NBER TECHNICAL PAPER SERIES

SEASONAL UNIT ROOTS IN AGGREGATE U.S. DATA

J. Joseph Beaulieu

Jeffrey A. Miron

Technical Paper No. 126

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 1992

We thank Ed Glaeser for research assistance, David Cutler, David Dickey, Eric Ghysels, Danny Quah, Jeff Wooldridge and three anonymous referees for comments, Sara Fisher-Ellison for help with density estimation, and Todd Clark for help with data. Miron is grateful for financial support from NSF Grant SES-8921506, a John M. Olin Fellowship from the National Bureau of Economic Research, and a Sloan Foundation Fellowship. Beaulieu acknowledges financial support from an NSF Graduate Fellowship. The Views expressed are not necessarily those of the Board of Governors or the staff of the Federal Reserve System. This paper is part of NBER's research program in Economic Fluctuations. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Technical Paper #126
August 1992

SEASONAL UNIT ROOTS IN AGGREGATE U.S. DATA

ABSTRACT

In this paper we provide evidence on the presence of seasonal unit roots in aggregate U.S. data. The analysis is conducted using the approach developed by Hylleberg, Engle, Granger and Yoo (1990). We first derive the mechanics and asymptotics of the HEGY procedure for monthly data and use Monte Carlo methods to compute the finite sample critical values of the associated test statistics. We then apply quarterly and monthly HEGY procedures to aggregate U.S. data. The data reject the presence of unit roots at most seasonal frequencies in a large fraction of the series considered.

Jeffrey A. Miron
Departments of Economics
Boston University
270 Bay State Road
Boston, MA 02115
and NBER

J. Joseph Beaulieu
Board of Governors of
the Federal Reserve
Washington, D.C. 20551

1. Introduction

The analysis of seasonal variation in economic times series is almost as old as macroeconomics itself (Mitchell (1927), Burns and Mitchell (1946), Macaulay (1938)). Despite this long history, however, there is little consensus on how seasonality should be treated in empirical research on aggregate fluctuations. The specification of seasonality varies significantly from one paper to the next, and few researchers provide an explicit justification for their handling of seasonal variation. Since the statistical properties of different seasonality models are distinct, and since seasonality in quantitatively important in many aggregate series, the imposition of one kind when another is present can result in serious biases or loss of information. It is therefore useful to establish what kind of seasonality is present in the data.

In this paper we provide evidence on the presence of seasonal unit roots in aggregate U.S. data. Of the three main definitions of seasonality offered in the literature (deterministic dummies, non-stationary stochastic seasonality due to seasonal unit roots, and stationary stochastic seasonality), it is the non-stationarity due to seasonal unit roots that raises the most troubling statistical issues. In addition, the investigation of seasonal unit roots logically precedes the examination of other kinds of seasonality, since such examinations can produce spurious results if seasonal unit roots are present but not accounted for. The analysis of seasonal integration also logically precedes the analysis of cointegration.

Our investigation of seasonal unit roots is conducted using the approach developed by Hylleberg, Engle, Granger and Yoo (1990) (HEGY); this is a general procedure that can test for unit roots at some seasonal frequencies without maintaining that unit roots are present at all seasonal frequencies. We first derive the mechanics of the HEGY procedure for monthly data and use Monte Carlo methods to compute the finite sample critical values of the associated test statistics. Franses (1990) provides a similar analysis. We then apply the HEGY procedure to both monthly and quarterly data on a number of aggregate U.S. data series. The data reject the presence of unit roots at most of the seasonal frequencies for a large fraction of the series considered.

The remainder of the paper is organized as follows. Section 2 reviews the testing procedure developed in the HEGY paper by deriving it for monthly data and comparing it to earlier approaches. This section also presents Monte Carlo finite sample critical values for the case of monthly data, supplementing the quarterly Monte Carlo results in HEGY. In Section 3 we apply the HEGY procedure to quarterly and monthly aggregate U.S. data. Section 4 concludes by discussing implications of the results.

2. Testing for Seasonal Unit Roots in Monthly Data

A recent paper by HEGY (1990) explains how to test for seasonal unit roots in processes that may also exhibit deterministic or stationary stochastic seasonality. The paper also shows how to test for a unit root at frequency zero when unit roots may be present at some or all of the seasonal frequencies. In this section we review the HEGY procedure by deriving its mechanics and the asymptotic distributions of its test statistics in the case of monthly rather than quarterly data. We then use Monte Carlo simulations to tabulate the finite sample distributions of test statistics for hypotheses about unit roots at seasonal and zero frequencies in monthly data and to investigate the size and power properties of the test under various data generating mechanisms.

2.1 The HEGY Test Procedure

Let x_t be the series of interest, generated by a general autoregression of the form

$$\varphi(B)x_t = \epsilon_t, \tag{1}$$

where $\varphi(B)$ is a polynomial in the backshift operator and ϵ_t is a white noise process. Let γ_k be the roots of the characteristic polynomial associated with $\varphi(B)$. Assume for the moment that deterministic terms, such as seasonal dummies or time trends, are known to be absent from the process for x_t. In general, some or all of the γ_k may be complex.

The frequency associated with a particular root is the value of α in $e^{\alpha i}$, the polar representation of the root. A root is seasonal if $\alpha = \frac{2\pi j}{S}$, $j = 1, \ldots, S-1$, where S is the number of observations per year. For monthly data, the seasonal unit roots are

$$-1;\ \pm i;\ -\tfrac{1}{2}(1 \pm \sqrt{3}i);\ \tfrac{1}{2}(1 \pm \sqrt{3}i);\ -\tfrac{1}{2}(\sqrt{3} \pm i);\ \tfrac{1}{2}(\sqrt{3} \pm i) \tag{2}$$

with these roots corresponding to 6, 3, 9, 8, 4, 2, 10, 7, 5, 1, and 11 cycles per year, respectively. The frequencies of these roots are $\pi, \pm\frac{\pi}{2}, \mp\frac{2\pi}{3}, \pm\frac{\pi}{3}, \mp\frac{5\pi}{6}$ and $\pm\frac{\pi}{6}$, respectively. We wish to know whether the polynomial in the backshift operator, $\varphi(B)$, has roots equal to one in absolute value at the zero or seasonal frequencies. In particular, the goal is to test hypotheses about a particular unit root without taking a stand on whether other seasonal or zero frequency unit roots may be present.

The testing procedure developed by HEGY (see Appendix A for details) consists essentially of linearizing the polynomial $\varphi(B)$ around the zero frequency unit root plus the $S-1$ unit roots given in (2). Thus, write $\varphi(B)$ as:

$$\varphi(B) = \sum_{k=1}^{S} \lambda_k \Delta(B) \frac{1-\delta_k(B)}{\delta_k(B)} + \Delta(B)\varphi^*(B) \tag{3}$$

where

$$\delta_k(B) = 1 - \frac{1}{\theta_k}B, \quad \lambda_k = \frac{\varphi(\theta_k)}{\prod_{j \neq k} \delta_j(\theta_k)}, \quad \Delta(B) = \prod_{k=1}^{S} \delta_k(B),$$

$\varphi^*(B)$ is a remainder with roots outside the unit circle, and the θ_k are the zero frequency unit root plus the $S-1$ seasonal unit roots. In the case of monthly data, substitution of (3) into (1) gives

$$\varphi(B)^* y_{13t} = \sum_{k=1}^{12} \pi_k y_{k,t-1} + \epsilon_t \tag{4}$$

where:

$$
\begin{aligned}
y_{1t} &= (1 + B + B^2 + B^3 + B^4 + B^5 + B^6 + B^7 + B^8 + B^9 + B^{10} + B^{11})x_t \\
y_{2t} &= -(1 - B + B^2 - B^3 + B^4 - B^5 + B^6 - B^7 + B^8 - B^9 + B^{10} - B^{11})x_t \\
y_{3t} &= -(B - B^3 + B^5 - B^7 + B^9 - B^{11})x_t \\
y_{4t} &= -(1 - B^2 + B^4 - B^6 + B^8 - B^{10})x_t \\
y_{5t} &= -\frac{1}{2}(1 + B - 2B^2 + B^3 + B^4 - 2B^5 + B^6 + B^7 - 2B^8 + B^9 + B^{10} - 2B^{11})x_t \\
y_{6t} &= \frac{\sqrt{3}}{2}(1 - B + B^3 - B^4 + B^6 - B^7 + B^9 - B^{10})x_t \\
y_{7t} &= \frac{1}{2}(1 - B - 2B^2 - B^3 + B^4 + 2B^5 + B^6 - B^7 - 2B^8 - B^9 + B^{10} + 2B^{11})x_t \qquad (5) \\
y_{8t} &= -\frac{\sqrt{3}}{2}(1 + B - B^3 - B^4 + B^6 + B^7 - B^9 - B^{10})x_t \\
y_{9t} &= -\frac{1}{2}(\sqrt{3} - B + B^3 - \sqrt{3}B^4 + 2B^5 - \sqrt{3}B^6 + B^7 - B^9 + \sqrt{3}B^{10} - 2B^{11})x_t \\
y_{10t} &= \frac{1}{2}(1 - \sqrt{3}B + 2B^2 - \sqrt{3}B^3 + B^4 - B^6 + \sqrt{3}B^7 - 2B^8 + \sqrt{3}B^9 - B^{10})x_t \\
y_{11t} &= \frac{1}{2}(\sqrt{3} + B - B^3 - \sqrt{3}B^4 - 2B^5 - \sqrt{3}B^6 - B^7 + B^9 + \sqrt{3}B^{10} + 2B^{11})x_t \\
y_{12t} &= -\frac{1}{2}(1 + \sqrt{3}B + 2B^2 + \sqrt{3}B^3 + B^4 - B^6 - \sqrt{3}B^7 - 2B^8 - \sqrt{3}B^9 - B^{10})x_t \\
y_{13t} &= (1 - B^{12})x_t.
\end{aligned}
$$

Appendix A shows that each y_{kt} can be written as a function of the frequency associated with that y_{kt}. In addition, for $k \geq 4$, y_{k-1t} can be written simply in terms of y_{kt} and y_{kt-1}, which makes construction of these variables easier.[1]

In order to test hypotheses about various unit roots, one estimates (4) by Ordinary Least Squares and then compares the OLS test statistics to the appropriate finite sample distributions based on Monte Carlo results. For frequencies 0 and π, one simply examines the relevant t-statistic for $\pi_k = 0$ against the alternative that $\pi_k < 0$. For the other roots, one tests $\pi_k = 0$, where k is even, with a two-sided test. The even coefficient is zero if the series contains a unit root at that

[1] Our implementation of the HEGY approach differs from that in Franses (1990) in that ours makes the set of regressors mutually orthogonal. This greatly facilitates the derivation of the asymptotic distribution.

frequency. It is not zero otherwise for the seasonal frequencies other than $\frac{\pi}{2}$. For $\frac{\pi}{2}$, the coefficient is not zero if no root exists at that frequency. Under the alternative, the even coefficient may be positive or negative. If one fails to reject $\pi_k = 0$, then one tests $\pi_{k-1} = 0$ versus the alternative that $\pi_{k-1} < 0$. The test is one-sided because the sensible alternative is that the series contains a root outside the unit circle. Under stationarity the true coefficient is less than zero. Another strategy is to test $\pi_{k-1} = \pi_k = 0$ by calculating an F-statistic. To show that no unit root exists at any seasonal frequency, π_k must not equal zero for $k = 2$ and for at least one member of each of the sets $\{3,4\}$, $\{5,6\}$, $\{7,8\}$, $\{9,10\}$, $\{11,12\}$. One cannot separately test for a unit root at a given seasonal frequency and its negative because of the aliasing problem.

The hypothesis tests are amendable to the case where the alternative includes a constant, seasonal dummies, or a time trend. Equation (4) becomes

$$\varphi(B)^{*} y_{13t} = \sum_{k=1}^{12} \pi_k y_{k,t-1} + m_0 t + m_1 + \sum_{k=2}^{12} m_k S_{kt} + \epsilon_t \ . \qquad (6)$$

The equation is still estimated by OLS, but the asymptotic and finite sample distributions change.

The advantage of the HEGY procedure over earlier approaches is that it allows one to distinguish processes that may be integrated at only some of the seasonal frequencies. Hasza and Fuller (1982) consider as their null the model

$$(1 - B)(1 - B^d) y_t = \epsilon_t,$$

where $d = 2, 4, 12$ depending on the number of observations per year. They suggest estimating the equations

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-d} + \alpha_3 y_{t-d-1} + \epsilon_t \ ,$$

$$y_t = \phi_1 y_{t-1} + \phi_2 (y_{t-d} - \phi_1 y_{t-d-1}) + \epsilon_t \ ,$$

and then testing the restriction $[\alpha_1, \alpha_2, \alpha_3] = [1, 1, -1]$ or $[\phi_1, \phi_2] = [1, 1]$ by calculating a standard F-statistic and using the proper distribution in their Table 5.1.[2]

[2] Wasserfallen (1986) applies the Hasza-Fuller test to aggregate macro series for several OECD countries. He finds

Interpretation of results from the Hasza–Fuller test is difficult for two reasons. First, the test imposes two unit roots at frequency zero under the null. Second, it is unclear how the performance of the test changes when only some of the seasonal frequencies possess a unit root. Rejection does not prove that *no* unit root exists at any frequency. Moreover, failure to reject does not help identify the frequencies that are integrated. One is left with the null that *all* of the frequencies possess unit roots, and frequency zero has two. Finally, as HEGY point out, the test imposes a particular form for the alternative: all of the seasonal roots have the same modulus. If the true process has no unit roots but the roots at the seasonal frequencies have different moduli, there will be residual autocorrelation and many lags of the dependent variable will be needed to whiten the errors. Thus, applications of the test are likely to suffer from low power and residual autocorrelation in the errors.

Dickey, Hasza and Fuller (1984) bypass the assumption of two unit roots at frequency zero by considering the null model $(1 - B^d)y_t = \epsilon_t$. They suggest estimating the equation

$$x_t = \rho x_{t-d} + b_t + \epsilon_t, \tag{7}$$

and testing whether $\rho = 1$, where b_t is either a constant or seasonal dummies. They publish the distribution of various statistics when the regression equation includes a constant and seasonal dummies.[3] As with the Hasza–Fuller test, this test suffers from the problems of interpreting rejections, low power and residual autocorrelation. The HEGY test procedure avoids these problems by including a regressor corresponding to each of the potential unit roots.

---

limited evidence of unit roots at seasonal frequencies. Ghysels (1990) applies the test to quarterly U.S. GNP and per capita GNP. When he excludes seasonal dummies and a time trend, he fails to reject the $(1-B)(1-B^4)$ model. When he includes seasonal dummies and a time trend he rejects the model, but autocorrelation in the residuals prevents him from putting much faith in the results.

