Periodic cointegration: Representation and inference

Boswijk, H.P.; Franses, P.H.

Publication date
1995

Published in
Review of Economics and Statistics

PERIODIC COINTEGRATION: REPRESENTATION AND INFERENCE

H. Peter Boswijk and Philip Hans Franses*

Abstract—This paper considers a new approach to the analysis of stable relationships between nonstationary seasonal time series. The basis of this approach is an error correction model in which both long-run effects and adjustment parameters are allowed to vary per season. First, we discuss theoretical arguments for such a periodic error correction model. We define periodic cointegration and compare this to the concept of seasonal cointegration. Next, we analyze statistical inference in the periodic error correction model. A sequential procedure is proposed, consisting of a test for periodic cointegration, an estimator of the cointegration parameters and adjustment coefficients, and a class of tests for the hypothesis that some of the parameters are constant over the seasons. The finite sample behavior of the proposed test statistics is analyzed in a limited Monte Carlo exercise. We conclude the paper with an application to a model of aggregate Swedish consumption.

I. Introduction

During the past decades, cointegration and error correction models (ECMs) have become much-favored tools in the econometric analysis of time series. One of the main attractions of these models is that they allow for a clear distinction between long-run relationships and short-run adjustment. Engle and Granger's (1987) definition of cointegration pertains to stable long-run relationships between non-stationary variables of a specific type, i.e., (non-seasonal) processes with a unit root at the zero frequency. For the analysis of seasonally observed time series, Hylleberg et al. (1990) and Engle et al. (1993) have extended this theory to cointegration at seasonal frequencies. The idea behind this extension is that several time series may have common non-stationary seasonal components. Osborn (1993) has argued that the associated ECM may be difficult to interpret from an economic point of view. The present paper proposes an alternative approach to cointegration between seasonal time series which attempts to overcome this difficulty.

The basis of our analysis is a *periodic error correction model*, i.e., an ECM in which both long-run parameters and adjustment coefficients are allowed to vary per season. If the individual time series contain a stochastic trend, then the periodic ECM implies *periodic cointegration*, i.e., stable but (possibly) seasonally varying long-run relationships. Such models were originally proposed by Birchenhall et al. (1989) and Franses and Kloek (1991). They are closely related to the univariate concept of periodic integration; see Osborn (1988) and Boswijk and Franses (1995, 1994). In the present paper we extend Boswijk's (1992) analysis of cointegration in conditional ECMs to periodic cointegration and error correction. Our procedure consists of a test for cointegration, an estimator of the cointegration parameters, and a class of tests for the hypothesis that some parameters are constant over the seasons. The applicability of our procedure is illustrated in a Monte Carlo experiment and an empirical application.

The outline of the paper is as follows. In section II, we compare seasonal cointegration and periodic cointegration as possible extensions of the standard ECM. We consider only quarterly time series; the models and methods discussed in this paper can be extended to monthly processes in a straightforward manner. In section III, we give the estimation and testing procedures, and we derive their asymptotic properties. The relevance of the asymptotics for the finite sample performance of the tests is checked via a limited Monte Carlo experiment in section IV. In section V we apply the methods to a model of real per capita consumption in Sweden. In the final section we conclude the paper with some remarks.

II. Error Correction and Cointegration for Seasonal Time Series

In this section we discuss seasonal and periodic cointegration and error correction as extensions

* University of Amsterdam and Erasmus University Rotterdam, respectively.

We wish to thank an anonymous referee for very constructive suggestions which have led to many improvements. We are also very grateful to Bengt Assarsson for providing us with the data, and to Marius Ooms for seasonally adjusting the data. Helpful comments from seminar participants in Tilburg, Groningen, Melbourne, Sydney, Canberra, Christchurch and Florence are gratefully acknowledged. Finally, we thank the Royal Netherlands Academy of Arts and Sciences and the Tinbergen Institute for financial support.

of the standard error correction model. Consider an observed bivariate time series $\{y_t, z_t,\ t = 1,\dots,n\}$, where the process $\{z_t\}$ is considered exogenous, whereas y_t is generated by the single-equation conditional error correction model

$$\Delta_1 y_t = \lambda(y_{t-1} - \theta z_{t-1}) + \beta\Delta_1 z_t + \epsilon_t, \qquad t = 1,\dots,n, \quad (1)$$

where λ is the error correction coefficient (or adjustment parameter), θ is the long-run multiplier, and β is the impact multiplier. Further, $\{\epsilon_t\}$ is white noise, i.e., a mean-zero uncorrelated process. If $-2 < \lambda < 0$, then the characteristic equation of (1) has its root outside the unit circle, which implies a stable long-run relationship $y = \theta z$. Assume that the explanatory variable is integrated of order 1 ($I(1)$). If the model is stable, then the linear combination $y_t - \theta z_t$ is stationary, although z_t (and y_t if $\theta \neq 0$) are integrated of order 1; i.e., y_t and z_t are cointegrated of order $(1, 1)$ with cointegrating vector $(1, -\theta)'$. On the other hand, if $\lambda = 0$, then there is no adjustment towards equilibrium and hence no cointegration.

