Journal of Business & Economic Statistics, July 1995, Vol. 13, No. 3

Are Seasonal Patterns Constant Over Time? A Test for Seasonal Stability

Fabio Canova
Department of Economics, Universitat Pompeu Fabra, E-08008 Barcelona, Spain, and Department of Economics, University of Catania, I-95100 Catania, Italy

Bruce E. Hansen
Department of Economics, Boston College, Chestnut Hill, MA 02167, and Department of Economics, University of Rochester, Rochester, NY 14627

This article introduces Lagrange multiplier tests of the null hypothesis of no unit roots at seasonal frequencies against the alternative of a unit root at either a single seasonal frequency or a set of seasonal frequencies. The tests complement those of Dickey, Hasza, and Fuller and Hylleberg, Engle, Granger, and Yoo that examine the null of seasonal unit roots. We derive an asymptotic distribution theory for the tests, and investigate their size and power with a Monte Carlo exercise. Application of the tests to three sets of seasonal variables shows that in most cases seasonality is nonstationary.

KEY WORDS: Instability; Lagrange multiplier tests; Monte Carlo; Parameter constancy; Seasonality; Unit roots.

The study of seasonal fluctuations has a long history in the analysis of economic time series. Traditionally, seasonal fluctuations have been considered a nuisance that obscures, the more important components of the series [presumably the growth and cyclical components; e.g., see Burns and Mitchell (1946)], and seasonal adjustment procedures have been devised and implemented to eliminate seasonality (Shiskin, Young, and Musgrave 1967). This view dominated applied time series econometrics until quite recently. In the past few years a new viewpoint has emerged. Seasonal fluctuations are not necessarily a nuisance, but they are an integral part of economic data and should not be ignored or obscured in the analysis of economic models. Contributors to this view include Ghysels (1988), Barsky and Miron (1989), Braun and Evans (1990), Chattarjee and Ravikumar (1992), and Hansen and Sargent (1993).

Many reasonable time series models of seasonality are conceivable. One approach is to model seasonality as deterministic, as did Barsky and Miron (1989), or as periodic with unchanged periodicity, as done by Hansen and Sargent (1993). A second approach is to model seasonality as the sum of a deterministic process and a stationary stochastic process (Canova 1992). A third approach is to model seasonal patterns as nonstationary by allowing for (or imposing) seasonal unit roots (Box and Jenkins 1976).

It is hard to know a priori which approach yields the best statistical description of the data. The assumption of stable seasonal patterns seems reasonable when one considers that Christmas has been in December for as many years as we can remember and that this period is historically the major retail season. On the other hand, selected series have shown monumental changes in the seasonal patterns, in which even the location of seasonal peaks and troughs has changed. Examples include the energy consumption series examined by Engle, Granger, and Hallman (1989), the Japanese consumption and income series examined by Engle, Granger, Hylleberg, and Lee (1993), the industrial production series examined by Canova (1993), and some of the gross domestic product series analyzed by Hylleberg, Jorgensen, and Sorensen (1993). Although there certainly are examples of such large changes in seasonal patterns, one might conjecture that they are relatively rare events and isolated to just a few of the many aggregate macroeconomic series.

It is unsatisfactory to rely on hunches, intuition, stylized facts, and/or ad hoc statistical techniques to determine which statistical model makes the best fit. We need simple statistical techniques that can discriminate between various forms of seasonality. One such testing framework was introduced by Dickey, Hasza, and Fuller (DHF) (1984) and Hylleberg, Engle, Granger, and Yoo (HEGY) (1990), who generalized the unit-root testing methodology of Dickey and Fuller (1979) to the seasonal case. They took the null hypothesis of a unit root at one or more seasonal frequencies and tested for evidence of stationarity. Rejection of their null hypothesis implies the strong result that the series has a stationary seasonal pattern. Due to the low power of the tests in moderate sample sizes, however, nonrejection of the null hypothesis unfortunately cannot be interpreted as evidence "for" the presence of a seasonal unit root.

A useful complement to the preceding testing methodology would be tests that take the null hypothesis to be *stationary* seasonality and the alternative to be *nonstationary* seasonality. In this context, rejection of the null hypothesis would imply the strong result that the data are indeed nonstationary, a conclusion that the DHF or HEGY tests cannot yield. Viewed jointly with these tests, such a procedure would allow researchers a more thorough analysis of their data. A family of such tests is introduced and studied in this article. Even though our null is stationary seasonality, we will for simplicity refer to our tests as seasonal unit-root tests.

The idea is perfectly analogous to that of testing the null of stationarity against the alternative of a unit root at the zero frequency. A Lagrange multiplier (LM) statistic for this null and alternative was recently proposed by Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) (1992). The KPSS test is analogous to the tests of Tanaka (1990) and Saikkonen and Luukkonen (1993), who examined the null hypothesis of a moving average (MA) unit root. In the same sense that HEGY generalized the Dickey–Fuller framework from the zero frequency to the seasonal frequencies, we generalize the KPSS framework from the zero frequency to the seasonal frequencies.

Another set of tests, which may appeal to applied macroeconomists, is to examine whether or not seasonal patterns can be accurately represented with a set of deterministic functions of time. Within our framework, we can also introduce tests of the proposition that the seasonal intercepts are constant over time. Under the null hypothesis of stationarity, seasonal intercepts represent the deterministic component of seasonality and are assumed to remain constant over the sample. In this case our tests apply the methodology of Nyblom (1989) and Hansen (1990), who designed LM tests for parameter instability. Interestingly, the LM test for joint instability of the seasonal intercepts is numerically identical to the LM test for unit roots at all seasonal frequencies. Thus the test we describe can be viewed as either a test for seasonal unit roots or for instability in the seasonal pattern, and both views are equally correct.

Our test statistics are precisely LM tests in models with iid Gaussian errors. Because this is not a reasonable assumption for time series applications, we show how to modify the test statistics (by using robust covariance matrix estimates) so that the tests can be applied to a wide class of data, including heteroscedastic and serially correlated processes. We only require relatively weak mixing conditions on the data. It is important to note that we exclude from the regression any trending regressors, such as a unit-root process or a deterministic trend. This is not simply for technical reasons because it is possible to show that the asymptotic distribution is not invariant to such variables. We also require that our dependent variable be free of trends. Thus, we are presuming that the data have already been appropriately transformed to eliminate unit roots at the zero frequency.

The test statistics are derived from the LM principle, which requires only estimation of the model under the null, so least squares techniques are all that is needed. The statistics are fairly simple functions of the residuals. The large-sample distributions are nonstandard but are free from nuisance parameters and only depend on one "degrees-of-freedom" parameter.

To study both the power and the size of the proposed tests, we conducted a Monte Carlo exercise, and we compared their performance with two other standard types of tests, a test for the presence of stochastic (stationary) seasonality and the HEGY tests for unit roots at seasonal frequencies. We show that our tests have reasonable size and power properties.

We apply the tests to the data set originally examined by Barsky and Miron (1989). We are interested in establishing if their maintained hypothesis that quarterly seasonal fluctuations in U.S. macrovariables are well approximated by deterministic patterns is appropriate or not. The second data set used is the set of quarterly industrial production indexes for eight industrialized countries used by Canova (1993). The third is a data set on stock returns on value-weighted indexes for seven industrialized countries. This last data set deserves special attention because "January effects" and other abnormal periodic patterns in stock returns have been repeatedly documented and known for a long time [see Thaler (1987) for a survey of these anomalies]. It is therefore of interest to examine whether the knowledge of these patterns has changed their properties or, in other words, if information about the existence of periodic patterns has led to structural changes due to profit-taking activities. We show that for 20 of the 25 series examined by Barsky and Miron the assumption of unchanged seasonality is problematic and that, in some cases, the economic significance of these changes is substantial. Similarly the seasonal patterns of the European industrial production indexes display important instabilities. On the other hand, we find that the seasonal pattern of stock returns has substantially changed only in Japan and in the United Kingdom.

The rest of the article is organized as follows. Section 1 describes the regression model. Two methods to model the deterministic component of seasonality are discussed. Section 2 derives LM tests for unit roots at seasonal frequencies and develops an asymptotic theory of inference for the tests. Section 3 presents LM tests for instability in the seasonal intercepts. Section 4 presents a Monte Carlo exercise. Three applications to economic data appear in Section 5. Conclusions are summarized in Section 6.

1. REGRESSION MODELS WITH STATIONARY SEASONALITY

1.1 Regression Equation

We start from a linear time series model with stationary seasonality:

$$y_i = \mu + x_i'\beta + S_i + e_i, \qquad i = 1, 2, \ldots, n. \tag{1}$$

In (1), y_i is real valued, x_i is a $k \times 1$ vector of explanatory variables, S_i is a real-valued deterministic seasonal component of period s, where s is a positive even integer, to be discussed in Subsection 1.2, and $e_i \sim (0, \sigma^2)$ is an error uncorrelated with x_i and S_i. The number of observations is n. If there are exactly T years of data, then $n = Ts$.

To distinguish between nonstationarity at a seasonal frequency and at the zero frequency, we must require that the dependent variable y_i not have a unit root at the zero frequency (or any other form of nonstationarity in the overall mean). This does not restrict the set of possible applications of the tests because it is widely believed that most macroeconomic time series are stationary at the zero frequency after suitable transformations, such as taking the first-difference of the natural logarithm. For some series, such as the price level, double-differencing may be necessary to eliminate the zero frequency unit root. In either case, the deterministic seasonal component S_i of the differenced series y_i can be related to the seasonal component of the original undifferenced series through a result of Pierce (1978, theorem 1).

The regressors x_i may be any nontrending variables that satisfy standard weak dependence conditions. To identify the regression parameters β, we exclude from x_i any variables that are collinear with S_i. In many cases, no x_i will be included. One suggestion we discuss in Section 1.3 is to use the first lag of the dependent variable, $x_i = y_{i-1}$.

When there are no regressors x_i, the error e_i represents the deviation y_i from its seasonal mean. Thus, e_i includes all of the random variation in the dependent variable, will be serially correlated, and may include fluctuations that are seasonal in nature. Because we have no desire to exclude a priori stationary stochastic seasonal patterns under the null hypothesis, our distributional theory is derived under mild mixing-type conditions for the error term e_i that allow for general forms of stochastic behavior, including stationary (and mildly heteroscedastic) stochastic seasonality.

1.2 Modeling Deterministic Seasonal Patterns

A common specification for the deterministic seasonal component in (1) is

$$S_i = d_i'\alpha, \tag{2}$$

where d_i is an $s \times 1$ vector of seasonal dummy indicators and α is an $s \times 1$ parameter vector (e.g., $s = 4$ for quarterly data and $s = 12$ for monthly data). Combined with (1), we obtain the regression model

$$y_i = x_i'\beta + d_i'\alpha + e_i, \qquad i = 1, 2, \dots, n, \tag{3}$$

where we have dropped the intercept μ from the model to achieve identification. The advantage of this formulation is that the coefficients α represent *seasonal* effects. Plotting recursive estimates of the subcoefficients α_s against time is often used to reveal the structure of seasonal patterns (Franses 1994).