[3] Ghysels (1990) applies the Dickey-Hasza-Fuller test to GNP data and rejects the $(1-B^4)$ specification in favor of a stationary alternative that does not include seasonal dummies and trends. Bhargava (1987b) also considers the test of $\rho = 1$ in equation (7), emphasizing the importance of the inclusion of seasonal dummies. He rejects the null of $\rho = 1$ in U.K. consumer expenditures. See also Bhargava (1987a).

2.2 Asymptotic Distributions of the Monthly HEGY Test Statistics

Derivation of the asymptotic distributions of the test statistics proposed above follows Chan and Wei (1988), Park and Phillips (1988), Phillips (1987), and Stock (1988). We consider first the distributions of the t-statistics when constant, seasonal dummy and trend terms are excluded and the regression is correctly specified. The goal is to develop the asymptotic distribution of t_1, t_2, t_k and t_{k+1} where $k \in \{3,5,7,9,11\}$. We prove that the asymptotic distributions of the five t_k are the same as are those of the five t_{k+1}. For ease of notation, we write that k is odd if $k \neq 1$ and $k \in \{3,5,7,9,11\}$ and that k is even if $k \neq 2$ and $k \in \{4,6,8,10,12\}$. W_k denotes a standard brownian motion independent of all other W_j for $k \neq j$. Lemma 1 gives the asymptotic distributions of these t-statistics.

Lemma (1):

$$t_k = \frac{\sum_{t=2}^{T} y_{kt-1}\epsilon_t}{\hat{\sigma}\left(\sum_{t=2}^{T} y_{kt-1}^2\right)^{\frac{1}{2}}} \xrightarrow{\mathcal{L}} \begin{cases} \dfrac{\int_0^1 W_1(r)dW_1(r)}{\left(\int_0^1 W_1(r)^2dr\right)^{\frac{1}{2}}} & \text{if } k=1; \\ \dfrac{\int_0^1 W_2(r)dW_2(r)}{\left(\int_0^1 W_2(r)^2dr\right)^{\frac{1}{2}}} & \text{if } k=2; \\ \dfrac{\int_0^1 W_k(r)dW_k(r)+\int_0^1 W_{k+1}(r)dW_{k+1}(r)}{\left(\int_0^1 W_k(r)^2dr+\int_0^1 W_{k+1}(r)^2dr\right)^{\frac{1}{2}}} & \text{if } k \text{ odd}; \\ \dfrac{\int_0^1 W_k(r)dW_{k-1}(r)-\int_0^1 W_{k-1}(r)dW_k(r)}{\left(\int_0^1 W_{k-1}(r)^2dr+\int_0^1 W_k(r)^2dr\right)^{\frac{1}{2}}} & \text{if } k \text{ even}. \end{cases}$$

Proof: The appendix provides the full proof; we sketch it here. Note that y_{kt-1} can be written in terms of sine or cosine weighted averages of the last twelve partial sums of ϵ_t. The numerator is the scalar sum of $y_{kt-1}\epsilon_t$ because $y_{k,t-1}$ is orthogonal to $y_{j,t-1}$ $\forall j \neq k$ by construction. The denominator is the sum of y_{kt-1}^2. One first sums over years keeping the cosine and sine terms constant. This leads to the type of analysis in Phillips (1987) or Stock (1988). One then sums over the trigonometry terms as in Ahtola and Tiao (1987). •

It follows from Lemma 1 that t_1 and t_2 have the same asymptotic distribution, that all of the odd seasonal statistics have the same distribution, and that all of the even seasonal statistics have the same distribution. The Itô calculus shows that t_1 and t_2 have the Dickey-Fuller (1979) distribution (Phillips, 1987; Stock, 1988). The construction of π_1 and π_2 in Franses (1990) is the same as here, so his statistics have the same distributions as ours. The distributions of the odd seasonal t-statistics are the negatives of the distributions in Corollary 3.3.8 in Chan and Wei (1988) and the same as for t_3 in Engle, Hylleberg, Granger and Lee (1991) (EHGL), while the distributions of the corresponding even statistics are the same as for t_4 in EGHL. The construction of π_3 and π_4 in Franses (1990) is also the same as here with the exception that he reverses our notation. As Chan and Wei prove, the distribution of t_3 is the same as for $\hat{t}_d$ with $d = 2$ in Dickey-Hasza-Fuller (1984) (see also EGHL). Finally, Ahtola and Tiao (1987) and Chan and Wei note that the distribution of the odd seasonal statistics are the same across frequencies.

We next consider the distributions of the t-statistics when deterministic terms are included in the regression. Define y_{kt}^{μ} as the residual from a regression of y_{kt} on a constant. Define y_{kt}^{τ}, y_{kt}^{ξ} and $y_{kt}^{\xi\tau}$ analogously, where τ stands for constant plus trend, ξ stands for constant plus eleven dummies, and $\xi\tau$ stands for constant, eleven dummies, and trend. The numerators and denominators of the respective regressions are partially defined in Lemma 1. Let N_k^x equal the part of the numerator different from the corresponding numerator in Lemma 1, $x \in \{\mu, \tau, \xi, \xi\tau\}$. Let D_k^x equal the square of the denominator different from the square of the corresponding denominator in Lemma 1.

Lemma (2):

$$(\mu) \quad N_k^{\mu} \equiv \begin{cases} -W_1(1)\int_0^1 W_1(r)dr & \text{if } k = 1; \\ 0 & \text{if } k \geq 2 \end{cases}$$

$$(\tau) \quad N_k^{\tau} \equiv \begin{cases} -4W_1(1)\int_0^1 W_1(r)dr + 6\int_0^1 W_1(r)dr\int_0^1 rdW_1(r) & \\ \quad -12\int_0^1 rW_1(r)dr\int_0^1 rdW_1(r) + 6W_1(1)\int_0^1 rW_1(r)dr & \text{if } k = 1; \\ 0 & \text{if } k \geq 2 \end{cases}$$

$$(\xi) \quad N_k^{\xi} \equiv \begin{cases} -W_1(1)\int_0^1 W_1(r)dr & \text{if } k = 1; \\ -W_2(1)\int_0^1 W_2(r)dr & \text{if } k = 2; \\ -\frac{1}{2}W_k(1)\int_0^1 W_k(r)dr - \frac{1}{2}W_{k+1}(1)\int_0^1 W_{k+1}(r)dr & \text{if } k \text{ odd}; \\ -\frac{1}{2}W_k(1)\int_0^1 W_{k-1}(r)dr + \frac{1}{2}W_{k-1}(1)\int_0^1 W_k(r)dr & \text{if } k \text{ even}; \end{cases}$$

$$
(\xi\tau)\; N_k^{\xi\tau} \equiv \begin{cases} -4W_1(1)\int_0^1 W_1(r)dr + 6\int_0^1 W_1(r)dr\int_0^1 rdW_1(r) & \\ \quad -12\int_0^1 rW_1(r)dr\int_0^1 rdW_1(r) + 6W_1(1)\int_0^1 rW_1(r)dr & \text{if } k=1; \\ -W_2(1)\int_0^1 W_2(r)dr & \text{if } k=2; \\ -\frac{1}{2}W_k(1)\int_0^1 W_k(r)dr - \frac{1}{2}W_{k+1}(1)\int_0^1 W_{k+1}(r)dr & \text{if } k \text{ odd}; \\ -\frac{1}{2}W_k(1)\int_0^1 W_k(r)dr + \frac{1}{2}W_{k-1}(1)\int_0^1 W_{k-1}(r)dr & \text{if } k \text{ even}. \end{cases}
$$

$$
(\mu)\quad D_k^{\mu} \equiv \begin{cases} -\left(\int_0^1 W_1(r)dr\right)^2 & \text{if } k=1; \\ 0 & \text{if } k \geq 2 \end{cases}
$$

$$
(\tau)\quad D_k^{\tau} \equiv \begin{cases} -4\left(\int_0^1 W_1(r)dr\right)^2 + 12\int_0^1 W_1(r)dr\int_0^1 rW_1(r)dr & \\ \qquad -12\left(\int_0^1 rW_1(r)dr\right)^2 & \text{if } k=1; \\ 0 & \text{if } k \geq 2 \end{cases}
$$

$$
(\xi)\quad D_k^{\xi} \equiv \begin{cases} -\left(\int_0^1 W_1(r)dr\right)^2 & \text{if } k=1; \\ -\left(\int_0^1 W_2(r)dr\right)^2 & \text{if } k=2; \\ -\frac{1}{4}\left(\int_0^1 W_k(r)dr\right)^2 - \frac{1}{4}\left(\int_0^1 W_{k+1}(r)dr\right)^2 & \text{if } k \text{ odd}; \\ -\frac{1}{4}\left(\int_0^1 W_{k-1}(r)\right)^2 dr - \frac{1}{4}\left(\int_0^1 W_k(r)dr\right)^2 & \text{if } k \text{ even}; \end{cases}
$$

$$
(\xi\tau)\; D_k^{\xi\tau} \equiv \begin{cases} -4\left(\int_0^1 W_1(r)dr\right)^2 + 12\int_0^1 W_1(r)dr\int_0^1 rW_1(r)dr & \\ \qquad -12\left(\int_0^1 rW_1(r)dr\right)^2 & \text{if } k=1; \\ -\left(\int_0^1 W_2(r)dr\right)^2 & \text{if } k=2; \\ -\frac{1}{4}\left(\int_0^1 W_k(r)dr\right)^2 - \frac{1}{4}\left(\int_0^1 W_{k+1}(r)dr\right)^2 & \text{if } k \text{ odd}; \\ -\frac{1}{4}\left(\int_0^1 W_{k-1}(r)\right)^2 dr - \frac{1}{4}\left(\int_0^1 W_k(r)dr\right)^2 & \text{if } k \text{ even}. \end{cases}
$$

Proof: The proof is similar to that for Lemma 1 and is developed fully in the Appendix. See also Park and Phillips (1988). •

Given these results, the distributions of the t-statistics when the regression includes deterministic terms are as follows:

Lemma (3):

$$
t_k = \frac{\sum_{t=2}^T y_{kt-1}^x \epsilon_t^x}{\hat{\sigma}\left(\sum_{t=2}^T y_{kt-1}^2\right)^{\frac{1}{2}}} \xrightarrow{\mathcal{L}} \begin{cases} \frac{\int_0^1 W_1(r)dW_1(r) + N_1^x}{\left(\int_0^1 W_1(r)^2 dr + D_1^x\right)^{\frac{1}{2}}} & \text{if } k=1; \\ \frac{\int_0^1 W_2(r)dW_2(r) + N_2^x}{\left(\int_0^1 W_2(r)^2 dr + D_2^x\right)^{\frac{1}{2}}} & \text{if } v=2; \\ \frac{\int_0^1 W_k(r)dW_k(r) + \int_0^1 W_{k+1}(r)dW_{k+1}(r) + N_k^x}{\left(\int_0^1 W_k(r)^2 dr + \int_0^1 W_{k+1}(r)^2 dr + D_k^x\right)^{\frac{1}{2}}} & \text{if } k \text{ odd}; \\ \frac{\int_0^1 W_k(r)dW_{k+1}(r) - \int_0^1 W_{k+1}(r)dW_k(r) + N_k^x}{\left(\int_0^1 W_{k-1}(r)^2 dr + \int_0^1 W_k(r)^2 dr + D_k^x\right)^{\frac{1}{2}}} & \text{if } k \text{ even}, \end{cases}
$$

where $x = \mu, \tau, \xi, \xi\tau$.

Proof: Directly from Lemmas 1 and 2 and the continuous mapping theorem (Billingsley, 1988). •

Lemma 3 shows that all of the odd statistics have the same distributions when different deterministic regressors are included in the regression. The same is true for the even statistics. One can also see that t_1 is invariant to the inclusion of seasonal dummies as long as a constant is included. The distributions of t_1^{μ} and t_1^{τ} are the same as in Park and Phillips (1988) or Stock (1988), while the distributions of $t_2, \ldots, t_{12}$ are independent of constant and trend terms. The explanation is that the terms y_{kt} $(k \geq 2)$ can be written as functions of cosine and sine waves that repeat every twelve periods and sum to zero over those periods. These terms are asymptotically orthogonal to terms that are not periodic, such as a constant or a time trend. Also, the distribution of t_2 when dummies are included in the regression is the same as that of t_1 when only a constant is included.

The distributions of the F-statistics follow from Lemmas 1-3. Write

$$F_k = (1/(2\hat{\sigma}^2))b'(X'X)b$$

with

$$b = \begin{bmatrix} \beta_k \\ \beta_{k+1} \end{bmatrix} \qquad X = \begin{bmatrix} y_{k,1} & y_{k+1,1} \\ \vdots & \vdots \\ y_{k,T} & y_{k+1,T} \end{bmatrix}.$$

This implies

Lemma (4):

$$F_k^x \xrightarrow{\mathcal{L}} \tfrac{1}{2}(t_k^{x\,2} + t_{k+1}^{x\,2}),$$

where $x = \cdot, \mu, \tau, \xi$, and $\xi\tau$ and $k \in \{3, 5, 7, 9, 11\}$.

Proof: The orthogonality of $y_{k,t-1}$ and $y_{k+1,t-1}$ means that $X'X$ is diagonal. Writing out the above in scalar notation gives:

$$F_k^x = (1/(2\hat{\sigma}^2))\left(\beta_k^{x\,2} \sum_{t=2}^{T} y_{k,t-1}^{x\,2} + \beta_{k+1}^{x\,2} \sum_{t=2}^{T} y_{k+1,t-1}^{x\,2}\right) = \tfrac{1}{2}(t_k^{x\,2} + t_{k+1}^{x\,2}).$$

The asymptotics then follow from Lemmas 1-3 and the continuous mapping theorem. See also EGHL. •

2.3 Finite Sample and Asymptotic Critical Values for the Test Statistics

Table A1 contains critical values from the finite sample distributions of the t and F statistics needed to employ the HEGY procedure with monthly data. The critical values were obtained by simulating 24,000 regressions of the form (4), with various combinations of constants, seasonal dummies and trends included. The fundamental series were generated by $y_t = y_{t-12} + \epsilon_t$, with ϵ_t standard normal. These results complement those in HEGY for quarterly data.

The critical values for t_{odd}, t_{even} and F were calculated by combining the observations on all five similar statistics. For instance, t_{odd} is computed by stacking t_3, t_5, t_7, t_9 and t_{11} and calculating the order statistic for that $(120,000 \times 1)$ vector. Thus, t_1 and t_2 are based on 24,000 observations while t_{odd}, t_{even} and F are based on 120,000 observations. The first rationale for this procedure is that the asymptotic distributions of these statistics are the same for each of the five pairs of coefficients, as shown above. In addition, investigations of the finite sample distributions for a subset of the cases considered below indicates that these distributions are similar for a given number of simulations and converge as the number of simulations increases.