A. *Seasonal Integration and Cointegration*

The assumption made above that y_t and z_t are $I(1)$ implies that these series contain a stochastic trend, which appears to be a stylized fact of many observed economic series. In addition, seasonal time series often display a slowly changing seasonal pattern. In the standard univariate analysis of such seasonal time series, see Box and Jenkins (1976), non-stationary fluctuations are removed before the analysis. This can be accomplished by taking annual differences, i.e., applying the fourth-difference filter Δ_4, where $\Delta_s \equiv (1 - L^s)$, with L the lag operator (such that $L^k x_t = x_{t-k}$, $k \in \mathbb{N}$). Hylleberg et al. (1990, henceforth HEGY) propose a test procedure to check the validity of this transformation. Their test is based on the decomposition $(1 - L^4) = (1 - L)S(L)$, where $S(L) = (1 + L + L^2 + L^3) = (1 + L)(1 - iL)(1 + iL)$. The crucial aspect is that non-seasonal non-stationarity is removed by the $(1 - L)$ filter, whereas seasonal non-stationarity may be removed by $S(L)$. Applying both Δ_1 and Δ_4, as suggested by Box and Jenkins' (1976) airline model, often appears to be superfluous in practice, see, e.g., Osborn (1990). HEGY show that an AR(4) model for a univariate series $\{x_t\}$ can always be rewritten as

$$\Delta_4 x_t = \pi_1 x_{1,t-1} + \pi_2 x_{2,t-1} + \pi_3 x_{3,t-2} + \pi_4 x_{3,t-1} + \epsilon_t, \qquad t = 1,\dots,n, \quad (2)$$

where

$$x_{1t} = S(L)x_t,\ x_{2t} = -(1 - L + L^2 - L^3)x_t,$$

and

$$x_{3t} = -(1 - L^2)x_t.$$

For AR(p) models with $p > 4$, the right-hand side of (2) should be augmented with lagged $\Delta_4 x_t$ terms. If $\pi_i = 0$, $i = 1,\dots,4$, then an AR model for $\Delta_4 x_t$ results, so that the Δ_4 filter is required to obtain stationarity. In that case x_t is said to be seasonally integrated. If, on the other hand, $\pi_1 = 0$ but $\pi_i \neq 0$, $i = 2, 3, 4$, then (2) can be rewritten as an AR model for $\Delta_1 x_t$. Other combinations of π-values are also possible. HEGY provide critical values for t-tests for the hypotheses $\pi_1 = 0$ and $\pi_2 = 0$, and an F-test for $\pi_3 = \pi_4 = 0$, *inter alia*.

HEGY also propose a multivariate extension of this approach, involving seasonal cointegration (see also Engle et al., 1993). For a bivariate time series $x_t = (y_t, z_t)'$, define the transformations x_{it}, $i = 1, 2, 3$, similarly to the univariate case above. The simplest seasonal ECM reads

$$\begin{aligned}\Delta_4 y_t = {} & \gamma_{11}\alpha_1' x_{1,t-1} + \gamma_{12}\alpha_2' x_{2,t-1} \\ & + (\gamma_{14} - \gamma_{13}L)\alpha_3' x_{3,t-1} \\ & + (\gamma_{13} + \gamma_{14}L)\alpha_4' x_{3,t-1} + \epsilon_t, \qquad (3)\end{aligned}$$

together with a similar equation for $\Delta_4 z_t$. In comparison with the non-seasonal ECM in (1), there are four target relationships in (3), formulated in terms of the transformed series y_{it} and z_{it}. As demonstrated by Osborn (1993), the model corresponds to

$$\Delta_4 y_t = \sum_{i=1}^{4} \lambda_i (y_{t-i} - \theta_i z_{t-1}) + \epsilon_t, \qquad t = 1,\dots,n. \quad (4)$$

Hence the seasonal cointegration model can be interpreted as a model with four different target relationships and adjustment parameters, each associated with a different lag. An alternative formulation, which allows parameter variation with respect to the seasons instead of the lags, is considered next.

