$
$$+ \frac{1}{T}\sum_{t=2}^{T}\Delta e_t\,(t\text{-}2)\,\overline{\Delta e} + \frac{1}{T}\sum_{t=2}^{T}\overline{\Delta e}\,(t\text{-}2)\,\overline{\Delta e}$$
$$\rightarrow \sigma_\varepsilon^2(\beta\text{-}1) + o_p(1)$$

For the above result, we note $\frac{1}{T}\sum_{t=2}^{T}\Delta e_t(e_{t-1} - e_1) \approx \frac{1}{T}\sum_{t=2}^{T}(e_{t-1} - e_1)e_{t-1}$ since $e_1\frac{1}{T}\sum_{t=2}^{T}\Delta e_t = e_1\overline{\Delta e}$ $\rightarrow 0.$ Also, $\frac{1}{T}\sum_{t=2}^{T}(e_{t-1} - e_1)e_{t-1} \rightarrow \gamma_1 - \sigma_\varepsilon^2 = \sigma_\varepsilon^2(\beta\text{-}1)$ where β is the AR coefficient in the DGP (5). It can be easily seen that all remaining terms are degenerate. The second term in the last equation in (A.19) can be shown as $o_p(1)$ by employing the results that $\frac{1}{\sqrt{T}}\sum_{t=2}^{T} e_t \sin(2\pi kt/T) \rightarrow$ $\sigma\int_0^1 \cos(2\pi kr)W(r)dr$, and $\frac{1}{\sqrt{T}}\sum_{t=2}^{T} e_t \sin(2\pi kt/T) \rightarrow \sigma\left[W(1)+(2\pi k)\int_0^1 \sin(2\pi kr)W(r)dr\right.$. The third and fourth terms in the last equation in (A.19) can be shown as $o_p(1)$ by utilizing these results.

Thus,

$$\frac{1}{T}\sum_{t=2}^{T} \Delta\hat{S}_t\hat{S}_{t-1} \rightarrow \sigma_\varepsilon^2(\beta - 1) \qquad \textbf{(A.20)}$$

The result in Lemma 3 follows by combining the results in (A.18) and (A.20).