Table A1 also contains the asymptotic critical values of the monthly HEGY test statistics. To calculate these values, we approximate the functions of brownian motion using partial sums of normal random variables. For example, $W(1)$ is approximated as

$$W(1) = \frac{1}{\sqrt{5000}} \sum_{j=1}^{5000} \epsilon_j,$$

where the ϵ_j are independent, standard normal random variables. This procedure is then repeated 100,000 times.

The standard errors of the estimated critical values are less than or equal to .02 for t_1 and t_2, less than or equal to .01 for t_{odd} and t_{even}, and less than or equal to .03 for F. These standard errors are based on the asymptotic distribution of the normalized central order statistic (Bickel and Doksum, 1977, p. 400). In calculating the standard errors, a kernel estimate of the density using the Epanechnikov kernel with bandwidth equal to $(40\sqrt{\pi})^{.2}N^{-.2}$ was used (N = 24,000 or 120,000).

Examination of the results in Table A1 suggests that the finite sample distributions display all the characteristics of the asymptotic distributions described above. First, the distribution of each of the eleven seasonal coefficients is unaffected by the presence of a constant or a constant plus trend; these terms only affect the distribution of t_1. Second, the distribution of t_1 is unaffected by the presence of seasonal dummies as long as a constant is included in the regression. Third, the distribution of t_2 when dummies are included is similar to that of t_1 when a constant is included. In addition, the results in the table indicate that there are not large differences between the finite sample and asymptotic distributions.

2.4 Size and Power Issues

One problem that arises in carrying out these tests is the treatment of residual autocorrelation in x_t. If $\varphi(B)$ is allowed to be of order greater than S, then $\varphi^*(B) \neq 1$, so additional lags of y_{13t} must be included on the right hand side of (4). These extra lags do not affect the asymptotic distribution of the test statistics (so long as they correctly estimate the remaining AR component of x_t), but they do affect the finite sample distribution. In particular, if the model has a true MA component, the correct number of lags in $\varphi(B)$ is infinite. This implies a finite sample bias that disappears as the number of lags grows to infinity. Stated differently, the testing procedure approximates a true ARMA model with a high-order AR model (Said and Dickey, 1984). Schwert (1987), however, maintains that there exist substantial biases in zero frequency unit root tests for samples of moderate length.

In Table A2 we present some suggestive evidence on this source of finite sample bias. Ghysels.

Lee and Noh (1992) provide a similar discussion. The top half of the table reports the probability of rejecting the null of integration when the true process is integrated but contains an MA(1) component that is approximated as a finite order AR component. The two specifications presented in the table model the error term as $\epsilon_t = \eta_t + \rho\eta_{t-1}$, where $\rho = \pm.85$ and the regressions include twelve lags of the dependent variable. To save computation time, the table reports results based on a series twenty years long, with regressions that contain a constant and seasonal dummies but no trend. Inclusion of a trend does not significantly affect the results.

Two consistent results emerge in the table. First, there is some bias in the size of the test. At a test size of 5 percent, t_3, t_5 and t_7 reject only about 1.5 percent of the time, and t_9 and t_{11} only about 3.3 percent. The even statistics are more troublesome. When $\rho = -.85$, the distributions of t_4 and t_8 shift to the left. 17 and 8 percent of the densities, respectively, are to the left of the 5 percent critical value. The distributions of t_6 and t_{10} shift to the right. 20 and 13 percent of the densities are to the right of the 5 percent right-hand side critical value. The F-statistics uniformly reject at a rate lower than the implied critical value. The highest is $F_{9,10}$ at 3.9 percent for a 5 percent test; the lowest is $F_{7,8}$ at 1.8 percent. The distributions of t_1 and t_2 are not substantially affected. At a test size of 5 percent, one rejects a unit root for t_1 4.9 percent of the time and for t_2 3.4 percent of the time.

Second, the difference between the effect of $\rho = -.85$ versus $\rho = .85$ can be predicted by examining the spectrum of an MA(1) process (this argument ignores the twelve lags of the dependent variable that also appear in the regression). If $x_t = \eta_t + \rho\eta_{t-1}$ the spectrum of x_t given ρ is

$$S_{x,\rho}(\omega) = \sigma_\eta^2(1 + \rho^2 + 2\rho\cos\omega).$$

It is immediately clear that for an MA(1), $S_{x,\rho}(\omega) = S_{x,-\rho}(\pi - \omega)$. This implies that the distribution of statistics associated with frequency ω when $\rho = .85$ will be the same as those associated with $\pi - \omega$ when $\rho = -.85$. Table A2 verifies this prediction. The rejection probabilities for t_1 ($\omega = 0$)

and t_2 ($\omega = \pi$) flip. The pair t_5, t_6 ($\omega = \frac{2\pi}{3}$) flips with the pair t_7, t_8 ($\omega = \frac{\pi}{3}$), and the pair t_9, t_{10} ($\omega = \frac{5\pi}{6}$) flips with the pair t_{11}, t_{12} ($\omega = \frac{\pi}{6}$). The distribution of t_3 ($\omega = \frac{\pi}{2}$) appears not to change while t_4 changes its center so that it is offset to the other side of the original distribution.

The results of a different model are reported in the bottom panel of Table A2. Here the fundamental series is generated by $(1 - B^{12})x_t = (1 - .85B^{12})\epsilon_t$, where ϵ_t is white noise. The series is still integrated at all the respective frequencies but only fifteen percent of the shock is permanent. The regressions that generate the t and F statistics include twelve lags of the dependent variable, a constant and eleven seasonal dummies. As in the previous cases the statistics for a unit root at frequency 0 and π are not much affected by the additional correlation and lags of the dependent variable. At a test size of 5 percent one rejects 4.4 and 4.5 percent of the time, respectively. For t_3–t_{12} the distributions shift to the right. For the left-hand side statistics one rejects much less often than is implied by the size of the test. For the even statistics one rejects more for the right-hand side. As in the MA(1) cases, the F-statistics have lower sizes.

One should avoid drawing a strong conclusion about the robustness of these statistics to residual correlation in the errors from the few simulations reported above. Still, the statistics we focus on the most, t_1, t_2 and the five F-statistics, do not have true sizes higher than implied by the critical values in Table A1. The results also suggest that the $F-$statistics are better behaved than the sequential t-tests on $t_3, \ldots, t_{12}$.

A second area of investigation is the power of these statistics to some alternative. As is typically the case, there are many alternatives one would like to consider. In the top half of Table A3 we report the probabilities of rejecting the null of unit roots when the true x_t process is generated by

$$(1 - \rho^{12}B^{12})x_t = \epsilon_t. \tag{8}$$

In this case the series contains roots at frequency zero and all of the seasonal frequencies, but all twelve roots lie outside of the unit circle. In the bottom half of the table we report the probabilities of rejecting the null of unit roots when the true x_t process is generated by

$$(1 - B^4)(1 + \rho^4 B^4 + \rho^8 B^8)x_t = \epsilon_t. \tag{9}$$

In this case the series contains roots at the same frequencies as in (8), but for frequencies 0, π and $\frac{\pi}{2}$ the roots lie on the unit circle. The stationary roots have modulus $\frac{1}{\rho}$, as in (8).

In all cases the regressions contain a constant and eleven dummies but no trend. Experimentation suggests that the inclusion of a trend does not change the results, so we omit this term to reduce the computational burden. The fact that we include seasonal dummies in the regressions means that the results are robust to processes that include an arbitrary seasonal dummy pattern.

The tables suggest several conclusions about the power of these tests. First, when the frequency in question possesses a unit root, the probability of rejection is close to the size of the test, independent of frequency and independent of the size of the roots at other frequencies. For instance, when the process is generated as in (9), one rejects unit roots at frequencies 0, π, and $\frac{\pi}{2}$ only 5-6% of the time at a 5 percent critical value. Second, the power of the test is moderate when the root lies close to the unit circle. For instance, when the series is generated as in (8) with the modulus equal to $\frac{1}{.95} = 1.05$, the test rejects unit roots at frequencies 0 and π 45 percent of the time and at the remaining frequencies 71 percent of the time. As one increases the modulus of the root, the probability of rejection increases substantially. When the modulus is $\frac{1}{.85} = 1.18$, the test rejects at frequencies 0 and π 99 percent of the time and at the remaining frequencies 100 percent of the time. Finally, the test has more power at frequencies with complex roots than at frequencies 0 and π.

3. Evidence on Seasonal Unit Roots

We now present results of applying the HEGY test procedure to aggregate U.S. time series. The quarterly series that we consider are real GNP, consumption, fixed investment, and government purchases (for these regressions we use the critical values presented in HEGY). The monthly series are the nominal money stock, the price level, real and nominal interest rates, industrial production, real retail sales, real wages, and the unemployment rate. All series are in log levels, except for nominal rates (log of gross rates) and real rates (ex-post net rates). The sample period is always the longest subperiod of the post-WWII period for which data are available; details are provided in the tables. The sources and exact definitions of the series are given in the Appendix B.

The estimation equations include a constant, eleven seasonal dummies, time, and lags of the dependent variable. We allow for seasonal dummies in all tests because the loss of power that results from their inclusion when unnecessary is insignificant compared to the bias that results from their omission when necessary. We report OLS t- and F-statistics.

For the results presented in the tables, the set of lags was determined by first estimating the equation with three years of lags and then excluding those lags that failed to enter significantly at the 15 percent level. This approach trades off the loss of power that results from including unnecessary lags against the bias that results from excluding necessary lags. We have also used the Schwarz (1987) and Akaike (1974, 1976) criteria to determine the number of lags. We compared the criteria for all unrestricted autoregressions of order P, with P equal 36 months or fewer. The Schwarz criterion almost always chooses a specification with fewer than a year of lags, while the Akaike criterion (which is known to overparameterize (Granger and Newbold, 1986)), usually chooses the specification with almost three years of lags. Based on an earlier version of this paper, which simply reported results for 0, 1, 2 or 3 years of lags, we conclude that if we chose the lag length based on the Schwarz criterion we would obtain results similar to those reported in Tables 1 and 2 below. If we allowed for the much larger set of lags implied by the Akaike criterion, we would reject much less often.

In the quarterly data on the national income accounts, reported in Table 1, there is no series for which the data fail to reject a unit root for at least one of the seasonal frequencies at the 5 percent level or better. At $\frac{\pi}{2}$, the F-statistics reject at the 5 percent level for GNP, Investment and Government Purchases and at the 10 percent level for Consumption. In addition, the t-statistic for either π_3 or π_4 is below the 5 percent critical value for all four series. At frequency π we reject at the 5 percent level for GNP and Investment and at the 10 percent level for Government Purchases, but we fail to reject for Consumption. At frequency zero we reject at the 5 percent level for Investment and Government Purchases but fail to reject for GNP and Consumption.

In the monthly data on measures of real activity, reported in Table 2, we generally reject seasonal unit roots more strongly than in the quarterly data in Table 1. For Real Retail Sales we reject unit roots at the 5 percent level at all the seasonal frequencies (including frequency π) using the t-statistics, and we reject at frequencies $\frac{\pi}{6}$, $\frac{\pi}{3}$, $\frac{\pi}{2}$, $\frac{2\pi}{3}$, and $\frac{5\pi}{6}$ using the F-statistics. For Industrial Production we reject at every seasonal frequency using the t-statistics and in every case except $\frac{\pi}{6}$ using the F-statistics. The data on Unemployment and Money reject unit roots less often than those for Retail Sales and Industrial Production, although they reject at several frequencies in both cases. For both series, the data fail to reject at the 5 percent level at frequency π. For Unemployment, they also fail to reject at frequency $\frac{5\pi}{6}$ using the t- and F-statistics and at frequencies $\frac{\pi}{6}$, $\frac{\pi}{2}$, and $\frac{5\pi}{6}$ using the F-statistics. At frequencies $\frac{\pi}{6}$, $\frac{\pi}{3}$, $\frac{\pi}{2}$, and $\frac{2\pi}{3}$, the data reject at the 5 percent level using the t-statistics. They also reject at $\frac{2\pi}{3}$ at 5 percent and at $\frac{\pi}{3}$ at 10 percent using the F-statistic. For Money, both the t and F-statistics reject at all frequencies other than π.

The data on monthly price variables, also reported in Table 2, provide stronger evidence against seasonal unit roots than the data for quantities. The data reject at the 5 percent level or better using the $t-$ and F-statistics at all seasonal frequencies, except for the F-statistic for the Real Wage at frequency $\frac{\pi}{6}$, in which case the data still reject at 10 percent. The data reject the zero frequency unit root at the 5 percent level for the Real Rate and at the 10 percent level for the Nominal Rate.

To summarize, for most series we reject unit roots at most frequencies, and there is no series for which we fail to reject unit roots for at least one of the seasonal frequencies. The strongest evidence for a seasonal unit root is at frequency π, but even in this case we reject more often than not at the 5 percent level. Moreover, as discussed earlier, tests at this frequency have lower power than tests for seasonal unit roots at other frequencies. Generally, we fail to reject the hypothesis of a unit root at frequency zero. Osborn (1990) and Lee and Siklos (1990) provide similar results for the U.K. and Canada, respectively, while Beaulieu and Miron (1990a,b) obtain similar conclusions for disaggregated U.S. manufacturing data and aggregate OECD data.

4. Conclusion

We conclude by discussing the significance of these results. As noted by many authors in the context of the zero frequency, there are potentially serious problems of statistical inference when one or more series contains a unit root. For example, the variance of a series with a unit root is infinite, the distribution of the estimated first order autocorrelation changes when the true autocorrelation equals one, and two independent integrated series can display spurious correlation (Granger and Newbold, 1986). These problems are not alleviated by blind differencing to ensure stationarity (Quah and Wooldridge, 1988).

Many of these same issues arise when one considers unit roots at frequencies other than zero. In prior considerations of seasonality, different parameterizations have been used to explore the importance and effects of seasonality in univariate and multivariate contexts. The use of seasonal dummy variables, as in Barsky and Miron (1989), is not appropriate if the observed seasonality is generated by an integrated process. Models of seasonal cointegration, such as in HEGY, EGHL, and Lee and Siklos (1989), require the series to be seasonally integrated. Finally, the appropriateness of applying the filter $(1 - B^d)$ to a series with a seasonal component, as advocated by Box and Jenkins (1970), depends on the series being integrated at frequency zero *and all* of the seasonal frequencies.

Given these observations, the implication of our results is that the mechanical application of the seasonal difference filter is likely to produce serious misspecification in many instances. There is sufficient evidence against seasonal unit roots that empirical researchers should check for their presence using procedures such as the one discussed above, rather than simply imposing seasonal unit roots *a priori*. In addition, since the power of these tests is never perfect, it is presumably also sensible to ask whether the economic problem under consideration suggests mechanisms that would plausibly give rise to seasonal unit roots.