A mathematically equivalent formulation is obtained using the trigonometric representation

$$S_i = \sum_{j=1}^{q} f_{ji}'\gamma_j, \tag{4}$$

where $q = s/2$, and for $j < q$, $f_{ji}' = (\cos((j/q)\pi i), \sin((j/q)\pi i))$, while for $j = q$, $f_{qi} = \cos(\pi i) = (-1)^i$, where the latter holds because $\sin(\pi i)$ is identically 0 for all integer i. Stacking the q elements of (4) in a vector, we have $S_i = f_i'\gamma$, where

$$\gamma = \begin{pmatrix} \gamma_1 \\ \cdot \\ \cdot \\ \gamma_q \end{pmatrix}, \quad f_i = \begin{pmatrix} f_{1i} \\ \cdot \\ \cdot \\ f_{qi} \end{pmatrix}. \tag{5}$$

Note that both γ and f_i have $s - 1$ elements. Inserting in (1), we have the regression equation

$$y_i = \mu + x_i'\beta + f_i'\gamma + e_i, \qquad i = 1, 2, \dots, n. \tag{6}$$

Note that f_i is a mean zero process because for any n that is an integer power of s, $\sum_{i=1}^{n} f_i = 0$. It is also a full-rank process (by the properties of trigonometric series) so that if we define the $s \times (s - 1)$ matrix

$$R_1 = \begin{pmatrix} f_1' \\ \cdot \\ \cdot \\ f_s' \end{pmatrix},$$

then $f_i = R_1'd_i$. Because $1 = \iota' d_i$, where ι is an $(s \times 1)$ vector of ones, we have $(f_i' \; 1)' = R'd_i$, where $R = [R_1 \; \iota]$. Thus, (3) and (6) are equivalent, and $(\gamma' \mu)' = R^{-1}\alpha$.

The formulation (6) is useful because it allows seasonality to be interpreted as *cyclical.* By construction, the elements of f_i are cyclical processes at the seasonal frequencies: $(j/q)\pi$, $j = 1, \dots, q$, and the coefficients γ_j represent the contribution of each cycle to the seasonal process S_i. The dummy formulation (2) is primarily employed in applied macroeconomics (Barsky and Miron 1989), but the trigonometric representation (4) is common in the time series literature (e.g., see Granger and Newbold 1986, p. 36; Hannan 1970, p. 174; Harvey 1990, p. 42).

1.3 Lagged Dependent Variables

The distribution theory we present in Sections 2 and 3 will not be affected if the regressors x_i include lagged dependent variables. But if lagged variables capture one or more seasonal unit roots, the tests we present may have no power. Essentially, what must be excluded are lags of the dependent variable that capture seasonal unit roots. This may be easier to see if we take (1), where the x_i are exclusively lags of the dependent variable:

$$y_i = \mu + \beta(\ell)y_{i-1} + S_i + e_i, \tag{7}$$

with $\beta(\ell) = \beta_1 + \cdots + \beta_\zeta \ell^{\zeta-1}$. When $\zeta = 1$ and $\beta_1 \neq -1$, the autoregressive polynomial $\beta(\ell)$ will not be able to extract a seasonal pattern from y_i. But if $\zeta \geq 2$, $\beta(\ell)$ may absorb some of the seasonal roots. Thus, testing for a seasonal unit root in S_i will be useless.

This discussion should not be interpreted as suggesting that all lagged dependent variables should be excluded from x_i. Indeed, exclusion of lagged dependent variables means that the error e_i will be serially correlated in most applications. Because the inclusion of a single lag of the dependent variable in x_i will reduce this serial correlation (we can think of this as a form of pre-whitening), yet not pose a danger of extracting a seasonal root, we recommend that x_i contain y_{i-1}. The fact that the e_i may be serially correlated will be accounted for at the stage of inference.

1.4 Estimation and Covariance Matrices

Both (3) and (6) are valid regression equations and can be consistently estimated (under standard regularity conditions) by ordinary least squares (OLS). Let the estimates from (3) be denoted $(\widehat{\beta}, \widehat{\alpha})$ and the estimates from (6) be denoted $(\widehat{\mu}, \widehat{\beta}, \widehat{\gamma})$. Due to the equivalence in parameterization, the estimates of β and the regression residuals $\widehat{e}_i$ in the two equations are identical.

Our tests will require a consistent estimate of Ω, the long-run covariance matrix of $d_i e_i$:

$$\Omega = \lim_{n\to\infty} \frac{1}{n} E(D_n D_n'), \qquad D_n = \sum_{i=1}^{n} d_i e_i,$$

where Ω depends on the stochastic structure of the error. When e_i is serially uncorrelated and seasonally homoscedastic, then

$$\Omega = \frac{1}{n}\sum_{i=1}^{n} d_i d_i' E(e_i^2) = \frac{\sigma^2}{s} I_s$$

(since $(1/n)\sum_{i=1}^{n} d_i d_i' = I_s/s$), which can be estimated by $\widehat{\Omega} = (\widehat{\sigma}^2/s)I_s$. When e_i is uncorrelated but possibly seasonally heteroscedastic, then $\Omega = \mathrm{diag}\{\sigma_1^2/s, \ldots, \sigma_s^2/s\}$, where σ_j^2 is the variance of e_i for the jth season. In many cases of interest, however, e_i is likely to be serially correlated so that we need an estimate of Ω that is robust to serial correlation as well. Following Newey and West (1987), we suggest a kernel estimate of the form

$$\widehat{\Omega} = \sum_{k=-m}^{m} w\left(\frac{k}{m}\right) \frac{1}{n} \sum_{i} d_{i+k}\widehat{e}_{i+k} d_i' \widehat{e}_i, \tag{8}$$

where $w(\cdot)$ is any kernel function that produces positive semidefinite covariance matrix estimates, such as the Bartlett, Parzen, or quadratic spectral. It is desirable to select the bandwidth number m sufficiently large to be able to capture the serial correlation properties of the data. Andrews (1991) proposed methods for minimum mean squared error estimation of such variances, and Hansen (1992a) gave sufficient conditions for consistent estimation. Here we assume that $m \to \infty$ as $n \to \infty$ such that $m^5/n = O(1)$ as recommended by Andrews (1991) for the efficient Parzen and quadratic spectral kernels.

We will also require an estimate of the long-run covariance matrix of $f_i e_i$:

$$\Omega^f = \lim_{n\to\infty} \frac{1}{n} E(F_n F_n'), \qquad F_n = \sum_{i=1}^{n} f_i e_i.$$

Since $f_i = R_1' d_i$, we see that $F_n = R_1' D_n$, and thus $\Omega^f = R_1' \Omega R_1$. Hence a consistent estimate is given by

$$\widehat{\Omega}^f = R_1' \widehat{\Omega} R_1 = \sum_{k=-m}^{m} w\left(\frac{k}{m}\right) \frac{1}{n} \sum_{i} f_{i+k}\widehat{e}_{i+k} f_i' \widehat{e}_i. \tag{9}$$

2. TESTING FOR SEASONAL UNIT ROOTS

2.1 The Testing Problem

Our goal here is to develop tests of the hypothesis that (6) is valid against the alternative that there is a seasonal unit root in S_i. To do so rigorously, we have to write down a specific alternative hypothesis. Hannan (1970, p. 174) suggested that one reasonable model for changing seasonal patterns can be obtained by allowing the coefficients γ to vary over time as a random walk, in which case (6) is

$$y_i = \mu + x_i'\beta + f_i'\gamma_i + e_i, \tag{10}$$

with

$$\gamma_i = \gamma_{i-1} + u_i, \tag{11}$$

γ_0 fixed, and u_i iid. If the covariance matrix of u_i is full ranked, then Models (10)–(11) imply that y_i has unit roots at each seasonal frequency. His model reduces to the stationary seasonal model (6) when the covariance matrix of u_i is identically 0.

We would like to generalize Hannan's model to allow for unit roots potentially at only a subset of the seasonal frequencies. This is equivalent to allowing only a subset of the vector γ_i to be time varying. We can do so by defining a full-rank $(s-1)\times a$ matrix A that selects the a elements of γ_i that we wish to test for nonstationarity. For example, to test whether the entire vector γ is stable, set $A = I_{s-1}$, and to test for a unit root only at frequency $j/q\pi$, set $A = (\widetilde{0}\, I_2\, \widetilde{0})'$ (commensurate with γ_i) and for frequency π, set $A = (\widetilde{0}\ 1)'$. Then modify (11) as

$$A'\gamma_i = A'\gamma_{i-1} + u_i. \tag{12}$$

We assume that, for some increasing sequence of sigma fields $\mathcal{F}_i$, $\{u_i, \mathcal{F}_i\}$ is an $a\times 1$ martingale difference sequence (MDS) with covariance matrix $E(u_i u_i') = \tau^2 G$, where $G = (A'\Omega^f A)^{-1}$ is a full-rank $a \times a$ matrix and $\tau^2 \geq 0$ is real valued. When $\tau^2 = 0$, the parameter $\gamma_i = \gamma_0$ and the model has no seasonal unit roots. When $\tau^2 > 0$, y_i has a unit root at the seasonal frequencies determined by A.

Our model specification is closely related to the trigonometric seasonal model of Harvey [1990, eq. (2.3.49)]. For quarterly data ($s = 4$) Harvey specified $S_i = S_{1i} + S_{2i}$, where $(1+\ell^2)S_{1i} = \xi_{1i}$, $(1+\ell)S_{2i} = \xi_{2i}$, and ξ_{1i} and ξ_{2i} are mutually independent. Hence his model also produced unit roots at the seasonal frequencies but imposed a somewhat different set of correlations across seasonal fluctuations. Bell (1993) showed that the models of Hannan and Harvey generate equivalent structures for the seasonal components.

2.2 The Hypothesis Test

When $\tau^2 = 0$, S_i is purely deterministic and stationary in the models (10)–(12). This suggests considering the hypothesis test of $H_0: \tau^2 = 0$ against $H_1: \tau^2 > 0$. Nyblom (1989) showed that this testing problem is particularly easy to implement in a correctly specified probability model using maximum likelihood estimation. Hansen (1990) extended Nyblom's analysis to general econometric estimators, and Hansen (1992b) developed their specific form for linear regression models. Because (6) is linear, these techniques are directly applicable.