References

Ahtola, J.A. and G.C. Tiao, 1987, Distribution of least squares estimators of autoregressive parameters for a process with complex roots on the unit circle, Journal of Time Series Analysis 8, 1-14.

Akaike, H., 1974, A new look at statistical model identification, IEEE Transactions on Automatic Control AC-19, 716-23.

Akaike, H., 1976, Canonical correlation analysis of time series and the use of an information criterion," in System Identification : Advances and Case Studies, R.K. Mehra and D.G. Lainiotis, eds., (Academic Press, New York), 52-107.

Barsky, R.B. and J.A. Miron, 1989, The seasonal cycle and the business cycle, Journal of Political Economy 97, 503-534.

Beaulieu, J.J. and J.A. Miron, 1990a, Seasonality in U.S. manufacturing, NBER Working Paper #3450.

Beaulieu, J.J. and J.A. Miron, 1990b, A cross country comparison of seasonal cycles and business cycles, NBER Working Paper #3459.

Bhargava, A., 1987a, An econometric analysis of the U.S. postwar G.N.P., mimeo, University of Pennsylvania.

Bhargava, A., 1987b, On the specification of regression models in seasonal differences, mimeo, University of Pennsylvania.

Bickel, P.J. and K.A. Doksum, 1977, Mathematical Statistics: Basic Ideas and Selected Topics (Holden Day, San Francisco, CA).

Box, G.E.P. and G.M. Jenkins, 1970, Time series analysis, forecasting and control (Holden Day, San Francisco, CA).

Burns, A.F. and W.C. Mitchell, 1946, Measuring business cycles (National Bureau of Economic Research, New York).

Chan, N.H. and C.Z. Wei, 1988, Limiting distributions of least squares estimates of unstable autoregressive processes, Annals of statistics16, 367-401.

Dickey, D.A., D.P. Hasza, and W.A. Fuller, 1984, Testing for unit roots in seasonal time series, Journal of the American Statistical Association 79, 355-367.

Engle, R.F., C.W.J. Granger and J.J. Hallman, 1989, Merging short– and long–run forecasts: an application of seasonal cointegration to monthly electricity sales forecasting, Journal of Econometrics 40, 45-62.

Engle, R.F., C.W.J. Granger, S. Hylleberg, and H.S. Lee, 1990, Seasonal cointegration: the Japanese consumption function, 1961.1-1987.4, mimeo, University of California, San Diego.

Franses, P.H., 1990, Testing for seasonal unit roots in monthly data, Report 9032/A of the Econometric Institute, Erasmus University, Rotterdam.

Ghysels, E., 1990, Unit root tests and the statistical pitfalls of seasonal adjustment: the case of U.S. post–war real GNP, Journal of Business and Economic Statistics 8, 145-52.

Ghysels, E., H. Lee and J. Noh, 1991, Testing for unit roots in seasonal time series: some theoretical extensions and a monte carlo investigation, mimeo, University of Montreal.

Granger, C.W.J. and P. Newbold, 1986, Forecasting economic time series, 2nd Ed., (Academic Press, New York).

Hasza, D.P. and W. Fuller, 1982, Testing for nonstationarity parameter specifications in seasonal time series models, Annals of Statistics 10, 1209-1216.

Hilmer, S.C. and G.C. Tiao, 1982, An ARIMA-model based approach to seasonal adjustment, Journal of the American Statistical Association 77, 63-70.

Hylleberg, S., R. Engle, C.W.J. Granger, and B.S. Yoo, 1990, Seasonal integration and co-integration, Journal of Econometrics 44, 215-238.

Lee, H.S. and P.L. Siklos, 1990, The influence of seasonal adjustment on cointegration: the case of the permanent income hypothesis for canada, 1947-87, mimeo, UCSD.

Macaulay, F.R., 1938, Some theoretical problems suggested by movements of interest rates, bond yields and stock prices in the United States since 1856, (National Bureau of Economic Research, New York).

Mitchell, W.C., 1927, The business cycle: the problem and its setting, (National Bureau of Economic Research, New York).

Osborn, D.R., 1990, A survey of seasonality in UK macroeconomic variables, International Journal of Forecasting, forthcoming.

Park, J. Y. and P.C.B. Phillips, 1988, Statistical inference in regressions with integrated processes: part 1, Econometric Theory, 4, 468–497.

Phillips, P.C.B., 1987, Time series regression with a unit root, Econometrica 55, 277-301.

Phillips, P.C.B., 1988a, Weak convergence to the matrix stochastic integral $\int_0^1 BdB'$, Journal of Multivariate Analysis 24, 252-264.

Phillips, P.C.B, 1988b, Weak convergence of sample covariance matrices to stochastic integrals via martingale aprproximations, Econometric Theory 4, 528-533.

Quah, D. and J.M. Wooldridge, 1988, A common error in the treatment of trending time series. M.I.T. Working Paper #483.

Said, S.E. and D.A. Dickey, 1984, Testing for unit roots in autoregressive-moving average models of unknown order, Biometrika 71, 599-608.

Schwarz, G., 1978, Estimating the dimension of a model, Annals of Mathematical Statistics VI, 461-464.

Schwert, G.W., 1987, Effects of model specification on tests for unit roots in macroeconomic data, Journal of Monetary Economics 20, 73-103.

Stock, J.H., 1988, A class of tests for integration and cointegration, mimeo, Kennedy School of Government, Harvard University.

Wasserfallen, W., 1986, Non-stationarities in macroeconomic time series - further evidence and implications, Canadian Journal of Economics 19, 498-510.

Table 1: Results of Tests for Seasonal Unit Roots in Quarterly Aggregate Series

			0	π	$\frac{\pi}{2}$		$\frac{\pi}{2}$
	Period	*Lags*	π_1	π_2	π_3	π_4	$F_{3,4}$
G.N.P.	49:2–87:4	3	-1.74	-2.93	-2.32	-4.44	11.63
Consumption	50:2–87:4	5	-1.67	-1.71	-2.75	-2.54	6.42
Investment	48:2–87:4	1	-3.73	-4.58	-2.65	-7.13	29.00
Gov't Purchases	50:4–87:4	6	-5.52	-2.64	-4.01	-2.29	9.00

Table 2: Results of Tests for Seasonal Unit Roots in Monthly Aggregate Series

			0	π	$\frac{\pi}{2}$		$\frac{2\pi}{3}$		$\frac{\pi}{3}$		$\frac{5\pi}{6}$		$\frac{\pi}{6}$		$\frac{\pi}{2}$	$\frac{2\pi}{3}$	$\frac{\pi}{3}$	$\frac{5\pi}{6}$	$\frac{\pi}{6}$
	Period	*Lags*	π_1	π_2	π_3	π_4	π_5	π_6	π_7	π_8	π_9	π_{10}	π_{11}	π_{12}	$F_{3,4}$	$F_{5,6}$	$F_{7,8}$	$F_{9,10}$	$F_{11,12}$
Retail Sales	61:11-88:2	9	-1.87	-4.57	-5.65	-2.08	-7.25	-0.81	-3.54	-3.36	-7.49	-0.48	-3.06	-3.15	18.15	27.26	12.31	28.14	9.60
Indus. Prod.	50:2 -88:12	9	-2.58	-3.21	-2.38	-4.00	-3.92	3.66	0.04	-3.68	-4.70	1.21	0.72	-2.30	10.12	12.36	6.97	11.67	2.85
Unemploy.	52:1 -88:12	13	-1.89	-1.65	-2.01	-2.27	-3.38	2.17	-0.53	-3.38	-2.57	1.60	0.55	-2.56	4.49	8.17	5.88	4.57	3.40
Money	61:6 -88:12	8	-1.49	-2.53	-1.25	-4.12	-4.65	1.80	-2.15	-5.27	-4.61	2.34	-0.39	-4.95	9.11	12.54	16.28	13.50	12.43
Price	51:1 -88:12	13	-2.32	-5.45	-3.90	-5.10	-6.23	3.88	-3.65	-8.57	-10.09	1.38	-0.58	-4.63	20.22	32.11	37.63	51.63	10.83
Nominal	48:5 -88:12	3	-3.14	-8.95	-0.94	-9.21	-9.57	6.56	-1.81	-9.94	-9.16	4.06	-2.44	-6.40	42.90	58.70	50.18	56.67	22.76
Real Rate	51:1 -88:9	21	-3.44	-4.71	-0.32	-4.58	-9.70	5.98	-3.11	-3.40	-4.55	-2.00	-4.77	-2.55	10.54	47.08	14.15	12.38	14.73
Real Wage	67:6 -88:12	8	-2.18	-5.91	-4.47	-4.24	-5.74	2.03	-2.88	-6.54	-7.63	0.83	-1.61	-3.06	21.59	19.14	26.99	29.34	5.98

Notes:

1. *Lags* is the number of lagged values of the dependent variable included in the regression.
2. All series are in log levels, except the nominal rate (logs of gross rates) and the real rate (ex-post net rates).
3. The estimation equations include a constant, eleven seasonal dummies, a time trend, and the listed number of lags of the dependent variable.
4. Standard errors are OLS standard errors.

APPENDIX A: TECHNICAL DETAILS

A 1. Construction of $y_{k,t}$

For seasonal integration in monthly data, $(1 - B^{12})$ is the relevant polynomial. It can be factored as

$$(1 - B^{12}) = (1 - B)(1 + B)(1 + B^2)(1 + B + B^2)(1 - B + B^2)(1 + \sqrt{3}B + B^2)(1 - \sqrt{3}B + B^2)$$

with corresponding roots, $\theta_k = 1$ plus the eleven roots given in equation (2). Expanding $\varphi(B)$ about these θ_k using equation (3):

$$\begin{aligned}
\varphi(B) = {} & \lambda_1 B(1 + B)(1 + B^2)(1 + B^4 + B^8) \\
& + \lambda_2(-B)(1 - B)(1 + B^2)(1 + B^4 + B^8) \\
& + \lambda_3(-iB)(1 - iB)(1 - B^2)(1 + B^4 + B^8) \\
& + \lambda_4(iB)(1 + iB)(1 - B^2)(1 + B^4 + B^8) \\
& + \lambda_5(-\tfrac{1}{2}B)(1 - \sqrt{3}i + 2B)(1 - B + B^2)(1 - B^2 + B^6 - B^8) \\
& + \lambda_6(-\tfrac{1}{2}B)(1 + \sqrt{3}i + 2B)(1 - B + B^2)(1 - B^2 + B^6 - B^8) \\
& + \lambda_7(\tfrac{1}{2}B)(1 - \sqrt{3}i - 2B)(1 + B + B^2)(1 - B^2 + B^6 - B^8) \\
& + \lambda_8(\tfrac{1}{2}B)(1 + \sqrt{3}i - 2B)(1 + B + B^2)(1 - B^2 + B^6 - B^8) \\
& + \lambda_9(-\tfrac{1}{2}B)(\sqrt{3} - i + 2B)(1 - \sqrt{3}B + B^2)(1 + B^2 - B^6 - B^8) \\
& + \lambda_{10}(-\tfrac{1}{2}B)(\sqrt{3} + i + 2B)(1 - \sqrt{3}B + B^2)(1 + B^2 - B^6 - B^8) \\
& + \lambda_{11}(\tfrac{1}{2}B)(\sqrt{3} - i - 2B)(1 + \sqrt{3}B + B^2)(1 + B^2 - B^6 - B^8) \\
& + \lambda_{12}(\tfrac{1}{2}B)(\sqrt{3} + i - 2B)(1 + \sqrt{3}B + B^2)(1 + B^2 - B^6 - B^8) \\
& + \varphi^*(B)(1 - B^{12}).
\end{aligned} \qquad (A1)$$

Because $\varphi(B)$ is real, the pairs $(\lambda_3, \lambda_4), (\lambda_5, \lambda_6), (\lambda_7, \lambda_8), (\lambda_9, \lambda_{10})$, and $(\lambda_{11}, \lambda_{12})$ must be complex conjugates. Let π_k be defined implicitly as follows:

$$\lambda_1 = -\pi_1 \qquad \lambda_5 = \tfrac{1}{2}(-\pi_5 + i\pi_6) \qquad \lambda_9 = \tfrac{1}{2}(-\pi_9 + i\pi_{10})$$

$$\lambda_2 = -\pi_2 \qquad \lambda_6 = \tfrac{1}{2}(-\pi_5 - i\pi_6) \qquad \lambda_{10} = \tfrac{1}{2}(-\pi_9 - i\pi_{10})$$

$$\lambda_3 = \tfrac{1}{2}(-\pi_3 + i\pi_4) \qquad \lambda_7 = \tfrac{1}{2}(-\pi_7 + i\pi_8) \qquad \lambda_{11} = \tfrac{1}{2}(-\pi_{11} + i\pi_{12})$$

$$\lambda_4 = \tfrac{1}{2}(-\pi_3 - i\pi_4) \qquad \lambda_8 = \tfrac{1}{2}(-\pi_7 - i\pi_8) \qquad \lambda_{12} = \tfrac{1}{2}(-\pi_{11} - i\pi_{12}).$$

Substituting π_k for λ_k gives

$$\begin{aligned}
\varphi(B) = &- \pi_1 B(1+B)(1+B^2)(1+B^4+B^8) \\
&- \pi_2(-B)(1-B)(1+B^2)(1+B^4+B^8) \\
&- (\pi_4 + \pi_3 B)(-B)(1-B^2)(1+B^4+B^8) \\
&- \tfrac{1}{2}(\sqrt{3}\pi_6 - (1+2B)\pi_5)B(1-B+B^2)(1-B^2+B^6-B^8) \\
&- \tfrac{1}{2}(\sqrt{3}\pi_8 - (1-2B)\pi_7)(-B)(1+B+B^2)(1-B^2+B^6-B^8) \\
&- \tfrac{1}{2}(\pi_{10} - (\sqrt{3}+2B)\pi_9)B(1-\sqrt{3}B+B^2)(1+B^2-B^6-B^8) \\
&- \tfrac{1}{2}(\pi_{12} - (\sqrt{3}-2B)\pi_{11})(-B)(1+\sqrt{3}B+B^2)(1+B^2-B^6-B^8) \\
&+ \varphi^*(B)(1-B^{12}).
\end{aligned} \qquad (A2)$$

From equations (1) and (A2) one obtains regression equation (5). Note that we can rewrite equation (5) as

$$y_{1t} = \sum_{j=1}^{12} \cos(0j\pi)B^{j-1} \qquad y_{7t} = \sum_{j=1}^{12} \cos(\tfrac{j\pi}{3})B^{j-1} = -\tfrac{\sqrt{3}}{3}(y_{8t} - 2y_{8t-1})$$

$$y_{2t} = \sum_{j=1}^{12} \cos(j\pi)B^{j-1} \qquad y_{8t} = -\sum_{j=1}^{12} \sin(\tfrac{j\pi}{3})B^{j-1}$$

$$y_{3t} = \sum_{j=1}^{12} \cos(\tfrac{j\pi}{2})B^{j-1} = y_{4t-1} \qquad y_{9t} = \sum_{j=1}^{12} \cos(\tfrac{5j\pi}{6})B^{j-1} = -(\sqrt{3}y_{10t} + 2y_{10t-1})$$

$$y_{4t} = -\sum_{j=1}^{12} \sin(\tfrac{j\pi}{2})B^{j-1} \qquad y_{10t} = \sum_{j=1}^{12} \sin(\tfrac{5j\pi}{6})B^{j-1}$$

$$y_{5t} = \sum_{j=1}^{12} \cos(\tfrac{2j\pi}{3})B^{j-1} = -\tfrac{\sqrt{3}}{3}(y_{6t} + 2y_{6t-1}) \qquad y_{11t} = \sum_{j=1}^{12} \cos(\tfrac{j\pi}{6})B^{j-1} = -(\sqrt{3}y_{12t} - 2y_{12t-1})$$

$$y_{6t} = \sum_{j=1}^{12} \sin(\tfrac{2j\pi}{3})B^{j-1} \qquad y_{12t} = -\sum_{j=1}^{12} \sin(\tfrac{j\pi}{6})B^{j-1}.$$

This shows explicitly how $y_{k,t}$ is related to its particular frequency.