Following these articles, a good test for H_0 versus H_1 takes the form of rejecting H_0 for large values of

$$L = \frac{1}{n^2}\sum_{i=1}^{n} \widehat{F}_i' A (A'\widehat{\Omega}^f A)^{-1} A' \widehat{F}_i$$
$$= \frac{1}{n^2} \mathrm{tr}\left((A'\widehat{\Omega}^f A)^{-1} A' \sum_{i=1}^{n} \widehat{F}_i \widehat{F}_i' A \right), \tag{13}$$

where $\widehat{F}_i = \sum_{t=1}^{i} f_t \widehat{e}_t$, $\widehat{e}_t$ are the OLS residuals from (6), $\widehat{\Omega}^f$ is defined in (9), and $\mathrm{tr}(Q)$ stands for the trace of Q. When e_i are iid normal and the x_i are strictly exogenous, L is the LM test for H_0 against H_1. When these assumptions are relaxed, L can be interpreted as an "LM-like" test derived from the generalized least squares criterion function or as an asymptotic equivalent of the true LM test. In addition, when e_i is directly observed (rather than a residual), L is an asymptotic approximation to the locally most powerful test for H_0 versus H_1, suggesting that L should have good power for local departures from the null of no seasonal unit roots.

To be precise, (13) is the LM statistic for H_0 against H_1 under the assumption that $G = (A'\Omega^f A)^{-1}$ [recall that $E(u_i u_i') = \tau^2 G$, where u_i is the error in (12)]. This choice for G may seem arbitrary, but it is guided by the fact that only this choice produces an asymptotic distribution for L that is free of nuisance parameters and hence allows the tabulation of critical values for use in applications. This technique is not without precedent. Indeed, the same criterion is used to construct the standard Wald test. It may be helpful to digress briefly on this point. The general form of the Wald test for the hypothesis $\gamma = 0$ against $\gamma \neq 0$ is $W = \widehat{\gamma}' G \widehat{\gamma}$. The matrix G determines the direction of the hypothesis test. Indeed, for any G the power of the Wald test W against an alternative γ is determined by the noncentrality parameter $\gamma' G \gamma$, and thus the power is maximized against alternatives γ, which are proportional to the eigenvector of G corresponding to its largest eigenvalue. Hence G could (in principle) be selected to maximize power against particular directions of interest. This is never done in practice. Instead, we set $G = \widehat{V}^{-1}$, where $\widehat{V}$ is a consistent estimate of the asymptotic covariance matrix of $\widehat{\gamma}$. This is not because the eigenvector of $\widehat{V}^{-1}$ corresponding to its largest eigenvalue is a particularly interesting direction for the alternative γ but because it is the unique choice, which yields an asymptotic distribution for W free of nuisance parameters. The same reasoning applies to our L tests. Although better power against particular alternatives could in principle be obtained by selecting an appropriate matrix G, this would result in a test with unknown asymptotic size and would hence be useless in practice.

The large-sample distribution of L was studied by Nyblom (1989) and Hansen (1990, 1992b). To simplify the presentation, we introduce the following notation. Let $\rightarrow_d$ denote convergence in distribution, W_p denote a vector standard Brownian bridge of dimension p, and let $\mathrm{VM}(p)$ denote a random variable obtained by the transformation

$$\mathrm{VM}(p) = \int_0^1 W_p(r)' W_p(r) dr. \tag{14}$$

When $p = 1$, the distribution of $\mathrm{VM}(p)$ simplifies to that known as the *Von Mises goodness-of-fit distribution* widely used in the statistical literature (e.g., see Anderson and Darling 1952), so we will refer to $\mathrm{VM}(p)$ as the *generalized Von Mises distribution with p degrees of freedom.* Critical values are given in Table 1.

Theorem 1. Under H_0, $L \rightarrow_d \mathrm{VM}(a)$.

Table 1. Critical Values for VM(p)

	Significance level					
p	*1%*	*2.5%*	*5%*	*7.5%*	*10%*	*20%*
1	.748	.593	.470	.398	.353	.243
2	1.070	.898	.749	.670	.610	.469
3	1.350	1.160	1.010	.913	.846	.679
4	1.600	1.390	1.240	1.140	1.070	.883
5	1.880	1.630	1.470	1.360	1.280	1.080
6	2.120	1.890	1.680	1.580	1.490	1.280
7	2.350	2.100	1.900	1.780	1.690	1.460
8	2.590	2.330	2.110	1.990	1.890	1.660
9	2.820	2.550	2.320	2.190	2.100	1.850
10	3.050	2.760	2.540	2.400	2.290	2.030
11	3.270	2.990	2.750	2.600	2.490	2.220
12	3.510	3.180	2.960	2.810	2.690	2.410

Source: Hansen (1990), table 1.

Proof. The proof is omitted and available on request from the authors.

Theorem 1 shows that the large-sample distribution of the L statistic does not depend on any nuisance parameters other than a (the rank of A), which refers to the number of elements of γ that are being tested for constancy.

2.3 Joint Test for Unit Roots at All Seasonal Frequencies

If the alternative of interest is seasonal nonstationarity, then we should simultaneously test for unit roots at all seasonal frequencies. This can be accomplished by using Statistic (13) with $A = I_{s-1}$. This yields the statistic

$$L_f = \frac{1}{n^2} \sum_{i=1}^{n} \widehat{F}_i' (\widehat{\Omega}^f)^{-1} \widehat{F}_i = \frac{1}{n^2} \mathrm{tr}\left((\widehat{\Omega}^f)^{-1} \sum_{i=1}^{n} \widehat{F}_i \widehat{F}_i' \right). \tag{15}$$

The subscript f on L indicates that the test is for nonstationarity at *all* seasonal frequencies.

We would like to emphasize that, although the form of the statistic L_f is nonstandard, it is quite simple to calculate. It only requires estimation under the null hypothesis of stationary seasonality, and it is calculated directly from the OLS residuals and the trigonometric coefficients f_i.

The large-sample distribution of L_f follows directly from Theorem 1.

Theorem 2. Under H_0, $L_f \rightarrow_d \mathrm{VM}(s-1)$.

Theorem 2 indicates that the large-sample distribution of the test for unit roots at all seasonal frequencies is given by the generalized Von Mises distribution with $s - 1$ df. This result shows that not only is the test statistic L_f easy to calculate but that its large-sample distribution theory takes a simple form. For quarterly data, the appropriate critical values are found in Table 1 using the row corresponding to $p = s - 1 = 3$. For monthly data, the appropriate critical values are found using the row corresponding to $p = s - 1 = 11$.

2.4 Tests for Unit Roots at Specific Seasonal Frequencies

Writing (6) to emphasize the seasonal components at individual seasonal frequencies, we have

$$y_i = \mu + x_i'\beta + \sum_{j=1}^{q} f_{ji}'\gamma_j + e_i, \qquad i = 1, 2, \ldots, n. \tag{16}$$

Recall that the jth coefficient γ_j corresponds to the seasonal cycle for the frequency $(j/q)\pi$. Testing for a seasonal unit root at frequency $(j/q)\pi$ therefore reduces to testing for a unit root in γ_j. This corresponds to the hypothesis test H_0 versus H_1, where the A matrix has the block diagonal form $A = (\tilde{0}\ I_2\ \tilde{0})'$ for $j < q$ and $A = (\tilde{0}\ 1)'$ for $j = q$, where the 1s correspond to the subvector γ_j. Letting $\widehat{\Omega}_{jj}^f$ denote the jth block diagonal element of $\widehat{\Omega}^f$ (commensurate with γ), we find that the test statistic L reduces to

$$L_{(\pi j/q)} = \frac{1}{n^2} \sum_{i=1}^{n} \widehat{F}_{ji}' \left(\widehat{\Omega}_{jj}^f\right)^{-1} \widehat{F}_{ji}, \tag{17}$$

for $j = 1, \ldots, q$, where $\widehat{F}_{ji} = \sum_{t=1}^{i} f_{jt}\widehat{e}_t$ is the subvector of $\widehat{F}_i$ partitioned conformably with γ.

Again, we would like to emphasize the convenience of the statistics $L_{(\pi j/q)}$. They can be computed as by-products of the calculation of the joint test L_f because $L_{(\pi j/q)}$ only make use of subcomponents of the vector $\widehat{F}_i$ and of the matrix $\widehat{\Omega}^f$. Their asymptotic distributions are readily obtained:

Theorem 3. Under H_0,

1. For $j < q$, $\quad L_{\pi j/q} \rightarrow_d \text{VM}(2)$.
2. $L_\pi \rightarrow_d \text{VM}(1)$.

Theorem 3 states that the large-sample distributions of the tests for seasonal unit roots are given by the generalized Von Mises distribution with 2 df for frequencies different than π and with 1 df for frequency π. This stems from the dimensionality of the subvectors γ_j in the two cases. For quarterly data, the two seasonal frequencies are at $\pi/2$ (annual) and π (biannual).

The $L_{(\pi j/q)}$ tests are useful complements to the joint test L_f. If the joint test rejects, it could be due to unit roots at any of the seasonal frequencies. The $L_{(\pi j/q)}$ tests are specifically designed to detect at which particular seasonal frequency nonstationarity emerges.

3. TESTING FOR NONCONSTANT SEASONAL PATTERNS

3.1 The Testing Problem

The tests for seasonal unit roots proposed in Section 2 were derived from the trigonometric seasonal model (6). To study whether the seasonal intercepts α have changed over time, we return to the more conventional seasonal dummy model (3), which we modify as

$$y_i = x_i'\beta + d_i'\alpha_i + e_i, \qquad i = 1, 2, \ldots, n. \tag{18}$$

There are many forms of potential nonstationarity for α_i that could be considered. Here we consider stochastic variation of a martingale form:

$$A'\alpha_i = A'\alpha_{i-1} + u_i, \tag{19}$$

where α_0 is fixed and $\{u_i, \mathcal{F}_i\}$ is an MDS with covariance matrix $E(u_i u_i') = \tau^2 G$. The $s \times a$ matrix A selects the elements of α that we allow to stochastically vary under the alternative hypothesis. Note that when $\tau = 0$ the coefficient vector is fixed at α_0 for the entire sample.

This specification of coefficient variation is quite general. One special case is the Gaussian random walk, under which the seasonal intercepts α_i slowly (but continuously) evolve over time. Another interesting case discussed by Nyblom (1989) is when the martingale differences u_i come from an "infrequent innovation" process. Let $u_i = \delta_i v_i$, where v_i is iid $\mathcal{N}(0, I_a)$ and δ_i is a discrete random variable equaling 1 with probability ψ and 0 with probability $1 - \psi$ (and v_i and δ_i are independent). For ψ sufficiently small, the martingale α_i will be constant for most observations but will exhibit infrequent and unpredictable "structural breaks." If desired, (19) can be generalized to a random array α_{ni}, which can have exactly one "structural break" of unknown timing (in all or a subset of the seasonal intercepts α_{ni}) in a given sample.

As in Section 2.2, the LM test for $H_0 : \tau = 0$ against $H_1 : \tau \neq 0$ is given by the statistic [setting $G = (A'\widehat{\Omega}A)^{-1}$]

$$\begin{aligned} L &= \frac{1}{n^2} \sum_{t=1}^{n} \widehat{D}_t' A (A'\widehat{\Omega}A)^{-1} A' \widehat{D}_t \\ &= \frac{1}{n^2} \operatorname{tr}\left[(A'\widehat{\Omega}A)^{-1} A' \sum_{t=1}^{n} \widehat{D}_t \widehat{D}_t' A \right], \end{aligned} \tag{20}$$

where $\widehat{D}_t = \sum_{i=1}^{t} d_i \widehat{e}_i$, $\operatorname{tr}(Q)$ is the trace of Q, and $\widehat{\Omega}$ is defined in (8).

3.2 Testing for Instability in an Individual Season

Testing the stability of the ath seasonal intercept (where $1 \leq a \leq s$) can be achieved by choosing A to be the unit vector with a 1 in the ath element and zeros elsewhere. This produces the test statistic

$$L_a = \frac{1}{\widehat{\Omega}_{aa} n^2} \sum_{t=1}^{n} \widehat{D}_{at}^2, \tag{21}$$

where $\widehat{D}_{at}$ is the ath element of $\widehat{D}_t$, and $\widehat{\Omega}_{aa}$ is the ath diagonal element of $\widehat{\Omega}$.

Theorem 4. Under H_0, $L_a \rightarrow_d \text{VM}(1)$ for each $a = 1, \ldots, s$.

Theorem 4 shows that the asymptotic distribution of the test for instability in an individual seasonal intercept is given by the generalized Von Mises distribution with 1 df, for which critical values are given in the first row of Table 1.

To calculate these test statistics, note that, because the ath dummy variable is 0 for all but one out of every s observations, the cumulative sum $\widehat{D}_{at}$ is only a function of the residuals from the ath season. Thus the test statistic L_a can be calculated using only the residuals from the ath season. To see this, let $j = 1, \ldots, T_1$ denote the annual observations for the ath season, and let $\widetilde{e}_j, j = 1, \ldots, T_1$, denote the OLS residuals for this season. Then

$$L_a = \frac{1}{T_1^2 \widetilde{\sigma}_1^2} \sum_{j=1}^{T_1} \left(\sum_{t=1}^{j} \widetilde{e}_t \right)^2, \tag{22}$$

where $\widetilde{\sigma}_1^2 = \sum_{k=-m}^{m} w(k/m) 1/T_1 \sum_i \widetilde{e}_{i+k} \widetilde{e}_i$.

Hence the statistics L_a are essentially the KPSS statistic applied to the seasonal subseries (only the observations from the ath season are used). Thus, the KPSS test is for instability in the average level of the series, but the L_a tests are for instability in the seasonal subseries.

3.3 Joint Test for Instability in the Seasonal Intercepts

Just as we computed a joint test for unit roots at all the seasonal frequencies, we can construct joint tests for instability in all the seasonal intercepts. One straightforward test statistic can be obtained by taking (20) with $A = I$, yielding

$$L_J = \frac{1}{n^2} \sum_{t=1}^{n} \widehat{D}_t' \widehat{\Omega}^{-1} \widehat{D}_t. \tag{23}$$

Standard analysis shows that, under H_0, $L_J \to_d \mathrm{VM}(s)$. Note that L_J is a test for instability in any of the seasonal intercepts so that it will have power against zero-frequency movements in y_i. In other words L_J is a joint test for instability at the zero frequency as well as at the seasonal frequencies. This is an undesirable feature because rejections of H_0 could be a consequence of long-run instability at the zero frequency. This objection could also be raised with individual test statistics L_a, but the problem appears more acute with the joint test L_J.

To cope with this problem one could test for variation in the joint seasonal intercept process that keeps the overall mean constant. Specifically, decompose the seasonal intercepts α in (3) into an overall mean and deviations from the mean. We can write this as

$$\alpha = \iota_s \mu + H\eta, \tag{24}$$

where ι_s is an s vector of ones, $\mu = \iota_s'\alpha/s$ is the overall mean, η is the $(s-1) \times 1$ vector of deviations from μ for the first $s-1$ seasons (the deviation for the sth season is redundant), and H is the $s \times (s-1)$ matrix

$$H = \begin{pmatrix} I_{s-1} \\ -\iota_{s-1}' \end{pmatrix},$$

where ι_{s-1} is an $(s-1)$ vector of ones. This is simply a reparameterization of the model (3), which can now be written as

$$y_i = \mu + x_i'\beta + d_i^{*\prime}\eta + e_i, \tag{25}$$

where $d_i^* = H'd_i$ is a full-rank, $(s-1)$-dimensional, mean-zero deterministic seasonal process. We can test for stability of the seasonal intercepts, holding constant the overall mean μ, by testing for the stability of the coefficients η via the specification

$$y_i = \mu + x_i'\beta + d_i^{*\prime}\eta_i + e_i \tag{26}$$

and

$$\eta_i = \eta_{i-1} + u_i, \tag{27}$$

where $u_i \sim (0, \tau^2 G)$, using the methods outlined in the previous sections. Rejection of the null hypothesis implies that some seasonal intercepts have changed. Note that Models (26)–(27) allow for time variation in the seasonal pattern, but the seasonal intercepts are constrained to sum to the constant μ. As the associate editor has pointed out, this unusual specification embeds as special cases the periodic models of seasonality studied by Osborn and Smith (1989) and Hansen and Sargent (1993).

The fact that (26)–(27) are unusual is not crucial, however, once one considers the algebraic structure of the models and test statistics. Because f_i and d_i^* are linear combinations of one another, in the sense that, for some invertible matrix B, $d_i^* = B'f_i$, (6) can be written as

$$y_i = \mu + x_i'\beta + f_i'\gamma + e_i, \tag{28}$$

where $\gamma = B^{-1}\eta$. By uniqueness of the representation, this γ is the same as the coefficients in (6). By linearity, testing the joint stability of η in (25) using the alternative (26)–(27) is algebraically equivalent to testing the joint stability of γ in (28) against (10)–(11). It follows that the joint test for seasonal instability obtained from (26)–(27) is exactly L_f.

To put the finding in another way, we have found that either construction—testing for instability as viewed through the lens of seasonal intercepts or from the angle of seasonal unit roots—gives exactly the same joint test. There is no need to choose one approach or the other because both yield the same answer. Thus the appearance of the alternative model given by (26)–(27) as overly restrictive is an artifact of the analysis of the seasonal dummy model and not a substantive restriction.

4. A MONTE CARLO EXPERIMENT

To examine the performance of our proposed test statistics, we conducted a small Monte Carlo exercise. We consider two quarterly models, one roughly consistent with our specification and the other consistent with the setup of HEGY. The first model is

$$y_t = by_{t-1} + \sum_{j=1}^{2} f_{jt}'\gamma_{jt} + \epsilon_t, \qquad \epsilon_t \sim \mathcal{N}(0,1), \tag{29}$$

and

$$\gamma_t = \gamma_{t-1} + u_t, \qquad u_t \sim \mathcal{N}(0, \tau^2 G), \tag{30}$$

where $\gamma_0 = [0,0,0]$. The second model is

$$(1 - bL)(1 + g_2 L)\left(1 + g_3 L^2\right) \times [y_t - 10 + 4.0d_1 - 4.0d_2 + 6.0d_3] = \epsilon_t, \tag{31}$$

where $\epsilon_t \sim \mathcal{N}(0,1)$ and d's are seasonal dummies. For the first model [(29)–(30)] we use three data-generating processes (DGP) under the alternative:

$$\mathrm{DGP1}: G = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{32}$$

$$\mathrm{DGP2}: G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{33}$$

and

$$\text{DGP3}: G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (34)$$

The model implied by DGP1 is exactly that for which the test L_π is designed: When $\tau = 0$, there are no unit roots, but when $\tau \neq 0$, there is a unit root at frequency π. DGP2 is designed for the $L_{\pi/2}$ test, in that $\tau = 0$ implies no unit roots and $\tau \neq 0$ implies a pair of complex conjugate roots at frequency $\pi/2$. DGP3 has no unit roots when $\tau = 0$ but has unit roots at both seasonal frequencies when $\tau \neq 0$.

For the second model (31) we consider several specifications—a case in which seasonality is deterministic ($g_2 = g_3 = 0$), another in which there is a deterministic and a stochastic component to seasonality ($0 < g_2$ and $g_3 < 1$), a third in which there is a unit root at π ($g_2 = 1$ and $g_3 < 1$), a fourth in which there are a pair of unit roots at $\pi/2$ ($g_2 < 1$ and $g_3 = 1$), and finally one with unit roots at both π and $\pi/2$ ($g_2 = g_3 = 1$).

For both seasonal models we vary the first-order autoregressive parameter among $b = [.5, .95, 1.0]$, and the sample size among $T = [50, 150]$. For model (29)–(30) we select the strength of the seasonal unit-root component among $\tau = [.00, .10, .20]$. For model (31) we select the strength of the stochastic seasonal component among g_2, $g_3 = [.0, .5, .95, 1.0]$. The choices for the first-order autoregressive parameter and sample size were selected to correspond to typical macroeconomic time series (and our applications).

For each parameter configuration, we created 1,000 independent samples for each DGP and calculated the tests for unit roots at the seasonal frequencies π and $\pi/2$, the joint test at both frequencies, and the tests for instability in the four seasonal dummies. Because the alternative here is a seasonal unit root, the tests for instability in the four individual seasons had similar performances, so we only report the results for the first seasonal dummy, D1. We ran the tests on the level of simulated data, but we also experimented, for the case in which $b = 1$, running the test on the first difference of the simulated data. Furthermore, we follow the suggestion of Section 1.3 and include one lag of the dependent variable. Thus the model for OLS estimation is

$$y_t = \mu + \beta y_{t-1} + f_t'\gamma + e_t. \quad (35)$$

To implement the tests, we need to select estimates of the long-run covariance matrix $\widehat{\Omega}$, which reduces to the choice of kernel and lag truncation number m. In the simulations reported here, we use the Bartlett kernel and, following Andrews (1991, table 1), select $m = 3$ if $T = 50$ and $m = 5$ if $T = 150$. The effect of selecting other values is discussed later.

The performance of our LM tests is compared with two alternative testing methodologies. The first is a simple t test for the existence of stochastic seasonality, which is obtained from testing $\delta_2 = 0$ in the model

$$(1 - \delta_1\ell)\left(1 - \delta_2\ell^4\right)\left(y_t - \sum_{j=1}^{4}\alpha_j d_{jt}\right) = e_t. \quad (36)$$

Within our Monte Carlo design, the null hypothesis for this test is the same as for our LM tests (although this would not be true in more general models because our tests take the null to include *stationary* stochastic fluctuations) and thus provides a valid basis for comparison.