A 2. Proofs of Lemmas

This section finishes the proofs of the asymptotic distributions of the various statistics. In the proofs, however, a large amount of notation must be fixed, and a couple standard convergence facts established. The first two subsections set-up the proofs which are given in the third subsection.

A 2.1 Weak Convergence to Brownian Motion

The strategy is similar to that in Chan and Wei (1988), Park and Phillips (1988), Phillips (1987), and Stock (1988). The major assumption is that ϵ_t is a martingale difference sequence with constant variance equal to σ^2 and sufficient conditions on other moments such that one can apply an invariance principle. The condition $\sup_t E|\epsilon_t|^\delta < \infty$ for some $\delta > 2$ is sufficient. (See Stock, 1988 and Phillips, 1987.) In addition to avoid excessive notation, the assumption of a balanced sample, that is each month has the same number of observations, is convenient. Because an asymptotic distribution is the goal, to ensure a balanced sample one can imagine that a few observations on either end of the sample are always available. Denote the total number of observations as T and let $t = 12j + l$. $j = 0 \ldots J-1$ denotes the year and l the month. Finally, initial values do not matter asymptotically and are set to zero for convenience.

Given these assumptions one can write the partial sum of the errors in a particular month converging to a brownian motion independent of another month's. Independence follows from the ϵ_t being the fundamental Wold innovations. In addition, given the martingale difference assumption, weak convergence to stochastic integrals of the form $\int_0^1 W(r)dW(r)$ is assumed. These results are summarized in Lemma A1.

Lemma (A1):

$$
\begin{aligned}
\frac{1}{\sigma}T^{-\frac{1}{2}}\sum_{t=0}^{T}\epsilon_t &\xrightarrow{\mathcal{L}} W_1(1) \\
\frac{1}{\sigma}T^{-\frac{1}{2}}\sum_{j=0}^{J-1}\epsilon_{12j+l} &\xrightarrow{\mathcal{L}} \frac{1}{\sqrt{12}}B_l(1) \\
\frac{1}{\sqrt{12}}\sum_{l=0}^{11}B_l(r) &= W_1(r) \qquad\qquad (A3) \\
\frac{1}{\sigma}J^{-\frac{3}{2}}\sum_{m=0}^{J-1}\sum_{s=0}^{m}\epsilon_{12s+l} &\xrightarrow{\mathcal{L}} \int_0^1 B_l(r)dr \\
\frac{1}{\sigma}J^{-1}\sum_{m=0}^{J-1}\sum_{s=0}^{m}\epsilon_{12s+l-n}\epsilon_{12m+l} &\xrightarrow{\mathcal{L}} \int_0^1 B_{l-n}(r)dB_l(r),
\end{aligned}
$$

where $B_k(r)$ are mutually independent standard brownian motions. $W_1(r)$ is also a standard brownian motion.

Proof:

For the first two convergences see Park and Phillips (1988). The equality is implied by that which is above it. For the next convergence result see Park and Phillips (1988). For the last see Chan and Wei (1988) or Phillips (1988a.b). •

The $\frac{1}{\sqrt{12}}$ term is needed to insure that the variance of differences are conventional.[4]

A 2.2 Trigonometry

As is seen in section A1 of the appendix, the right-hand side regressors can be written as weighted sums of the observed series where the weights are sine or cosine functions. In deriving

[4] In other words, if $\frac{1}{\sqrt{12}}\sum_{l=1}^{12}B_l(r) = W_1(r)$ then the variance of differences must be equal also. By this one can write

$$\text{Var}\left(\frac{1}{\sqrt{12}}\sum_{l=1}^{12}B_l(s) - \frac{1}{\sqrt{12}}\sum_{l=1}^{12}B_l(r)\right) = \frac{1}{12}\sum_{l=1}^{12}(s-r) = \text{Var}(W_1(s) - W_1(r)).$$

the numerator and denominator of the various t-statistics, terms of the form

$$\sum_{l=0}^{11}\sum_{s=1}^{12}\operatorname{trg}(\theta_k(l-s))X_{l-s}Z_l$$
$$\sum_{m=1}^{J}\sum_{l=0}^{11}\Big(\sum_{s=1}^{12}\operatorname{trg}(\theta_k(l-s))X_{l-s}\Big)^2$$

are needed, where trg denotes either the sine or cosine function. The index, $k = 1\ldots 12$, corresponds to $y_1 \ldots y_{12}$, and θ_k is the frequency associated with the root used to make y_k. Note that y_4, y_8 and y_{12} are actually equal to the negative of the sum. For these three cases, one can replace the root with its negative though such details will not be expressed in the rest of this appendix. The subsequent analysis would be unchanged because of the symmetry of the distributions of $y_{4,6,8,10,12}$.

Define the vectors $\mathcal{X} \equiv [X_0, X_1 \ldots X_{11}]'$, $\mathcal{Z} \equiv [Z_0, Z_1 \ldots Z_{11}]'$. Also define the matrix

$$A_k \equiv \begin{bmatrix} \operatorname{trg}(\theta_k\cdot 0) & \operatorname{trg}(\theta_k\cdot 1) & \operatorname{trg}(\theta_k\cdot 2) & \ldots & \operatorname{trg}(\theta_k\cdot 11) \\ \operatorname{trg}(\theta_k\cdot 11) & \operatorname{trg}(\theta_k\cdot 0) & \operatorname{trg}(\theta_k\cdot 1) & \ldots & \operatorname{trg}(\theta_k\cdot 10) \\ \operatorname{trg}(\theta_k\cdot 10) & \operatorname{trg}(\theta_k\cdot 11) & \operatorname{trg}(\theta_k\cdot 0) & \ldots & \operatorname{trg}(\theta_k\cdot 9) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \operatorname{trg}(\theta_k\cdot 1) & \operatorname{trg}(\theta_k\cdot 2) & \operatorname{trg}(\theta_k\cdot 3) & \ldots & \operatorname{trg}(\theta_k\cdot 0) \end{bmatrix}. \tag{A4}$$

Similar matrices are found in Ahtola and Tiao (1987). With these definitions one can write

$$\begin{aligned} \sum_{l=0}^{11}\sum_{s=1}^{12}\operatorname{trg}(\theta_k(l-s))X_{l-s}Z_l &= \mathcal{X}'A_k\mathcal{Z} \\ \sum_{m=1}^{J}\sum_{l=0}^{11}\Big(\sum_{s=1}^{12}\operatorname{trg}(\theta_k(l-s))X_{l-s}\Big)^2 &= \mathcal{X}'A_k'A_k\mathcal{X}. \end{aligned} \tag{10}$$

Because $A_k = -A_k'$ if trg = sin and $A_k = A_k'$ if trg = cos, the second equality can be rewritten as $\mathcal{X}'A_kA_k'\mathcal{X}$.

Though it is rather inconvenient to show analytically, simple numerical calculations show that A_k has rank equal to one if $k = 1, 2$ and has rank equal to two if $k \geq 3$. Define the singular value decomposition of $A_k = U_kD_kV_k$ where D_k is diagonal with only the first or first two diagonal entries non-zero. U_k and V_k are orthonormal, and thus $U_k'U_k = V_k'V_k = I$. Replacing in (10) yields

$$\mathcal{X}'U_kD_kV_k\mathcal{Z} \quad \text{and} \quad \mathcal{X}'U_kD_kD_kU_k'\mathcal{Z}. \tag{A5}$$

The elements of X and Z will converge to simple functions of the twelve brownian motions, either $B_s(r)$, $dB_s(r)$ or $\int_0^1 B_s(r)dr$. Post multiplying X with U_k and pre-multiplying Z with V_k defines a vector of linear combinations of $B_s(r)$. Including the matrix D in the middle and multiplying leaves only one or two brownian motions.

Consider the case where $v = 1, 2$. The rank of A_k equals one with $D_{k;\,1,1} = 12$. Because D_k is all zeros except for that first entry, only the first column of U_k and the first row of V_k matter. Let $u_{k;\,c1}$ denote the first column of U_k (12×1), and $v_{k;\,r1}$ the first row of V_k (1×12).

$$u_{1\,c1} = v'_{1\,r1} = \frac{1}{\sqrt{12}}[1, 1, 1, 1, \ldots 1, 1]'$$

$$u_{2\,c1} = v'_{2\,r1} = \frac{1}{\sqrt{12}}[1, -1, 1, -1, \ldots 1, -1]'$$

For example, take $\mathcal{X} = \mathcal{Z} = B(1) = [B_0(1), B_1(1), \ldots B_{11}(1)]'$. Then

$$B(1)'U_1 D_1 V_1 B(1) = B(1)'u_{1\,c1} 12 v_{1\,r1} B(1) = 12W_1(1)^2$$

$$B(1)'U_2 D_2 V_2 B(1) = B(1)'u_{2\,c1} 12 v_{2\,r1} B(1) = 12W_2(1)^2.$$

One can verify that the $W_1(1)$ above is the same as is defined in Lemma A1. The $W_2(1)$ in the second line is by definition.

Consider now the case where $v \geq 3$. The rank of A_k equals two with $D_{k;\,1,1} = D_{k;\,2,2} = 6$. Because D_k is all zeros except for the first two entries, only the first two columns of U_k and the first two rows of V_k matter. The exact form of $u_{k;\,ci}$ differs across k, but a couple properties hold for all odd k and for all even k. Because U_k and V_k are orthonormal

$$u'_{k;\,ci} u_{k;\,ci} = v_{k;\,ri} v'_{k;\,ri} = 1; \qquad u'_{k;\,c1} u_{k;\,c2} = v_{k;\,r1} v'_{k;\,r2} = 0.$$

Moreover, for all odd k, that is trg = cos,

$$v_{k;\,r1} u_{k;\,c1} = v_{k;\,r2} u_{k;\,c2} = 1; \qquad v_{k;\,r1} u_{k;\,c2} = v_{k;\,r2} u_{k;\,c1} = 0,$$

and for all even k, that is trg = sin,

$$v_{k;\,r1} u_{k;\,c2} = 1; \qquad v_{k;\,r2} u_{k;\,c1} = -1; \qquad v_{k;\,r1} u_{k;\,c1} = v_{k;\,r2} u_{k;\,c2} = 0.$$

Because of this perfect correlation (positive or negative) for k odd, one can write

$$B(r)'u_{k;\,c1} = W_k(r) = v_{k;\,r1}B(r) \quad \text{and} \quad B(r)'u_{k;\,c2} = W_{k+1}(r) = v_{k;\,r2}B(r).$$

For k even, one can write

$$B(r)'u_{k;\,c1} = W_{k-1}(r) = -v_{k;\,r2}B(r) \quad \text{and} \quad B(r)'u_{k;\,c2} = W_k(r) = v_{k;\,r1}B(r).$$

These definitions of $W_l(r)$ for $l \geq 3$ will be the convention for what follows.

For example, take $\mathcal{X} = B(1) = [B_0(1), B_1(1), \ldots B_{11}(1)]'$ and $\mathcal{Z} = d\mathcal{X}$. Then if k is odd

$$\begin{aligned} B(1)'U_kD_kV_kB(1) &= B(1)'u_{k;\,c1}6v_{k;\,r1}dB(1) + B(1)'u_{k;\,c2}6v_{k;\,r2}dB(1) \\ &= 6W_k(1)dW_k(1) + 6W_{k+1}(1)dW_{k+1}(1). \end{aligned}$$

If k is even

$$\begin{aligned} B(1)'U_kD_kV_kB(1) &= B(1)'u_{k;\,c1}6v_{k;\,r1}dB(1) + B(1)'u_{k;\,c2}6v_{k;\,r2}dB(1) \\ &= 6W_{k-1}(1)dW_k(1) - 6W_k(1)dW_{k-1}(1). \end{aligned}$$

The specific use of these trigonometric expressions will be clearer in the proofs.

A 2.3 The Statistics

First to consider is the distribution of the t-statistics when no other terms are included in the regression, and the regression is correctly specified. The goal is to develop the asymptotic distribution of t_1, t_2, t_k and t_{k+1} where $k \in \{3, 5, 7, 9, 11\}$. It will be proved that the asymptotic distribution of the five t_k are the same and the five t_{k+1} are the same.

For all k the numerator of the t-statistic is $\sum_{t=2}^{T} y_{k,t-1}\epsilon_t$. Because $x_t - x_{t-12} = \epsilon_t$, $x_t = \sum_{j=0}^{[t/12]} \epsilon_{12j+l}$, where $l = t - 12[t/12]$ and $[\cdot]$ rounds towards zero. The denominator of the t-statistic is $\left(\sum_{t=2}^{T} y_{k,t-1}^2\right)^{\frac{1}{2}}$. Lemma 1 gives the asymptotic distribution of these t-statistics. In the proof the distribution of the numerator and denominator is established separately, which gives the total distribution by the continuous mapping theorem (Billingsley, 1968).