The second alternative testing procedure is that developed by HEGY. The approach is based on testing the nullity of the coefficients ρ_j in the auxiliary regression

$$A(\ell)w_{1t} = \rho_1 w_{2t-1} + \rho_2 w_{3t-1} + \rho_3 w_{4t-1} + \rho_4 w_{4t-2} + \sum_{j=1}^{4}\alpha_j d_{jt} + e_t, \quad (37)$$

where $w_{1t} = (1 - \ell^4)y_t$, $w_{2t} = (1 + \ell + \ell^2 + \ell^3)y_t$, $w_{3t} = -(1 - \ell + \ell^2 - \ell^3)y_t$, and $w_{4t} = -(1 - \ell^2)y_t$. We examine augmented Dickey–Fuller t statistics for the hypothesis $\rho_2 = 0$ (unit root at frequency π) and the HEGY F statistic for the hypothesis $\rho_3 = \rho_4 = 0$ (a pair of conjugate complex unit roots at frequency $\pi/2$). For each experiment we use six-lag augmentation; that is, $A(\ell) = 1 - a_1\ell - a_2\ell^2 - a_3\ell^3 - a_4\ell^4 - a_5\ell^5 - a_6\ell^6$. As an anonymous referee and an associate editor have pointed out, model (29)–(30) generates moving average (MA) components in y_t so that the lag augmentation should be sufficiently long for the HEGY test to be reasonably powerful. Our LM tests and the HEGY tests take the opposite null and alternative hypotheses and are thus not directly comparable.

Other Monte Carlo experiments to evaluate tests for unit roots at seasonal frequencies have been conducted by Hylleberg (1992) and Ghysels, Lee, and Noh (GLN) (1992). Hylleberg also contrasted the HEGY tests with our tests for structural stability in the dummies but used a simple AR process for the DGP of the data. GLN examined the relative performance of the HEGY and DHF tests.

The results of the experiments are contained in Tables 2–4 for Models (29)–(30) and Tables 5–6 for Model (31). Each table reports the percentage rate of rejection of the relevant null hypothesis at the asymptotic 5% significance level.

4.1 First Seasonal Model

4.1.1 Size of the Test. Table 2 reports the results for Models (29)–(30) when $\tau = 0$, which corresponds to the hypothesis of no seasonal unit roots.

All of our LM tests have good size, especially for $T = 150$. In nearly every case, the rejection frequency is close to or slightly above the nominal level of 5%. The size of the tests does not appear to be very sensitive to the magnitude of b.

The test for stochastic seasonality also takes the null $\tau = 0$; thus, the rejection frequencies in Table 2 are also the finite sample sizes of the test. The results are somewhat mixed. The test tends to underreject, regardless of the sample size, when $b < 1$, but it overrejects when $b = 1$.

The HEGY tests take the null of unit roots at the seasonal frequencies, so the parameter configurations of Table 2 lie in the alternative hypothesis for these tests. Thus, we should expect the statistics to reject frequently. Indeed, the tests for frequency π and $\pi/2$ reject in nearly every trial when $T = 150$, but the power of the tests is substantially reduced when $T = 50$.

Table 2. First Seasonal Model, Size and Power of Asymptotic 5% Tests, Monte Carlo Comparison, 1,000 Replications: $\tau = 0$

Sample		HEGY			Dummy	Seasonals		
size	b	π	$\pi/2$	Stationary	D1	π	$\pi/2$	Joint
150	.50	97.6	100.0	2.2	4.8	4.6	5.6	5.8
150	.95	98.8	100.0	3.4	6.6	6.4	4.0	5.2
150	1.00	98.4	100.0	7.4	3.6	6.4	5.4	5.6
50	.50	29.6	19.2	2.3	6.4	4.6	8.6	6.0
50	.95	38.8	26.4	2.2	5.4	4.8	5.6	4.6
50	1.00	29.6	37.2	8.4	4.2	7.6	7.4	8.2

NOTE: "HEGY" is the Hylleberg, Engle, Granger, and Yoo (1990) test for unit roots at seasonal frequencies, "Stationary" is a test for the presence of stochastic seasonality. "Dummy" is the test for the instability of the first seasonal dummy proposed in Section 3, and "Seasonals" are the tests for unit roots at seasonal frequencies proposed in Section 2. The HEGY test is run with a six-lag augmentation.

4.1.2 Power Under the Alternative. Tables 3 and 4 report the rejection frequency of the tests under the hypothesis of seasonal unit roots. In Table 3 we set $\tau = .1$, and in Table 4 $\tau = .2$.

Our proposed LM tests perform remarkably well. First, examine the test for nonconstancy in the first seasonal dummy (D1). Because all three alternative models induce seasonal unit roots into the model, this will appear as an unstable seasonal intercept, and we should expect this test to reject the null of stationarity. Indeed, for $T = 150$, the statistic rejects in 28%–44% of the trials when $\tau = .1$, and in 54%–72% of the trials when $\tau = .2$. As expected, the power is less for $T = 50$.

Second, examine the test for a unit root at frequency π. For illustration, take $T = 150$ and $\tau = .1$. Under DGP1 (a unit root at the frequency π), the test rejects in 73%–77% of the trials. When both seasonal unit roots are present (DGP3), the test rejects in only 43%–45% of the trials, indicating an adverse effect of the presence of a contaminating unit root. A remarkable result is that under DGP2, when there is a unit root at the frequency $\pi/2$ (the wrong seasonal frequency), the test rejects in only 3%–4% of the trials. This is good news, for it implies that the statistic has no tendency to "spuriously reject" due to the presence of another seasonal unit root.

The results are similar for the test at frequency $\pi/2$. Again taking $T = 150$ and $\tau = .1$, we note that in the presence of a unit root at $\pi/2$ (DGP2) the rejection frequency is 64%–65%, and in the presence of both seasonal unit roots (DGP3) the rejection frequency is 41%–46%. In the presence of the wrong unit root (DGP1), the rejection rate is an excellent 2%–3%.

The joint test also performed quite well, having power against a unit root at frequency π or $\pi/2$ close to that found by the L_π or $L_{\pi/2}$ test. When unit roots are present at both frequencies, then typically the joint test has greater power than either individual test.

Table 3. First Seasonal Model, Size and Power of Asymptotic 5% Tests, Monte Carlo Comparison, 1,000 Replications: $\tau = .1$

Sample			HEGY			Dummy	Seasonals		
size	b	DGP	π	$\pi/2$	Stationary	D1	π	$\pi/2$	Joint
150	.5	DGP1	40.0	100.0	38.6	39.8	73.4	3.6	61.4
		DGP2	98.8	87.2	25.2	30.8	3.4	64.8	59.6
		DGP3	77.2	96.2	22.4	28.2	43.2	44.8	65.2
150	.95	DGP1	42.4	100.0	49.6	42.4	77.8	2.0	70.4
		DGP2	98.6	93.6	29.8	28.6	4.4	64.6	61.8
		DGP3	77.8	99.0	29.4	31.0	45.6	46.2	65.8
150	1.00	DGP1	42.0	100.0	21.8	44.8	74.2	3.2	67.2
		DGP2	97.6	95.2	25.0	28.2	3.0	65.4	61.6
		DGP3	75.8	98.8	20.0	31.8	45.2	41.8	64.0
50	.50	DGP1	19.4	18.2	5.6	11.4	30.0	6.8	21.4
		DGP2	24.4	15.4	3.0	7.4	4.0	18.8	14.2
		DGP3	24.0	15.6	4.2	6.4	11.6	3.0	14.6
50	.95	DGP1	21.8	31.0	8.8	13.2	30.0	4.4	20.6
		DGP2	34.0	24.4	5.4	7.2	6.6	14.6	13.6
		DGP3	31.6	26.2	4.6	9.2	10.8	10.8	13.0
50	1.00	DGP1	18.4	34.2	12.4	9.4	31.4	5.6	21.8
		DGP2	28.8	28.8	9.8	5.0	5.6	18.4	13.8
		DGP3	26.2	28.4	10.6	4.8	14.6	10.4	15.6

NOTE: See note to Table 2.

Table 4. First Seasonal Model, Size and Power of Asymptotic 5% Tests, Monte Carlo Comparison, 1,000 Replications: $\tau = .2$

Sample size	b	DGP	HEGY π	HEGY $\pi/2$	Stationary	Dummy D1	Seasonals π	Seasonals $\pi/2$	Seasonals Joint
150	.5	DGP1	13.0	100.0	78.6	59.6	90.0	.6	82.4
		DGP2	98.4	39.0	77.0	57.8	.2	91.8	90.0
		DGP3	40.4	65.2	77.8	54.8	67.8	73.2	92.0
150	.95	DGP1	14.2	100.0	89.0	69.6	93.0	1.0	88.0
		DGP2	96.2	49.0	81.4	58.4	1.6	92.6	91.6
		DGP3	38.0	71.6	86.4	58.2	69.6	77.8	94.0
150	1.00	DGP1	15.2	100.0	39.6	72.4	94.4	1.0	87.8
		DGP2	96.0	55.4	64.8	54.2	1.6	91.2	88.4
		DGP3	38.6	79.2	56.4	57.0	75.2	73.8	95.2
50	.50	DGP1	12.2	21.4	18.6	21.0	55.8	2.4	40.4
		DGP2	37.0	10.0	10.0	17.0	2.8	47.2	38.6
		DGP3	20.6	13.0	10.0	17.4	27.2	24.8	36.6
50	.95	DGP1	14.6	28.0	27.0	26.8	60.0	2.8	46.6
		DGP2	33.8	12.2	15.4	16.8	2.8	40.6	34.2
		DGP3	24.8	16.6	13.8	19.0	31.0	23.8	39.0
50	1.00	DGP1	11.8	39.0	19.0	21.4	59.8	2.2	47.8
		DGP2	28.8	17.4	15.2	16.0	2.2	45.4	37.6
		DGP3	15.6	23.2	14.2	15.2	27.0	26.0	36.6

NOTE: See note to Table 2.

Summarizing, the power performance of the LM tests, with the first seasonal design, is essentially independent of the parameter b and increases with τ and T, and the tests reject against the correct alternatives.

The rejection frequency of the test for stationary stochastic seasonality is similar to that of the LM test for a nonstationary dummy variable, and both are lower than the power of the LM test for seasonal unit roots when $\tau = .1$. The power dramatically improves when $\tau = .2$. It also drops when $b = 1$ regardless of the size of the other parameters. Because the test misbehaved when $b = 1$ under the null hypothesis as well, however, this result may be an artifact of size distortion.

The HEGY tests have trouble dealing with this design. First consider the test for a unit root at frequency π. Since $\tau > 0$, DGP1 and DGP3 lie in the test's null hypothesis (because there is a unit root at π), while DGP2 lies in the alternative hypothesis. Yet for $T = 150$, the rejection rate against DGP1 when $\tau = .1$ is 40%–42%, and when $\tau = .2$ it is reduced to 13%–15%. Against DGP3, the rejection rate when $\tau = .1$ is 75%–77% and is 38%–40% when $\tau = .2$. These are rejections under the null, and hence this indicates massive size distortion.

Second, consider the test for a pair of complex conjugate unit roots at the frequency $\pi/2$, for which DGP2 and DGP3 are included under the null hypothesis and DGP1 is the alternative. Under the local alternative $\tau = .1$, the test rejects in nearly every trial under the null (87%–100% when $T = 150$), indicating that the statistic cannot discriminate between the null and the alternative. In this case also, this massive distortion diminishes when we increase τ to .2, even though it is still very sizable (39%–79% rejection rates when $T = 150$). Apparently, HEGY tests find it hard to deal with designs in which unit roots appear as large masses as opposed to sharp peaks in the spectrum.