Lemma (1):

$$t_k = \frac{\sum_{t=2}^{T} y_{k,t-1}\epsilon_t}{\hat{\sigma}\left(\sum_{t=2}^{T} y_{k,t-1}^2\right)^{\frac{1}{2}}} \xrightarrow{\mathcal{L}} \begin{cases} \frac{\int_0^1 W_1(r)dW_1(r)}{\left(\int_0^1 W_1(r)^2 dr\right)^{\frac{1}{2}}} & \text{if } k = 1; \\ \frac{\int_0^1 W_2(r)dW_2(r)}{\left(\int_0^1 W_2(r)^2 dr\right)^{\frac{1}{2}}} & \text{if } k = 2; \\ \frac{\int_0^1 W_k(r)dW_k(r)+\int_0^1 W_{k+1}(r)dW_{k+1}(r)}{\left(\int_0^1 W_k(r)^2 dr+\int_0^1 W_{k+1}(r)^2 dr\right)^{\frac{1}{2}}} & \text{if } k \text{ odd}; \\ \frac{\int_0^1 W_k(r)dW_{k-1}(r)-\int_0^1 W_{k-1}(r)dW_k(r)}{\left(\int_0^1 W_{k-1}(r)^2 dr+\int_0^1 W_k(r)^2 dr\right)^{\frac{1}{2}}} & \text{if } k \text{ even.} \end{cases}$$

Proof:

As will be typical of what follows, since the goal is an asymptotic distribution, assume that one can observe a few observations in the past, specifically for $t = 0$, and in the future to guarantee a balanced sample. Keeping this in mind, write

$$y_{k,t-1} = \sum_{s=l-1}^{l-12} \text{trg}\big(\theta_k(t-s)\big) \sum_{j=0}^{[t/12]} \epsilon_{12j+s}.$$

The numerator of the t-statistic can be written

$$\begin{aligned} \sum_{t=0}^{T} y_{k,t-1}\epsilon_t &= \sum_{m=0}^{J}\sum_{l=0}^{11}\sum_{s=l-1}^{l-12} \text{trg}\theta_k(l-s)\sum_{j=1}^{m}\epsilon_{12j+s}\epsilon_{12m+l} \\ &= \sum_{l=0}^{11}\sum_{s=1}^{12} \text{trg}\theta_k(l-s)\sum_{m=0}^{J}\sum_{j=1}^{m}\epsilon_{12j+s}\epsilon_{12m+l} \qquad (A6) \\ &\xrightarrow{\mathcal{L}} \tfrac{\sigma^2}{12}T\sum_{l=0}^{11}\sum_{s=1}^{12} \text{trg}\theta_k(l-s)\int_0^1 B_{l-s}dB_l. \end{aligned}$$

That last step is from Lemma A1. The term $T/12$ appears because it indicates that the numerator must be normalized by J^{-1} to ensure convergence.

Let $B \equiv [B_0, B_1, \ldots B_{11}]'$ and the (12×12) matrix A be defined as in (A4). With this matrix one can rewrite (A6) as

$$\tfrac{\sigma^2}{12}T\sum_{l=0}^{11}\sum_{s=1}^{12} \text{trg}\big(\theta_k(t-s)\big)\int_0^1 B_{l-s}dB_l = \tfrac{\sigma^2}{12}T\int_0^1 B'A_k dB = \tfrac{\sigma^2}{12}T\int_0^1 B'U_kD_kV_k dB.$$

Post-multiply B by U_k and pre-multiply dB by V_k. This step is simply making linear combinations and is well-defined. Applying the analysis in the *Trigonometry* subsection finishes the asymptotics for the numerator.

As for the denominator note its square equals

$$\sum_{m=1}^{J}\sum_{l=0}^{11}\Big(\sum_{s=l-1}^{l-12}\mathrm{trg}(\theta_k(l-s))\sum_{j=1}^{m}\epsilon_{12j+s}\Big)^2.$$

Denote by $\mathcal{X}$ the vector $[\sum_{j=1}^{m}\epsilon_{12j+0},\ldots\sum_{j=1}^{m}\epsilon_{12j+11}]'$. The denominator can be rewritten as

$$\sum_{m=1}^{J}\mathcal{X}'A_k'A_k\mathcal{X},$$

or because $A'A = AA'$

$$\begin{aligned}\sum_{m=1}^{J}\mathcal{X}'A_k'A_k\mathcal{X} &= \sum_{m=1}^{J}\mathcal{X}'U_kD_kD_kU_k'\mathcal{X} \\ &\xrightarrow{\mathcal{L}} \frac{\sigma^2}{144}T^2\int_0^1 B'(r)U_kD_kD_kU_k'B(r)dr\end{aligned} \tag{A7}$$

The term $T^2/144$ appears because J^{-2} is the normalizing factor. Post-multiplying B' by U and pre-multiplying B by U' finishes the proof. Note that the brownian motions defined by this operation are the same as in the numerator. Taking the square root of (A7) gives the asymptotic distribution of the denominator, and taking the ratio of the asymptotic distribution of the numerator and the denominator gives the asymptotic distribution of the ratio by the continuous mapping theorem. Finally, by taking the ratio, the T's cancel while there is an extra σ in the numerator. This σ is canceled by dividing by $\hat{\sigma}$ in the construction of the t-statistic. That $\hat{\sigma}$ is estimated is no problem by the asymptotic equivalence lemma. •

Squaring the denominator and multiplying the numerator by two for $k \geq 3$ gives the distribution of $T\beta_k$, which is another statistic considered in the literature.

Next to consider are the t-statistics when additional terms are included in the regression. A straight forward application of Lemma 2.1 in Park and Phillips (1988) combined with the above analysis yields Lemma A2.

Lemma (A2):

$$(a) \qquad \tfrac{1}{\sigma}T^{-\frac{1}{2}}\sum_{t=0}^{T} y_{k,t-1} \quad \xrightarrow{\mathcal{L}} \quad \begin{cases} \int_0^1 W_1(r)dr & \text{if } k=1; \\ 0 & \text{if } k \geq 2; \end{cases}$$

$$(b) \qquad \tfrac{1}{\sigma}T^{-\frac{3}{2}}\sum_{t=0}^{T} t y_{k,t-1} \quad \xrightarrow{\mathcal{L}} \quad \begin{cases} \int_0^1 rW_1(r)dr & \text{if } k=1; \\ 0 & \text{if } k \geq 2; \end{cases}$$

$$(c) \qquad \tfrac{1}{\sigma}T^{-\frac{3}{2}}\sum_{t=0}^{T} t\epsilon_t \quad \xrightarrow{\mathcal{L}} \quad \int_0^1 r dW_1(r).$$

Proof (Lemma A2a):

Rewrite $\sum_{t=0}^{T} y_{k,t}$ as

$$\begin{aligned}
\sum_{t=0}^{T} y_{k,t} &= \sum_{l=1}^{12}\sum_{m=0}^{J-1} y_{k,12m+l} \\
&= \sum_{l=1}^{12}\sum_{m=0}^{J-1}\sum_{p=1}^{12}\sum_{q=0}^{m} \operatorname{trg}(12m+l-(12q+p)+1)\theta_k\epsilon_{12q+p} \\
&= \quad T^{-\frac{1}{2}}\sum_{l=1}^{12}\sum_{p=1}^{12}\operatorname{trg}(l-p+1)\theta_k\sum_{m=0}^{J-1}\sum_{q=0}^{m}\epsilon_{12q+p}.
\end{aligned}$$

Note

$$T^{-\frac{3}{2}}\sum_{m=0}^{J-1}\sum_{q=0}^{m}\epsilon_{12q+p} \quad \xrightarrow{\mathcal{L}} \quad \tfrac{1}{12\sqrt{12}}\sigma\int_0^1 B_p(r)dr$$

But for any χ_p:

$$\sum_{l=1}^{12}\sum_{p=1}^{12}\operatorname{trg}\big(\theta_k(l-s)\big)\chi_p = 0 \qquad \forall k \neq 1.$$

Thus:

$$T^{-\frac{3}{2}}\sum_{l=1}^{12}\sum_{p=1}^{12}\operatorname{trg}\theta_k(l-s)\sum_{m=0}^{J-1}\sum_{q=0}^{m}\epsilon_{12q+p} \quad \xrightarrow{\mathcal{L}} \quad \begin{cases} \frac{\sigma}{\sqrt{12}}\sum_{l=1}^{12}\int_0^1 B_l(r)dr, & \text{if } k=1; \\ 0, & \text{if } k \geq 2. \end{cases}$$

Interchanging the order of summation and integration, $\frac{\sigma}{\sqrt{12}}\sum_{l=1}^{12}\int_0^1 B_l(r)dr = \int_0^1 W_1(r)dr$, finishes the proof. •

Proof (*A2b*):

It is easiest to consider the two cases $k = 1, k \neq 1$ separately. Suppose $k = 1$ first. Then

$$T^{-\frac{5}{2}} \sum_{t=0}^{T} t y_{1t-1} = T^{-\frac{5}{2}} \sum_{t=0}^{T} \sum_{s=0}^{t} t\epsilon_s$$
$$\xrightarrow{\mathcal{L}} \int_0^1 rW_1(r)dr.$$

This result is directly from Park and Phillips (1988) or Stock (1988).

Now, let $k \neq 1$. Denote the seasonal mean of $y_{k,12m+l}$ as $\bar{y}^{\xi}_{k,l}$.

$$T^{-\frac{5}{2}} \sum_{t=0}^{T} t y_{1t-1} = T^{-\frac{5}{2}} \sum_{l=0}^{11} \sum_{m=0}^{J-1} (12m+l)\text{trg}(\theta_k(t-s)) \sum_{j=0}^{m} \epsilon_{12m+l-1}$$
$$= T^{-\frac{5}{2}} J \sum_{l=0}^{11} l\bar{y}^{\xi}_{k,l} + T^{-\frac{5}{2}} J \sum_{l=0}^{11} \sum_{p=0}^{11} \text{trg}\big(\theta_k(t-s)\big) \sum_{m=0}^{J-1} \sum_{q=0}^{m} \epsilon_{12q+p-1}.$$

The first term in the second line converges to zero as T grows large. The second term by the proof in (A2a) equals zero. •

Proof (*A2c*):

Direct application of Park and Phillips (1988) Lemma (2.1). •

Lemma A2 gives the asymptotic limits needed to derive the distribution of statistics involving a constant term or a constant and trend terms. Lemma A3 gives the asymptotics needed for regressions with seasonal dummies. Define $W_1(r)$ as in Lemma 1 and $W_2(r)$, A_k, U_k, D_k, V_k, and $W_k(r)$ as in the *Trigonometry* subsection. Define

$$\bar{\epsilon}^{\xi}_{12m+l} = \bar{\epsilon}^{\xi}_{l} \equiv J^{-1} \sum_{j=0}^{J-1} \epsilon_{12j+l} \qquad \bar{y}^{\xi}_{k,12m+l} = \bar{y}^{\xi}_{k} \equiv J^{-1} \sum_{j=0}^{J-1} y_{k,12j+l},$$

that is each is the monthly mean, or fitted values from the regression of the variable on seasonal dummies. Lemma A3 follows.

Lemma (A3):

$$\frac{1}{\sigma}T^{-1}J\sum_{l=0}^{11}\bar{y}^{\xi}_{k,l-1}\bar{\epsilon}^{\xi}_{l} \xrightarrow{\mathcal{L}} \begin{cases} W_1(1)\int_0^1 W_1(r)dr & \text{if } k=1; \\ W_2(1)\int_0^1 W_2(r)dr & \text{if } k=2; \\ \frac{1}{2}W_k(1)\int_0^1 W_k(r)dr + \frac{1}{2}W_{k+1}(1)\int_0^1 W_{k+1}(r)dr & \text{if } k \text{ odd}; \\ \frac{1}{2}W_k(1)\int_0^1 W_{k-1}(r)dr - \frac{1}{2}W_{k-1}(1)\int_0^1 W_k(r)dr & \text{if } k \text{ even}; \end{cases}$$

$$\frac{1}{\sigma}T^{-2}J\sum_{l=0}^{11}\bar{y}^{\xi\,2}_{k,l-1} \xrightarrow{\mathcal{L}} \begin{cases} \left(\int_0^1 W_1(r)dr\right)^2 & \text{if } k=1; \\ \left(\int_0^1 W_2(r)dr\right)^2 & \text{if } k=2; \\ \frac{1}{2}\left(\int_0^1 W_k(r)dr\right)^2 + \frac{1}{2}\left(\int_0^1 W_{k+1}(r)dr\right)^2 & \text{if } k \text{ odd}; \\ \frac{1}{2}\left(\int_0^1 W_{k-1}(r)dr\right)^2 + \frac{1}{2}\left(\int_0^1 W_k(r)dr\right)^2 & \text{if } k \text{ even}; \end{cases}$$

Proof:

Let

$$X_l \equiv J^{-\frac{3}{2}}\sum_{m=0}^{J-1} y_{l,12m+n} \xrightarrow{\mathcal{L}} \int_0^1 B_n(r)dr$$

$$\varepsilon_l \equiv J^{-\frac{1}{2}}\sum_{m=0}^{J-1} \epsilon_{12m+n} \xrightarrow{\mathcal{L}} B_n(1).$$

The convergence results are from Lemma 1. Let $\mathcal{X} \equiv [X_0, X_1 \ldots X_{11}]'$, and $\mathcal{E} \equiv [\varepsilon_0, \varepsilon_1 \ldots \varepsilon_{11}]'$. Write

$$\begin{aligned} T^{-1}J\sum_{l=0}^{11}\bar{y}^{\xi}_{k,l-1}\bar{\epsilon}^{\xi}_{l} &= T^{-1}J\sum_{l=0}^{11}\left(J^{-1}\sum_{m=0}^{J-1} y_{k,12m+l-1}\right)\left(J^{-1}\sum_{m=0}^{J-1}\epsilon_{12m+l}\right) \\ &= \frac{1}{12}\sum_{l=0}^{11}\left(J^{-\frac{3}{2}}\sum_{m=0}^{J-1} y_{k,12m+l-1}\right)\left(J^{-\frac{1}{2}}\sum_{m=0}^{J-1}\epsilon_{12m+l}\right) \\ &= \frac{1}{12}\mathcal{X}'A_k\mathcal{E} = \frac{1}{12}\mathcal{X}'U_kD_kV_k\mathcal{E}. \end{aligned}$$

Passing U into $\mathcal{X}$ from the right and V into $\mathcal{E}$ from the left and taking the limit finishes the proof.

To prove the second equality, replace ε_l with $y_{k,l-1}$. Then

$$T^{-2}J\sum_{l=0}^{11}\bar{y}^{\xi}_{k,l-1}\bar{\epsilon}^{\xi}_{l} = \mathcal{X}'A_kA_k'\mathcal{X}.$$

The rest of the proof follows both the proof of the first equality in the lemma and the development in the *Trigonometry* subsection. •

Armed with Lemmas A2 and A3 one can develop the asymptotics for the numerator and for the denominator. Define $y^{\mu}_{k,t}$ as the residual part of a regression of $y_{k,t}$ on a constant. Define

$\bar{y}^{\mu}_{k,t}$ as the fitted values from that regression. ($y_{k,t} = y^{\mu}_{k,t} + \bar{y}^{\mu}_{k,t}$). Analogous definition hold for $(y^{\tau}_{k,t}, \bar{y}^{\tau}_{k,t})$, $(y^{\xi}_{k,t}, \bar{y}^{\xi}_{k,t})$ and $(y^{\xi\tau}_{k,t}, \bar{y}^{\xi\tau}_{k,t})$. The numerator of the respective regressions is partially defined in Lemma 1. Let N^{μ}_k equal that part of the numerator different from that base case, likewise D^{μ}_k the denominator.