4.1.3 Some Robustness Experiments. The simulation results for the LM test reported previously used a consistent kernel-based estimate of the long-run covariance matrix. It is fairly straightforward to see that, under the alternative hypothesis of a seasonal unit root, the value of the L statistic will be *decreasing* (at least in large samples) as a function of m. Thus, selecting too large a value will have adverse effects on power. Selecting a too-small value, however, risks size distortion if there is unaccounted-for serial correlation in the errors. This is the same problem as arises in the LM test for a unit root at the zero frequency and was discussed by Kwiatkowski, Phillips, Schmidt, and Shin (1992).

To investigate the robustness of the results to alternative estimates of the covariance matrix, we have experimented with two other values for m. As an extreme choice, we set $m = \sqrt{n}$ [as done, for example, by Hylleberg (1992)]. This decreased the power of our LM tests by approximately 50%. We consider this an upper bound on the power loss. At the opposite extreme we used a naive OLS estimator, which can be viewed as setting the lag window $m = 0$ in the general expression, and it is optimal within our design. For this last choice, we found the gain of power of our test to be approximately 5%–10%. We conclude that a conservative choice of m along the lines of Andrews (1991) is important to retain significant finite-sample power.

We also experimented with the larger sample size $T = 300$ and for the case in which no lagged y_t was included in (35). As expected, all of the tests performed much better in terms of both size and power when T was larger, but no significant size or power distortion occurred when no lagged dependent variable was included in the estimated model.

We also examined the size and the power of our tests when a preliminary first-order differencing transformation was used

Table 5. Second Seasonal Model, Size and Power of Asymptotic 5% Tests, Monte Carlo Comparison, 1,000 Replications

Sample size	b	g2	g3	HEGY π	HEGY $\pi/2$	Stationary	Dummy D1	Seasonals π	Seasonals $\pi/2$	Seasonals Joint
150	.5	.0	.0	98.6	100.0	1.9	5.2	3.6	4.9	3.8
150	.95	.0	.0	98.6	100.0	1.5	3.5	3.8	3.6	3.5
150	1.0	.0	.0	97.7	99.7	1.8	5.7	3.1	3.8	3.8
150	.5	.5	.5	95.8	95.9	64.4	8.7	4.0	15.6	12.6
150	.95	.5	.5	96.9	98.6	96.1	17.6	5.0	15.8	12.9
150	1.00	.5	.5	97.7	98.4	98.2	20.2	6.2	16.8	15.2
150	.5	.95	.95	16.7	10.0	100.0	68.6	42.5	94.2	96.9
150	.95	.95	.95	17.1	7.7	100.0	65.3	51.1	92.5	96.6
150	1.00	.95	.95	17.8	9.7	100.0	68.3	55.9	90.8	95.4
50	.5	.0	.0	32.6	20.6	1.5	5.5	4.4	4.4	4.0
50	.95	.0	.0	35.0	25.0	1.8	6.4	4.2	5.6	4.7
50	1.0	.0	.0	36.3	28.8	.8	4.3	4.0	3.8	3.3
50	.5	.5	.5	23.5	9.0	14.9	12.7	4.0	21.1	16.9
50	.95	.5	.5	33.0	12.2	36.7	16.0	4.1	17.1	15.6
50	1.00	.5	.5	28.7	13.3	42.7	17.5	5.4	17.7	15.9
50	.5	.95	.95	11.1	4.6	98.4	55.1	48.0	83.4	90.5
50	.95	.95	.95	11.4	3.6	99.4	55.2	51.0	82.3	89.5
50	1.00	.95	.95	10.9	4.2	99.6	57.0	52.2	81.4	90.0

NOTE: See note to Table 2.

on the simulated data when $b = 1$. None of the results are changed by this modification. We believe that the inclusion of one lag of the dependent variable in the regressions effectively soaks up the unit root at the zero frequency without the need for any preliminary transformation.

4.2 Second Seasonal Model

4.2.1 Size of the Tests. Table 5 reports the results for Model (31) when $0 \leq g_2, g_3 < 1$, which corresponds to the null hypothesis of deterministic or deterministic-plus-stochastic stationary seasonality. In this case we should expect our tests to reject in 5% of the trials and the HEGY tests to reject often (because the design lies in the alternative).

The performance of all tests is reasonable when only deterministic seasonality is present ($g_2 = g_3 = 0$). When there are also stationary stochastic seasonal components (e.g., $g_2 = g_3 = .5$) our seasonal tests overreject in many cases, in particular the test for a unit root at $\pi/2$. As g_2 and g_3 approach unity (so that stochastic seasonality approaches the nonstationary region), both the seasonal unit root and the dummy tests exhibit size distortion and cannot distinguish a unit-root from a non-unit-root process. Again the size of the test is independent of T and, to a certain extent, b.

The performance of the stationary test is very similar to the one for dummies, even though the distortions become very large as b approaches 1.

The HEGY tests are reasonable regardless of b when g_2 and g_3 are small. As g_2 and g_3 increase, although remaining less than 1, the power of the tests decreases substantially. Reducing the sample size from 150 to 50 greatly affects the performance of the HEGY test but has practically no influence on our LM tests or on the stochastic seasonality test.

4.2.2 Power Under the Alternative. Table 6 reports the rejection frequency of the tests under the hypothesis of at least one seasonal unit root. (There is a unit root at frequency π when $g_2 = 1$ and a pair of conjugate unit roots at $\pi/2$ when $g_3 = 1$.) Therefore we should expect our tests and the stochastic stationarity test to reject frequently and the HEGY tests to reject in about 5% of the trials.

With this simple AR DGP, the performance of our dummy test is in general good, rejecting in about 80% of the trials when $T = 150$. The power is independent of b and lower when $T = 50$.

Our LM test performs reasonably well with this DGP even though the tests at π reject less frequently than expected when a unit root appears. Moreover, when g_2 and g_3 are close to unity, the test overrejects the null of stationarity. This is particularly evident when $g_2 = 1$ and $g_3 = .95$. The performance of the test is slightly worse when the sample size is small and is independent of b.

The power of the stochastic stationarity test is good with this design, and the test rejects whenever unit roots are present regardless of the size of the parameters b and T.

Finally, the HEGY tests perform well when $T = 150$, even though some power losses appear when there are near unit roots at the other seasonal frequency. As previously noted, the performance of the HEGY tests strongly depend on the sample size, but it is practically independent of b.

Overall, these results suggest that neither our new tests nor the HEGY tests are "superior" in either Monte Carlo design. For Models (29)–(30) our new LM tests perform as expected, just as the HEGY tests work somewhat well with the DGP (31). At this stage we see the two testing procedures as complementary to each other, and making meaningful comparisons between these testing frameworks will be

Table 6. Second Seasonal Model, Size and Power of Asymptotic 5% Tests, Monte Carlo Comparison, 1,000 Replications

Sample size	b	g2	g3	HEGY π	HEGY $\pi/2$	Stationary	Dummy D1	Seasonals π	Seasonals $\pi/2$	Seasonals Joint
150	.5	.5	1.0	97.3	3.9	100.0	84.3	.0	98.4	96.3
150	.5	1.0	.5	3.8	97.6	98.1	82.9	87.9	2.1	78.3
150	.95	.5	1.0	98.6	4.6	100.0	83.0	3.0	98.8	96.7
150	.95	1.0	.5	5.1	98.8	100.0	78.1	87.8	1.0	74.5
150	1.00	.5	1.0	98.9	4.8	100.0	85.0	.0	98.7	97.0
150	1.00	1.0	.5	6.1	98.9	100.0	83.4	87.2	1.0	72.5
150	.5	.95	1.0	17.3	4.8	100.0	84.5	41.4	98.5	99.1
150	.5	1.0	.95	6.0	9.9	100.0	76.0	83.1	89.8	98.5
150	.95	.95	1.00	18.1	4.0	100.0	81.1	43.7	97.8	99.2
150	.95	1.00	.95	5.4	7.2	100.0	75.4	81.3	89.0	98.7
150	1.00	.95	1.00	17.9	4.0	100.0	80.3	49.1	98.0	99.2
150	1.00	1.00	.95	6.5	10.3	100.0	81.2	84.7	85.8	97.9
150	.5	1.00	1.00	4.4	4.6	100.0	85.5	79.2	98.2	99.7
150	.95	1.00	1.00	6.7	4.6	100.0	79.2	78.6	97.7	99.8
150	1.00	1.00	1.00	5.0	4.1	100.0	82.8	81.8	96.5	99.7
50	.5	.5	1.0	32.6	3.4	98.4	65.4	3.2	40.6	86.6
50	.5	1.0	.5	8.8	12.7	94.6	56.9	67.5	15.4	64.5
50	.95	.5	1.0	36.0	2.8	99.2	62.5	2.7	89.9	86.5
50	.95	1.0	.5	8.0	13.7	91.1	54.0	66.3	11.7	62.1
50	1.00	.5	1.0	33.3	3.9	99.0	62.1	3.8	88.6	85.2
50	1.00	1.0	.5	8.9	13.3	94.9	61.2	68.3	11.7	64.6
50	.5	.95	1.0	12.1	3.4	99.3	63.9	48.4	90.0	93.9
50	.5	1.00	.95	8.7	4.7	99.3	60.9	67.8	82.7	94.1
50	.95	.95	1.0	11.8	3.0	99.7	61.8	50.1	88.7	94.3
50	.95	1.00	.95	9.3	3.5	99.7	63.0	67.8	81.3	93.1
50	1.00	.95	1.0	11.2	4.5	99.7	63.8	51.2	89.4	94.0
50	1.00	1.00	.95	9.6	4.2	99.6	64.7	63.4	81.2	92.7
50	.5	1.00	1.00	8.8	4.0	99.0	64.8	63.4	90.2	95.5
50	.95	1.00	1.00	4.2	3.5	99.6	63.9	66.1	86.8	95.7
50	1.00	1.00	1.00	9.0	3.0	99.8	64.0	65.3	86.4	95.4

NOTE: See note to Table 2.

an interesting subject for future research (see also Hylleberg 1992; Hylleberg and Pagan 1994).

5. APPLICATIONS

5.1 U.S. Post World War II Macroeconomic Series

The first data set we examine is that originally examined by Barsky and Miron (1989) in their study of the relationship between seasonal and cyclical fluctuations. The data set includes 25 variables that cover practically all of the major nonseasonally adjusted U.S. macroeconomic variables (total fixed investment, fixed residential investments, fixed nonresidential investments, fixed nonresidential structures, fixed nonresidential producer durables, total consumption, consumption of durables, consumption of nondurables, consumption of services, federal government expenditure, imports and exports, final business sales, changes in business inventories, Consumer Price Index (CPI), one-month treasury bill (T-bill) rates, M1, unemployment, labor force, employment, monetary base, money multiplier, and hours and wage rates). The original sources are described in the appendix of Barsky and Miron. The sample covers data from 1946,1 to 1985,4 except for M1 (starting date 1947,1), for unemployment and labor force (starting date 1948,1), employment (starting date 1951,1), the monetary base and the money multiplier (starting date 1959,1), and hours and wage (starting date 1964,1).