Lemma (2):

$$(\mu) \quad N^{\mu}_k \equiv \begin{cases} -W_1(1)\int_0^1 W_1(r)dr & \text{if } k=1; \\ 0 & \text{if } k \geq 2 \end{cases}$$

$$(\tau) \quad N^{\tau}_k \equiv \begin{cases} -4W_1(1)\int_0^1 W_1(r)dr + 6\int_0^1 W_1(r)dr\int_0^1 rdW_1(r) & \\ \quad -12\int_0^1 rW_1(r)dr\int_0^1 rdW_1(r) + 6W_1(1)\int_0^1 rW_1(r)dr & \text{if } k=1; \\ 0 & \text{if } k \geq 2 \end{cases}$$

$$(\xi) \quad N^{\xi}_k \equiv \begin{cases} -W_1(1)\int_0^1 W_1(r)dr & \text{if } k=1; \\ -W_2(1)\int_0^1 W_2(r)dr & \text{if } k=2; \\ -\frac{1}{2}W_k(1)\int_0^1 W_k(r)dr - \frac{1}{2}W_{k+1}(1)\int_0^1 W_{k+1}(r)dr & \text{if } k \text{ odd}; \\ -\frac{1}{2}W_k(1)\int_0^1 W_{k-1}(r)dr + \frac{1}{2}W_{k-1}(1)\int_0^1 W_k(r)dr & \text{if } k \text{ even}; \end{cases}$$

$$(\xi\tau) \; N^{\xi\tau}_k \equiv \begin{cases} -4W_1(1)\int_0^1 W_1(r)dr + 6\int_0^1 W_1(r)dr\int_0^1 rdW_1(r) & \\ \quad -12\int_0^1 rW_1(r)dr\int_0^1 rdW_1(r) + 6W_1(1)\int_0^1 rW_1(r)dr & \text{if } k=1; \\ -W_2(1)\int_0^1 W_2(r)dr & \text{if } k=2; \\ -\frac{1}{2}W_k(1)\int_0^1 W_k(r)dr - \frac{1}{2}W_{k+1}(1)\int_0^1 W_{k+1}(r)dr & \text{if } k \text{ odd}; \\ -\frac{1}{2}W_k(1)\int_0^1 W_{k-1}(r)dr + \frac{1}{2}W_{k-1}(1)\int_0^1 W_k(r)dr & \text{if } k \text{ even}. \end{cases}$$

$$(\mu) \quad D^{\mu}_k \equiv \begin{cases} -\left(\int_0^1 W_1(r)dr\right)^2 & \text{if } k=1; \\ 0 & \text{if } k \geq 2 \end{cases}$$

$$(\tau) \quad D^{\tau}_k \equiv \begin{cases} -4\left(\int_0^1 W_1(r)dr\right)^2 + 12\int_0^1 W_1(r)dr\int_0^1 rW_1(r)dr & \\ \qquad -12\left(\int_0^1 rW_1(r)dr\right)^2 & \text{if } k=1; \\ 0 & \text{if } k \geq 2 \end{cases}$$

$$(\xi) \quad D^{\xi}_k \equiv \begin{cases} -\left(\int_0^1 W_1(r)dr\right)^2 & \text{if } k=1; \\ -\left(\int_0^1 W_2(r)dr\right)^2 & \text{if } k=2; \\ -\frac{1}{4}\left(\int_0^1 W_k(r)dr\right)^2 - \frac{1}{4}\left(\int_0^1 W_{k+1}(r)dr\right)^2 & \text{if } k \text{ odd}; \\ -\frac{1}{4}\left(\int_0^1 W_{k-1}(r)\right)^2 dr - \frac{1}{4}\left(\int_0^1 W_k(r)dr\right)^2 & \text{if } k \text{ even}; \end{cases}$$

$$(\xi\tau) \; D^{\xi\tau}_k \equiv \begin{cases} -4\left(\int_0^1 W_1(r)dr\right)^2 + 12\int_0^1 W_1(r)dr\int_0^1 rW_1(r)dr & \\ \qquad -12\left(\int_0^1 rW_1(r)dr\right)^2 & \text{if } k=1; \\ -\left(\int_0^1 W_2(r)dr\right)^2 & \text{if } k=2; \\ -\frac{1}{4}\left(\int_0^1 W_k(r)dr\right)^2 - \frac{1}{4}\left(\int_0^1 W_{k+1}(r)dr\right)^2 & \text{if } k \text{ odd}; \\ -\frac{1}{4}\left(\int_0^1 W_{k-1}(r)\right)^2 dr - \frac{1}{4}\left(\int_0^1 W_k(r)dr\right)^2 & \text{if } k \text{ even}. \end{cases}$$

Proof (Lemma 2μ):

$$\sum_{t=0}^{T} y^{\mu}_{k,t-1}\epsilon^{\mu}_{t} = \sum_{t=0}^{T}(y_{k,t-1} - \bar{y}^{\mu}_{k,t-1})(\epsilon_t - \bar{\epsilon}^{\mu}_{t})$$

Since the averages are the same for all t

$$\sum_{t=0}^{T} \bar{y}^{\mu}_{k,t-1}\epsilon_t = \sum_{t=0}^{T} y_{k,t-1}\bar{\epsilon}^{\mu}_{t} = \sum_{t=0}^{T} \bar{y}^{\mu}_{k,t-1}\bar{\epsilon}^{\mu}_{t},$$

and therefore,

$$\sum_{t=0}^{T} y^{\mu}_{k,t-1}\epsilon^{\mu}_{t} = \sum_{t=0}^{T} y_{k,t-1}\epsilon_t - T\bar{y}^{\mu}\bar{\epsilon}^{\mu}.$$

Applying the proper normalization and Lemma A2 finishes the proof. •

Proof (2τ):

As in (2μ)

$$\sum_{t=0}^{T} y^{\tau}_{k,t-1}\epsilon^{\tau}_{t} = \sum_{t=0}^{T} y_{k,t-1}\epsilon_t - \sum_{t=0}^{T} \bar{y}^{\tau}_{k,t-1}\epsilon_t - \sum_{t=0}^{T} y_{k,t-1}\bar{\epsilon}^{\tau}_{t} + \sum_{t=0}^{T} \bar{y}^{\tau}_{k,t-1}\bar{\epsilon}^{\tau}_{t}.$$

Denote the OLS slope coefficient of ϵ_t on t and a constant as b_ϵ, likewise for b_y. As before ignore the fact that lags make some variables unavailable. Assume the observations are available if necessary.

$$b_\epsilon = \frac{\sum_{t=0}^{T}(\epsilon_t - \bar{\epsilon})(t - \bar{t})}{\sum_{t=0}^{T}(t - \bar{t})^2}.$$

The denominator in b_ϵ is $\sum_{t=0}^{T}(t - \bar{t})^2 = \frac{T(T+1)(2T+1)}{6} - \frac{T(T+1)^2}{4} \approx \frac{T^3}{12}$, while the numerator in b_ϵ is

$$\sum_{t=0}^{T} \epsilon_t(t - \bar{t}) = \sum_{t=0} \epsilon_t t - \bar{\epsilon}\tfrac{1}{2}(T+1).$$

This leaves

$$\sum_{t=0}^{T} y_{k,t-1}^{\tau}\epsilon_t^{\tau} = \sum_{t=0}^{T}\big(y_{k,t-1} - \bar{y} - b_y(t-\bar{t})\big)\big(\epsilon_t - \bar{\epsilon} - b_{\epsilon}(t-\bar{t})\big)$$

$$= \sum_{t=0}^{T} y_{k,t-1}\epsilon_t - \sum_{t=0}^{T}\big(\bar{y} + b_y(t-\bar{t})\big)\epsilon_t - \sum_{t=0}^{T}\big(\bar{\epsilon} + b_{\epsilon}(t-\bar{t})\big)y_{k,t-1}$$

$$+ \sum_{t=0}^{T}\big(\bar{y} + b_y(t-\bar{t})\big)\big(\bar{\epsilon} + b_{\epsilon}(t-\bar{t})\big)$$

$$= \sum_{t=0}^{T} y_{k,t-1}\epsilon_t - \sum_{t=0}^{T}\big(\bar{y} + b_y(t-\bar{t})\big)\epsilon_t$$

$$= \sum_{t=0}^{T} y_{k,t-1}\epsilon_t - 4T\bar{y}\bar{\epsilon} - 12T^{-3}\sum_{t=0}^{T} y_{k,t-1}t\sum_{t=0}^{T}\epsilon_t t$$

$$+ 6T^{-1}\bar{y}\sum_{t=0}^{T}\epsilon_t tr + 6T^{-1}\bar{\epsilon}\sum_{t=0}^{T} y_{k,t-1}t.$$

The second to last equality is from expanding the terms and noting that $\sum_{t=0}^{T}(t-\bar{t}) = 0$. Applying the proper normalization, and lemma A2 finishes the proof. •

Proof (2ξ):

$$\sum_{t=0}^{T} y_{k,t-1}^{\xi}\epsilon_t^{\xi} = \sum_{t=0}^{T} y_{k,t-1}\epsilon_t - \sum_{t=0}^{T}\bar{y}_{k,t-1}^{\xi}\epsilon_t - \sum_{t=0}^{T} y_{k,t-1}\bar{\epsilon}_t^{\xi} + \sum_{t=0}^{T}\bar{y}_{k,t-1}^{\xi}\bar{\epsilon}_t^{\xi}.$$

It is easily seen that the last two terms cancel leaving after a few simple manipulations

$$\sum_{t=0}^{T} y_{k,t-1}^{\xi}\epsilon_t^{\xi} = \sum_{t=0}^{T} y_{k,t-1}\epsilon_t - J\sum_{l=0}^{11}\bar{y}_{k,l-1}^{\xi}\bar{\epsilon}_l^{\xi},$$

where $\bar{y}_{k,l}^{\xi} = J^{-1}\sum_{m=0}^{J-1} y_{k,12m+l}$. Applying the proper normalization and lemma A3 finishes the proof. •

Proof ($2\xi\tau$):

Asymptotically the vector of seasonal dummies is orthogonal to the vector of trend terms. One can write $y_{k,t-1}^{\xi\tau} = y_{k,t-1} - \bar{y}_{k,t-1}^{\tau} - \bar{y}_{k,t-1}^{\xi} + \bar{y}_{k,t-1}^{\mu}$. The addition of the $\bar{y}_{k,t-1}^{\mu}$ is necessary because both $\bar{y}^{\tau}_{k,t-1}$ and $\bar{y}^{\xi}_{k,t-1}$ both take it out. Thus,

$$\begin{aligned}
\sum_{t=0}^{T} y^{\xi\tau}_{k,t-1}\epsilon^{\xi\tau}_t &= \sum_{t=0}^{T}(y_{k,t-1} - \bar{y}^{\tau}_{k,t-1} - \bar{y}^{\xi}_{k,t-1} + \bar{y}^{\mu}_{k,t-1})(\epsilon_t - \bar{\epsilon}^{\tau}_t - \bar{\epsilon}^{\xi}_t + \bar{\epsilon}^{\mu}_t) \\
&= \sum_{t=0}^{T}(y_{k,t-1} - \bar{y}^{\tau}_{k,t-1})(\epsilon_t - \bar{\epsilon}^{\tau}_t) + (-\bar{y}^{\xi}_{k,t-1}\epsilon_t - y_{k,t-1}\bar{\epsilon}^{\xi}_t + \bar{y}^{\xi}_{k,t-1}\bar{\epsilon}^{\xi}_t) \\
&\qquad + (\bar{y}^{\xi}_{k,t-1}\bar{\epsilon}^{\tau}_t + \bar{y}^{\tau}_{k,t-1}\bar{\epsilon}^{\xi}_t) + \bar{y}^{\mu}_{k,t-1}(\epsilon_t - \bar{\epsilon}^{\tau}_t - \bar{\epsilon}^{\xi}_t) \\
&\qquad + \bar{\epsilon}^{\mu}_t(y_{k,t-1} - \bar{y}^{\tau}_{k,t-1} - \bar{y}^{\xi}_{k,t-1}) + \bar{y}^{\mu}_{k,t-1}\bar{\epsilon}^{\mu}_t.
\end{aligned}$$

The asymptotics of the first term in parentheses is given in (2τ). The asymptotics of the second term is given in (2ξ) and equals

$$-J\sum_{l=0}^{11}\bar{y}^{\xi}_{k,l-1}\bar{\epsilon}^{\xi}_l.$$

The next term will be shown below to equal $2T\bar{y}^{\mu}\bar{\epsilon}+o_p(1)$, When the numerator is multiplied by T^{-1} the $o_p(1)$ term disappears. The final terms when added together equal $-T\bar{y}^{\mu}\bar{\epsilon}$. The asymptotics of this term is given in (2μ). Adding this up yields

$$\begin{aligned}
&\sum_{t=0}^{T} y_{k,t-1}\epsilon_t + T\big[-12T^{-\frac{5}{2}}\sum_{t=0}^{T} y_{k,t-1}t + T^{-\frac{3}{2}}\sum_{t=0}^{T}\epsilon_t t - 6T^{-\frac{3}{2}}\sum_{t=0}^{T} y_{k,t-1}T^{-\frac{1}{2}}\epsilon_t t \\
&\qquad - 6T^{-\frac{3}{2}}\sum_{t=0}^{T} y_{k,t-1}tT^{-\frac{1}{2}}\epsilon_t\big] - 3T\big[T^{-\frac{3}{2}}\sum_{t=0}^{T} y_{k,t-1}T^{-\frac{1}{2}}\epsilon_t\big] \\
&\qquad - J\big[\sum_{l=0}^{11} J^{-\frac{3}{2}}\sum_{m=0}^{J-1} y_{k,12m+l-1}J^{-\frac{1}{2}}\sum_{m=0}^{J-1}\epsilon_{12m+l}\big].
\end{aligned}$$

The asymptotics are straight forward.