In constructing an estimate of the covariance matrix $\widehat{\Omega}$, we use the Newey and West (1987) procedure using Bartlett windows with lag truncation $m = 5$ for all series but hours and wage, for which we choose $m = 4$. For all variables we run the tests on the log differences to maintain comparability with previous analyses, and one lag of the dependent variable is included among the regressors. Table 7 reports significant dummies, the value of the L_i statistic for testing the stability of each separate dummy coefficient ($i = 1, 2, 3, 4$), the values of the L_π and $L_{\pi/2}$ statistics for nonstationarity at the seasonal frequencies, and the joint test statistic L_f. For four variables that display unstable seasonal patterns (fixed investment, consumption, government expenditure, and unemployment rate), Figure 1 plots recursive least squares estimates of the dummy coefficients in the spirit of Franses (1994). Under the assumption of unchanged seasonal patterns, the plot should depict four almost parallel lines. If lines intersect (e.g., spring becomes summer) unit-root behavior at seasonal frequencies is likely to occur. If changes in seasonal patterns changed primarily in the intensity of the fluctuations, the lines should tend to converge or diverge.

Table 7. Test for Structural Stability in the Seasonal Pattern of U.S. Macroeconomic Variables. Sample 46,1–85,4

Series	Quarter 1	Quarter 2	Quarter 3	Quarter 4	π	$\pi/2$	Joint
Investments	x .78*	x .29	.89*	.46	.10	2.16*	2.29*
Residential investments	x .66*	x .09	.93*	x .04	.09	1.35*	1.52*
Nonresidential investments	x .33	x .28	x .26	x .35	.17	1.13*	1.37*
Nonresidential structures	x .43	x .38	x .43	x .55*	.99*	.38	1.36*
Nonresidential producer durables	.41	x .47*	x .19	x .30	.33	.94*	1.38*
Consumption	x 2.16*	x 1.00*	x .63*	x 1.66*	2.11*	2.73*	4.98*
Consumption durables	x .22	x .12	.36	x .31	.39	.57	.96
Consumption nondurables	x 1.26*	x 1.60*	x 1.10*	x 1.74*	1.70*	2.22*	3.92*
Consumption services	x 1.36*	x .70*	x 1.21*	x .98*	1.49*	1.24*	2.74*
Gross national product	x 1.05*	x 1.14*	x .57*	x .88*	.46	2.88*	3.44*
Government expenditure	x 1.15*	x .70*	x .07	x .67*	.99*	.80*	1.79*
Imports	.15	x .44	x .08	.29	.30	.54	.85
Exports	.19	x .17	x .38	x .22	.25	.33	.60
Final sales	x 2.03*	x .26	x .17	x 1.47*	2.09*	2.97*	5.01*
CPI	.39	x .67*	x .29	.29	.26	.61	.87
T-bill rate	.35	.20	.09	.07	.32	.34	.66
Business inventories	x .24	1.53*	.34	.60*	.47*	1.08*	1.56*
M1	x .07	2.10*	x .13	x .18	1.79*	1.77*	3.67*
Unemployment rate	x 1.11*	x 1.23*	.44	x .12	.88*	2.20*	3.10*
Labor force	x 1.33*	x .31	x .09	x .45	1.05*	1.21*	2.29*
Employment	x .36	x .31	x .09	.57*	.32	.84*	1.16*
Monetary base	x .58*	x .33	x .20	x .23	.12	.93*	1.08*
Monetary multiplier	.57*	.81*	.30	x .72*	.31	1.59*	1.98*
Hours	x .07	x .81*	x .30	x .72*	.08	1.35*	1.44*
Wage	x .29	x .43	x .16	x .36	.83*	.30	1.14*

NOTE: An "x" indicates a dummy that is significant at the 5% level. The numbers reported in the columns for Quarters 1–4 are the values of the L statistics for each quarter. The next two columns report the values of the L_π and $L_{\pi/2}$ statistics, and the last column reports the joint test for instability at both frequencies. An asterisk indicates significance at the 5% level.

The results indicate that 24 out of the 25 variables display statistically significant seasonal patterns (the one-month T-bill rate is the only exception) and that for 20 of these the seasonal pattern has changed over the sample (according to the joint L_f test). The four variables that possess stable seasonal patterns are consumption of durables, imports, exports, and CPI. We also find that changes occur almost equally in all seasons but the third, that for 18 variables the null of constant seasonality is rejected at the annual frequencies, and that at the biannual frequency the test rejects the null in 11 cases. These results indicate that the comparison of deterministic seasonal and stochastic cyclical patterns as done by Barsky and Miron (1989) may not be appropriate because there are important time variations neglected in the analysis. They may also be viewed as consistent with results recently obtained by Ghysels (1991) that show that the seasonal pattern displayed by this set of macroeconomic variables tends to change with business-cycle conditions.

It is encouraging to observe that the individual dummy stability tests give similar conclusions to "eyeball" tests on the recursive estimates displayed in Figure 1. The first-quarter fixed-investment dummy trends toward 0, and the test rejects its constancy. The test also rejects the constancy of all government-expenditure dummies, except for the third-quarter dummy, a result that conforms with the plot of the recursive estimate. For the consumption series, the first and the fourth quarter dummies are the largest in absolute value, they trend toward 0 over time, and the test rejects their stability at the 1% level. A similar picture arises for the unemployment rate, except that it is the first- and the second-quarter dummies that are "large" in absolute value. In general, for all four variables considered in Figure 1 there is a tendency for the overall mean to be constant, for seasonals to become milder,

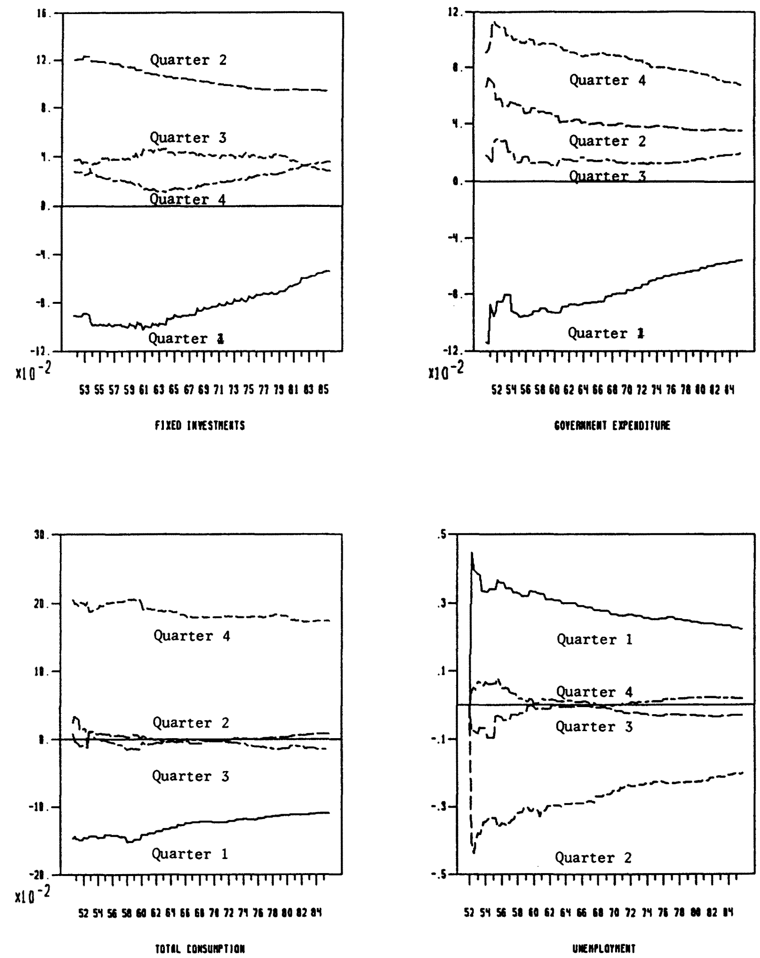

Figure 1. Recursive Estimates of Seasonal Dummies.

and for the intensity of the fluctuations to be reduced with some dummy coefficients turning insignificant in the last two decades. In addition, for the consumption and unemployment series, the coefficients of the dummies of two quarters change sign throughout the sample even though their value is always close to 0. Despite these large changes, none of the variables examined display a significant change in the location of seasonal peaks and troughs over time. Because these patterns are very typical of those we found among all the variables in the sample, one conclusion that emerges is that the intensity of seasonal fluctuations has substantially subsided in the past two decades, but no major seasonal inversion has occurred.

5.2 European Industrial Production

The second data set includes quarterly nonseasonally adjusted industrial production (IP) indexes for eight European countries (the United Kingdom, Germany, France, Italy, Spain, Austria, Belgium, and the Netherlands) for the sample 1960,3 to 1989,2. Canova (1993) described the original sources of the data. In this case, we also selected $m = 5$

Table 8. Test for Structural Stability in the Seasonal Pattern of Quarterly Industrial Production Indexes, Sample 60,1–89,2

Series	Quarter 1	Quarter 2	Quarter 3	Quarter 4	π	$\pi/2$	Joint
France	.13	x .57*	x .25	x .43	.57*	.39	.99
Germany	x .33	x 1.15*	.35	x .14	.47*	1.59*	2.06*
U.K.	x .19	1.16*	x .76*	x .12	1.09*	1.11*	2.20*
Italy	x .39	x .40	x 1.42*	x .78*	.85*	2.42*	3.20*
Austria	x .18	x .46	x .69*	x .15	.10	1.30*	1.45*
Belgium	1.19*	x .53*	x .81*	x .24	.36	1.59*	2.01*
Netherlands	.64*	x 1.15*	1.58*	x .92*	.21	2.72*	2.89*
Spain	.22	x .45	x 1.15*	x .56*	.87*	1.16*	2.11*

NOTE: An "x" indicates a dummy that is significant at the 5% level. The numbers reported in columns for Quarters 1–4 are the values of the L statistics for each quarter. The next two columns report the value of the L statistics at each of the two seasonal frequencies. The last column reports the joint test for instability at both seasonal frequencies. An asterisk indicates significance at the 5% level.

Table 9. Test for Structural Stability in the Seasonal Pattern of Monthly Stock Returns: Sample 50,1–89,9

	U.S.	*Japan*	*Germany*	*France*	*U.K.*	*Italy*	*Canada*
Significant dummies	1	1, 2, 11	7	1, 4, 7	7	1, 8	1
January	.07	.57*	.34	.05	.44	.61*	.08
February	.06	.39	.12	.40	.33	.10	.15
March	.14	.14	.18	.05	.28	.13	.19
April	.32	.11	.16	.28	.21	.14	.14
May	.33	.08	.25	.07	.24	.11	.16
June	.11	.30	.15	.11	.10	.12	.11
July	.06	.26	.22	.17	.08	.05	.05
August	.05	.12	.12	.09	.41	.26	.26
September	.30	.18	.28	.06	.11	.07	.06
October	.11	.52*	.12	.15	.12	.16	.11
November	.09	.33	.22	.07	.42	.15	.10
December	.17	.19	.14	.12	.41	.22	.08
Joint	1.74	3.02*	2.17	1.43	2.88*	1.84	1.25
$\pi/6$	.56	1.78*	.65	.70	1.44*	.70	.31

NOTE: The first 12 rows after the space report the values of the L statistic for each month. The next row reports the value of the L statistic for the joint test of instability at all seasonal frequencies. The last row reports the test for seasonal instability at frequency $\pi/6$. An asterisk indicates significance at the 5% level.

and estimated the model in log differences with a lag of the dependent variable. The results of testing for seasonal instability appear in Table 8.

The joint test indicates that all series, except possibly France, clearly display statistically significant changes in their seasonal patterns. The evidence of instability is stronger at the annual frequency, where the tests reject the null of no instability for all variables but the French IP index. At the biannual frequency, the test rejects the null for the IP index of the five largest countries. When we examine the stability of individual dummy coefficients, we find that, over the cross-section, all quarters appear to be subject to structural change but that the highest concentration of rejections of the null hypothesis of constancy emerges in the third quarter. This does not come as a surprise because the third quarter has been traditionally the vacation time in European countries, and in the last decade rescheduling programs have reduced the closing time of factories and offices to 10–14 days only, down from the 21 days which was the average in the 70s. Finally, the estimated coefficients of the dummies over three different decades and the recursive least squares plots (not presented for reasons of space) indicate changes in intensity, pattern and, in some cases, location of seasonal peaks and troughs over time.

5.3 Monthly Stock Returns

The third data we examine are a set of monthly stock returns on value-weighted indexes for seven industrialized countries (the United States, Japan, Germany, France, the United Kingdom, Italy, and Canada). This data set was obtained from the Citibase Tape and covers the period 1950,3 to 1989,9. As with the previous data sets, we set $m = 5$ and add one lagged dependent variable in the regression. The results of testing for the instability of the seasonal patterns in these variables are presented in Table 9.

All stock returns display some form of seasonality. The most significant seasonal dummies are for January returns (except for Germany and the United Kingdom). July and August returns have significant coefficients in four European countries. When we test for the structural stability of individual dummy coefficients, we find that significant time variations have emerged only for returns on a value-weighted index in Japan, the United Kingdom, and Italy. The joint test only rejects for Japan and the United Kingdom, where the rejection is due to a unit root at the annual frequency. It appears, therefore, that knowledge of predictable returns in four of the seven countries did not result in changes in these patterns, possibly indicating an inefficient propagation of information across these markets.

6. CONCLUSIONS

This article proposes a set of tests to examine the structural stability of seasonal patterns over time. The tests are built on the null hypothesis of unchanged seasonality and can be tailored to test for unit roots at seasonal frequencies or for time variation in seasonal dummy variables. We derive the asymptotic distribution of the statistics under general conditions that accommodate weakly dependent processes. A small Monte Carlo exercise demonstrates that the asymptotic distribution is a good approximation to the finite-sample distribution, and the test has good power against reasonable alternatives.

We apply the test to the three data sets. We find that in most cases deterministic dummies poorly capture the essence of seasonal variation in U.S. macroeconomic variables and that significant time variations are present in the seasonal patterns of the IP indexes of eight major industrialized countries and in the stock return indexes of some G-7 countries. The presence of seasonal time variations in quarterly U.S. macroeconomic variables partially invalidates some of the conclusions obtained by Barsky and Miron (1989), confirms recent findings of Ghysels (1991), and suggests the need for a more thorough and comprehensive examination of the statistical properties of macroeconomic variables.

The extension of our testing procedures to a vector of time series is straightforward. In that framework one can examine, for example, whether at least one of the seasonal intercepts of the system has changed. This extension would be analogous to that which Choi and Ahn (1993) made to the KPSS tests.

ACKNOWLEDGMENTS

Canova thanks the European Community for an EUI research grant, and Hansen thanks the National Science Foundation for a research grant. Part of this research was undertaken while Canova was associated with the European University Institute, Florence. The comments and the suggestions of John Geweke, Eric Ghysels, Sven Hylleberg, Adrian Pagan, two anonymous referees, the participants at seminars at the European University Institute, Institute of Advanced Studies, and the 1992 Winter Meetings of the Econometric Society are gratefully acknowledged. We are particularly grateful to an associate editor for extensive comments, in particular for his suggestion to use the trigonometric seasonal models of Hannan and Harvey to simplify our original presentation.

[Received April 1991. Revised November 1994.]

REFERENCES

Anderson, T. W., and Darling, D. A. (1952), "Asymptotic Theory of Certain 'Goodness of Fit' Criteria on Stochastic Processes," *The Annals of Mathematical Statistics*, 23, 193–212.

Andrews, D. (1991), "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817–858.

Barsky, R., and Miron, J. (1989), "The Seasonal and the Business Cycle," *Journal of Political Economy*, 97, 503–534.

Bell, W. R. (1993), "Empirical Comparison of Seasonal ARIMA and ARIMA Components (Structural) Time Series Models," unpublished paper presented at the 1993 Joint Statistical Meetings, San Francisco, August 8–12.

Braun, T., and Evans, C. (1990), "Seasonality and Equilibrium Business Cycle Theories," discussion paper 217, University of Virginia, Dept. of Economics.

Box, G., and Jenkins, G. (1976), *Time Series Analysis: Forecasting and Control*, San Francisco: Holden-Day.

Burns, A. F., and Mitchell, W. C. (1946), *Measuring Business Cycles*, New York: Columbia University Press (for the National Bureau of Economic Research).

Canova, F. (1992), "An Alternative Approach to Modeling and Forecasting Seasonal Time Series," *Journal of Business & Economic Statistics*, 10, 97–108

——— (1993), "Forecasting Time Series With Common Seasonal Patterns," *Journal of Econometrics*, 55, 173–200.

Chattarjee, S., and Ravikumar, B. (1992), "A Neoclassical Growth Model With Seasonal Fluctuations," *Journal of Monetary Economics*, 29, 59–86.

Choi, I., and Ahn, B. C. (1993), "Testing the Null of Stationarity for Multiple Time Series," unpublished manuscript, Ohio State University, Dept. of Economics.

Dickey, D., and Fuller, W. (1979), "Distribution of the Estimators for Autoregressive Time Series With a Unit Root," *Journal of the American Statistical Association*, 79, 355–367.

Dickey, D., Hasza, D., and Fuller, W. (1984), "Testing Seasonal Unit Roots in Seasonal Time Series," *Journal of the American Statistical Association*, 79, 355–367.

Engle, R., Granger, C., and Hallman, J. (1989), "Merging Short and Long Run Forecasts: An Application of Seasonal Cointegration to Monthly Electricity Sales Forecasts," *Journal of Econometrics*, 40, 45–62.

Engle, R., Granger, C., Hylleberg, S., and Lee, H. (1993), "Seasonal Cointegration: The Japanese Consumption Function," *Journal of Econometrics*, 55, 275–298.

Franses, P. H. (1994), "A Multivariate Approach to Modeling Univariate Seasonal Time Series," *Journal of Econometrics*, 63, 133–151.

Ghysels, E. (1988), "A Study Toward a Dynamic Theory of Seasonality for Economic Time Series," *Journal of the American Statistical Association*, 83, 168–172.

——— (1991), "On the Seasonal Asymmetries and Their Implications for Stochastic and Deterministic Models of Seasonality," unpublished manuscript, University of Montreal, Dept. of Economics.

Ghysels, E., Lee, H., and Noh, J. (1992), "Testing for Unit Roots in Seasonal Time Series: Some Theoretical Extension and a Monte Carlo Investigation," unpublished manuscript, University of Montreal, Dept. of Economics.

Granger, C. W. J., and Newbold, P. (1986), *Forecasting Economic Time Series* (2nd ed.), New York: Academic Press.

Hannan, E. J. (1970), *Multiple Time Series*, New York: John Wiley.

Hansen, B. E. (1990), "Lagrange Multiplier Tests for Parameter Instability in Non-linear Models," unpublished manuscript, University of Rochester, Dept. of Economics.

——— (1992a), "Consistent Covariance Matrix Estimation for Dependent Heterogeneous Processes," *Econometrica*, 60, 967–972.

——— (1992b), "Testing for Parameter Instability in Linear Models," *Journal of Policy Modeling*, 14, 517–533.

Hansen, L., and Sargent, T. (1993), "Seasonality and Approximation Errors in Rational Expectations Models," *Journal of Econometrics*, 55, 21–55.

Harvey, A. C. (1990), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge, U.K.: Cambridge University Press.

Hylleberg, S. (1992), "Tests for Seasonal Unit Roots: A Comparative Study," unpublished manuscript, submitted to *Journal of Econometrics*.

Hylleberg, S., Engle, R., Granger, C., and Yoo, S. (1990), "Seasonal Integration and Cointegration," *Journal of Econometrics*, 44, 215–238.

Hylleberg, S., Jorgensen, C., and Sorensen, N. K. (1993), "Seasonality in Macroeconomic Time Series," *Empirical Economics*, 18, 321–335.

Hylleberg, S., and Pagan, A. (1994), "Seasonal Integration and the Evolving Seasonal Model," unpublished manuscript, Austalian National University.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., and Shin, Y. (1992), "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root?" *Journal of Econometrics*, 44, 159–178.

Newey, W. K., and West, K. D. (1987), "A Simple, Positive Semidefinite, Heteroskedastic and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

Nyblom, J. (1989), "Testing for the Constancy of Parameters Over Time," *Journal of the American Statistical Association*, 84, 223–230.

Osborn, D., and Smith, J. (1989), "The Performance of Periodic Autoregressive Models in Forecasting Seasonal U.K. Consumption," *Journal of Business & Economic Statistics*, 7, 117–127.

Pierce, D. (1978), "Seasonal Adjustment When Both Deterministic and Stochastic Seasonality Are Present," in *Seasonal Analysis of Economic Time Series*, ed. A. Zellner, Washington, DC: U.S. Department of Commerce, Bureau of the Census, pp. 242–269.

Saikkonen, P., and Luukkonen, R. (1993), "Testing for a Moving Average Unit Root in Autoregressive Integrated Moving Average Models," *Journal of the American Statistical Association*, 88, 596–601.

Shiskin, J., Young, A., and Musgrave, J. (1967), "The *X*-11 Variant of the Census Method II Seasonal Adjustment Program," Technical Paper 15, Washington, DC, Government Printing Office.

Tanaka, K. (1990), "Testing for a Moving Average Unit Root," *Econometric Theory*, 6, 433–444.

Thaler, R. (1987), "Anomalies: The January Effect," *Journal of Economic Perspectives*, 1, 197–201.