The rest of the proof establishes the claim that

$$\sum_{t=0}^{T}\bar{y}^{\xi}_{k,t-1}\bar{\epsilon}^{\tau}_t + \sum_{t=0}^{T}\bar{y}^{\tau}_{k,t-1}\bar{\epsilon}^{\xi}_t = 2T\bar{\epsilon}^{\mu}\bar{y}^{\mu}_k + o_p(1).$$

Consider the first term; the second is derived analogously.

$$\sum_{t=0}^{T}\bar{y}^{\xi}_{k,t-1}\bar{\epsilon}^{\tau}_t = \bar{\epsilon}^{\mu}\sum_{t=0}^{T} y^{\xi}_{k,t-1} + b_{\epsilon}\sum_{t=0}^{T} y^{\xi}_{k,t-1}(t-\bar{t}).$$

As is shown in (2ξ), the first term equals $T\bar{\epsilon}^{\mu}\bar{y}^{\mu}$. The second term equals

$$\begin{aligned} b_\epsilon \sum_{t=0}^{T} y^{\xi}_{k,t-1}(t-\bar{t}) &= b_\epsilon \sum_{l=0}^{11} y^{\xi}_{k,l-1} \sum_{m=0}^{J-1} \left(12m + l - \tfrac{1}{2}(T+1)\right) \\ &= b_\epsilon \sum_{l=0}^{11} y^{\xi}_{k,l-1} \left(Jl - \tfrac{1}{2}(T+J)\right). \end{aligned}$$

Because the term at the end is of order T and not T^2 is why this whole term is $o_p(1)$. Factoring out the J and writing the definition of $y^{\xi}_{k,l-1}$ and of b_ϵ gives

$$\begin{aligned} &= b_\epsilon \sum_{l=0}^{11} \sum_{m=0}^{11} y_{k,12m+l-1}(l - \tfrac{13}{2}) \\ &= \left(12T^{-3} \sum_{t=0}^{T} \epsilon_t t - 6T^{-2} \sum_{t=0}^{T} \epsilon_t\right) \sum_{l=0}^{11} \sum_{m=0}^{11} y_{k,12m+l-1}(l - \tfrac{13}{2}) \\ &= \left(12T^{-\frac{3}{2}} \sum_{t=0}^{T} \epsilon_t t - 6T^{-\frac{1}{2}} \sum_{t=0}^{T} \epsilon_t\right) \sum_{l=0}^{11} T^{-\frac{3}{2}} \sum_{m=0}^{11} y_{k,12m+l-1}(l - \tfrac{13}{2}). \end{aligned}$$

In light of lemmas A1,A2 and A3, each term converges to some random variable. Multiplying by T^{-1} sends this to zero. To do its partner in the denominator, $\bar{y}^{\tau}_{k,t-1}\bar{\epsilon}^{\xi}_t$, switch ϵ and y_k, being careful about the factors of T that ensure convergence. •

With the numerator and denominator defined, the distribution of the t-statistics is simply the ratio by the continuous mapping theorem. These are given in Lemma 3 in the text.

Table A1

Critical Values from the Distributions of Test Statistics for Seasonal Unit Roots, data generating process $\Delta_{12} x_t = \epsilon_t$ nid(0,1)

Auxilliary Regressions	T	't': π_1 0.01	0.025	0.05	0.10	't': π_2 0.01	0.025	0.05	0.10	't': π_{odd} 0.01	0.025	0.05	0.10
No intercept	240	-2.51	-2.18	-1.89	-1.58	-2.53	-2.16	-1.87	-1.57	-2.50	-2.16	-1.88	-1.55
No seas.dum.	480	-2.52	-2.21	-1.91	-1.59	-2.52	-2.20	-1.91	-1.59	-2.52	-2.18	-1.90	-1.57
No trend	∞	-2.57	-2.24	-1.95	-1.62	-2.57	-2.24	-1.95	-1.62	-2.56	-2.23	-1.95	-1.59
Intercept	240	-3.35	-3.06	-2.80	-2.51	-2.48	-2.15	-1.89	-1.57	-2.51	-2.16	-1.87	-1.54
No seas.dum.	480	-3.40	-3.11	-2.85	-2.55	-2.54	-2.20	-1.91	-1.59	-2.56	-2.20	-1.90	-1.57
No trend	∞	-3.41	-3.12	-2.86	-2.57	-2.57	-2.24	-1.95	-1.62	-2.56	-2.23	-1.95	-1.59
Intercept	240	-3.32	-3.02	-2.76	-2.47	-3.28	-3.01	-2.76	-2.48	-3.83	-3.51	-3.25	-2.95
Seas.dum.	480	-3.37	-3.06	-2.81	-2.53	-3.37	-3.07	-2.81	-2.52	-3.86	-3.55	-3.29	-2.99
No trend	∞	-3.41	-3.12	-2.86	-2.57	-3.41	-3.12	-2.86	-2.57	-3.91	-3.61	-3.35	-3.05
Intercept	240	-3.87	-3.58	-3.32	-3.06	-2.52	-2.18	-1.88	-1.55	-2.49	-2.16	-1.88	-1.54
No seas.dum.	480	-3.92	-3.63	-3.37	-3.09	-2.55	-2.20	-1.93	-1.60	-2.53	-2.20	-1.91	-1.57
Trend	∞	-3.97	-3.67	-3.40	-3.12	-2.57	-2.24	-1.95	-1.62	-2.56	-2.23	-1.95	-1.59
Intercept	240	-3.83	-3.54	-3.28	-2.99	-3.31	-3.02	-2.75	-2.47	-3.79	-3.50	-3.24	-2.95
Seas.dum.	480	-3.85	-3.57	-3.32	-3.04	-3.40	-3.08	-2.84	-2.54	-3.85	-3.55	-3.29	-3.00
Trend	∞	-3.97	-3.67	-3.40	-3.12	-3.41	-3.12	-2.86	-2.57	-3.91	-3.61	-3.35	-3.05

Auxilliary Regressions	T	't': π_{even} 0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99	'F': π_{odd}, π_{even} 0.90	0.95	0.975	0.99
No intercept	240	-2.31	-1.95	-1.63	-1.27	1.25	1.61	1.93	2.29	2.34	3.03	3.71	4.60
No seas.dum.	480	-2.33	-1.96	-1.65	-1.28	1.27	1.63	1.94	2.32	2.38	3.08	3.78	4.70
No trend	∞	-2.30	-1.94	-1.63	-1.28	1.27	1.63	1.94	2.32	2.40	3.10	3.79	4.68
Intercept	240	-2.30	-1.93	-1.62	-1.27	1.24	1.60	1.91	2.28	2.32	3.01	3.68	4.60
No seas.dum.	480	-2.32	-1.95	-1.63	-1.27	1.27	1.62	1.93	2.30	2.36	3.06	3.76	4.66
No trend	∞	-2.30	-1.94	-1.63	-1.28	1.27	1.63	1.94	2.32	2.40	3.10	3.79	4.68
Intercept	240	-2.61	-2.21	-1.85	-1.45	1.46	1.86	2.20	2.60	5.27	6.26	7.19	8.35
Seas.dum.	480	-2.65	-2.25	-1.90	-1.49	1.49	1.91	2.25	2.63	5.42	6.42	7.38	8.60
No trend	∞	-2.72	-2.31	-1.95	-1.54	1.53	1.95	2.30	2.72	5.64	6.67	7.63	8.79
Intercept	240	-2.28	-1.93	-1.61	-1.25	1.24	1.59	1.90	2.26	2.30	2.97	3.64	4.53
No seas.dum.	480	-2.30	-1.94	-1.63	-1.27	1.25	1.61	1.92	2.28	2.36	3.05	3.72	4.62
Trend	∞	-2.30	-1.94	-1.63	-1.28	1.27	1.63	1.94	2.32	2.40	3.10	3.79	4.68
Intercept	240	-2.57	-2.18	-1.85	-1.45	1.45	1.86	2.19	2.60	5.25	6.23	7.14	8.33
Seas.dum.	480	-2.66	-2.27	-1.91	-1.49	1.49	1.90	2.25	2.64	5.44	6.43	7.35	8.52
Trend	∞	-2.72	-2.31	-1.95	-1.54	1.53	1.95	2.30	2.72	5.64	6.67	7.63	8.79

Table A2: Size of Monthly Unit Root Tests, Probability of Rejecting

$$(1 - B^{12})x_t = (1 + \rho B)\eta_t$$

		Left Side t-Statistics											
		0	π	$\frac{\pi}{2}$		$\frac{2\pi}{3}$		$\frac{\pi}{3}$		$\frac{5\pi}{6}$		$\frac{\pi}{6}$	
ρ	Percentile	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}
-.85	5%	.049	.034	.011	.172	.017	.006	.016	.085	.033	.004	.035	.057
.85	5%	.035	.044	.010	.008	.013	.085	.018	.006	.033	.053	.033	.005

		Right Side t-Statistics					*F-Statistics*				
		$\frac{\pi}{2}$	$\frac{2\pi}{3}$	$\frac{\pi}{3}$	$\frac{5\pi}{6}$	$\frac{\pi}{6}$	$\frac{\pi}{2}$	$\frac{2\pi}{3}$	$\frac{\pi}{3}$	$\frac{5\pi}{6}$	$\frac{\pi}{6}$
ρ	Percentile	t_4	t_6	t_8	t_{10}	t_{12}	$F_{3,4}$	$F_{5,6}$	$F_{7,8}$	$F_{9,10}$	$F_{11,12}$
-.85	95%	.009	.197	.023	.134	.042	.020	.032	.018	.039	.036
.85	95%	.171	.025	.191	.036	.142	.020	.014	.036	.034	.039

$$(1 - B^{12})x_t = (1 + \rho B^{12})\eta_t$$

		Left Side t-Statistics											
		0	π	$\frac{\pi}{2}$		$\frac{2\pi}{3}$		$\frac{\pi}{3}$		$\frac{5\pi}{6}$		$\frac{\pi}{6}$	
ρ	Percentile	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}
-.85	5%	.044	.045	.015	.054	.021	.036	.022	.040	.034	.034	.035	.037

		Right Side t-Statistics					*F-Statistics*				
		$\frac{\pi}{2}$	$\frac{2\pi}{3}$	$\frac{\pi}{3}$	$\frac{5\pi}{6}$	$\frac{\pi}{6}$	$\frac{\pi}{2}$	$\frac{2\pi}{3}$	$\frac{\pi}{3}$	$\frac{5\pi}{6}$	$\frac{\pi}{6}$
ρ	Percentile	t_4	t_6	t_8	t_{10}	t_{12}	$F_{3,4}$	$F_{5,6}$	$F_{7,8}$	$F_{9,10}$	$F_{11,12}$
-.85	95%	.054	.056	.054	.059	.059	.015	.022	.024	.035	.036

Notes to Table A2:

1. Left and Right hand statistics are t-statistics; F-statistics are the joint statistics calculated as an F-statistic. The percentiles give the percentage of the area underneath the simulated density to the left of the listed statistic. These statistics are based on 24,000 regressions with 20 years of data.
2. η_t is standard normal.
3. The regressions include twelve lags of the dependent variable, a constant and eleven seasonal dummies, but no trend.

Table A3: Power of Monthly Unit Root Tests, Probability of Rejecting

$$(1-\rho^{12}B^{12})y_t = \epsilon_t$$

		Left Side t-Statistics											
		0	π	$\frac{\pi}{2}$		$\frac{2\pi}{3}$		$\frac{\pi}{3}$		$\frac{5\pi}{6}$		$\frac{\pi}{6}$	
ρ	Percentile	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}
.95	5%	.447	.448	.796	.033	.795	.033	.795	.032	.794	.034	.794	.034
.85	5%	.987	.988	1.000	.028	1.000	.029	1.000	.029	1.000	.029	1.000	.028

		Right Side t-Statistics					*F-Statistics*				
		$\frac{\pi}{2}$	$\frac{2\pi}{3}$	$\frac{\pi}{3}$	$\frac{5\pi}{6}$	$\frac{\pi}{6}$	$\frac{\pi}{2}$	$\frac{2\pi}{3}$	$\frac{\pi}{3}$	$\frac{5\pi}{6}$	$\frac{\pi}{6}$
ρ	Percentile	t_4	t_6	t_8	t_{10}	t_{12}	$F_{3,4}$	$F_{5,6}$	$F_{7,8}$	$F_{9,10}$	$F_{11,12}$
.95	95%	.031	.032	.030	.030	.031	.716	.713	.713	.714	.711
.85	95%	.028	.027	.025	.024	.025	1.000	1.000	1.000	1.000	1.000

$$(1-B^4)(1+\rho^4B^4+\rho^8B^8)y_t = \epsilon_t$$

		Left Side t-Statistics											
		0	π	$\frac{\pi}{2}$		$\frac{2\pi}{3}$		$\frac{\pi}{3}$		$\frac{5\pi}{6}$		$\frac{\pi}{6}$	
ρ	Percentile	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}
.95	5%	.057	.052	.054	.052	1.000	.000	1.000	.000	1.000	.979	1.000	.979
.85	5%	.057	.058	.054	.052	1.000	.000	1.000	.000	1.000	.977	1.000	.976

		Right Side t-Statistics					*F-Statistics*				
		$\frac{\pi}{2}$	$\frac{2\pi}{3}$	$\frac{\pi}{3}$	$\frac{5\pi}{6}$	$\frac{\pi}{6}$	$\frac{\pi}{2}$	$\frac{2\pi}{3}$	$\frac{\pi}{3}$	$\frac{5\pi}{6}$	$\frac{\pi}{6}$
ρ	Percentile	t_4	t_6	t_8	t_{10}	t_{12}	$F_{3,4}$	$F_{5,6}$	$F_{7,8}$	$F_{9,10}$	$F_{11,12}$
.95	95%	.053	.966	.963	.000	.000	.053	1.000	1.000	1.000	1.000
.85	95%	.051	.965	.964	.000	.000	.051	1.000	1.000	1.000	1.000

Notes to Table A3:

1. Left and Right hand statistics are t–statistics; F–statistics are the joint statistics calculated as an F-statistic. The percentiles give the percentage of the area underneath the simulated density to the left of the listed statistic. These statistics are based on 24,000 regressions with 20 years of data.
2. ϵ_t is standard normal.
3. The regressions include a constant and eleven seasonal dummies but no trend.

APPENDIX B: DATA

The data we use are seasonally unadjusted, U.S. macroeconomic time series for the post-WWII period. All of the data were obtained through Data Resources, Incorporated. All series are measured in log levels, except for nominal rates (log of gross rates) and real rates (ex-post net rates).

The quarterly series are from the Bureau of Economic Analysis. They are Gross National Product, Personal Consumption Expenditures, Gross Private Domestic Fixed Investment, and Government Purchases of Goods and Services. All of these series are deflated by the CPI, which is from the Bureau of Labor Statistics (BLS).

The monthly series on real activity are Retail Sales, the Industrial Production Index, and the Unemployment Rate. The first two are from the Board of Governors while the third is from BLS. The Wage series is Average Hourly Earnings in Private, Non-Agricultural employment, also from BLS. The money stock is M1 and the Nominal Rate is the rate on 3-Month T-Bills; both series are from the Board of Governors. Retail Sales and Wages are deflated by the C.P.I.