B. Periodic Integration and Cointegration

Before we discuss the periodic cointegration model, we first consider some relevant concepts in univariate periodic time series. Consider again a univariate quarterly series $\{x_t\}$. The first-order *periodic autoregression* (PAR(1)) is given by

$$x_t = \phi_s x_{t-1} + \epsilon_t, \qquad t = 1, \ldots, n, \qquad s = 1, \ldots, 4, \quad (5)$$

with $\{\epsilon_t\}$ white noise. Thus in season s, the autoregressive parameter equals ϕ_s. Given this parameter variation, the process (5) is not covariance-stationary. Issues of stationarity and integration in periodic processes are analyzed most conveniently in a multivariate model for the annual process $\{X_T,\ T = 1, \ldots, N\}$, obtained by stacking the 4 quarterly observations of x_t in year T into a 4×1 vector X_T. Thus $X_{sT} = x_{4(T-1)+s}$, and $N = n/4$ denotes the total number of years in the sample (see, e.g., Tiao and Grupe (1980)). Following Franses (1994), we shall call $\{X_T\}$ the vector of quarters (VQ) process of $\{x_t\}$. The advantage of the VQ representation is that a periodic and hence time-varying model of $\{x_t\}$ can be written as a constant-parameter model of $\{X_T\}$. In particular, any periodic autoregression for $\{x_t\}$ implies a vector autoregression for $\{X_T\}$.

Building on earlier work of Osborn et al. (1988), Boswijk and Franses (1994) define the process $\{x_t\}$ to be periodically integrated of order 1 ($PI(1)$) if the characteristic equation of the VAR model of $\{X_T\}$ has exactly 1 unit root and all other roots outside the unit circle; if there are no unit roots, then x_t is $PI(0)$ or periodically stationary. The condition of 1 unit root implies that the 4 components of X_T have a single common trend and hence 3 cointegrating relationships. Thus, the seasons are not allowed to drift too far apart. This is in contrast with a seasonally integrated time series, where the factor $(1 - L)^4$ in the autoregressive polynomial implies that the corresponding VQ process is $I(1)$ but not cointegrated (see Osborn, 1993).

Franses (1994) develops a test procedure to choose between periodic integration and seasonal integration, based on an application of Johansen's (1991) cointegration analysis to a VAR model for the VQ process. Because such a VAR model may easily be overparametrized, Boswijk and Franses (1995) propose a likelihood ratio test for periodic integration of order 1 based on the original periodic autoregression. Boswijk and Franses show that a studentized version of this test has the same asymptotic distribution under the null hypothesis as the standard Dickey-Fuller unit root test, tabulated in Fuller (1976). The test has been generalized to higher order PAR processes by Boswijk and Franses (1994).

The periodically integrated AR (PIAR) process may be extended to periodic cointegration for multiple time series in various ways. The most general set-up is considered by Franses (1995a), who analyzes cointegration in VAR models for the VQ process of a quarterly $m \times 1$ vector time series $\{x_t\}$. However, the above-mentioned hazard of overparametrization is even more prominent in this case.

Therefore, following Birchenhall et al. (1989) and Franses and Kloek (1991), we analyze a class of models with somewhat more structure imposed. Consider the following periodic conditional error correction model

$$\Delta_4 y_t = \beta' \Delta_4 z_t + \lambda_s (y_{t-4} - \theta_s' z_{t-4}) + \epsilon_t, \qquad t = 1, \ldots, n, \quad s = 1, \ldots, 4, \quad (6)$$

where $\{\epsilon_t\}$ is a white noise process with variance σ^2. This model may be directly compared to (4). Note that the $\Delta_4 z_t$ term is the result of conditioning on z_t, which may also be performed in the seasonal cointegration model. Both models imply four different long-run relationships. However, in (3)–(4) the parameters of these relations vary with the lag, whereas in (6), they vary with the season. The motivation for this parameter variation is as follows. First, the seasonally varying adjustment parameter λ_s is a reflection of adjustment costs varying over the quarters. For example, this may be the case with employment in typically seasonal industries, such as tourism or construction. Further, the long-run parameter θ_s and the associated target relationship is not constant. This reflects seasonally varying preferences, e.g., a particular consumption bundle may not generate the same level of utility in the summer as in the winter.

Having discussed periodic integration and error correction, we are now able to define the concept of periodic cointegration. We use $\{D_{st}, t = 1,\ldots,n, s = 1,\ldots,4\}$ to denote a set of seasonal dummies.

DEFINITION 1: *Consider a quarterly m-vector process* $\{x_t\}$, *with a VQ process which is integrated of order* 1. *Then*

(*i*) x_t *is said to be fully periodically cointegrated of order* (1, 1), *if there exist* $m \times r$ *matrices* α_s *of full column rank,* $0 < r < m$, $s = 1,\ldots,4$, *such that the VQ process of* $\sum_{s=1}^{4} D_{st}\alpha_s' x_t$ *is stationary.*

(*ii*) x_t *is said to be partially periodically cointegrated of order* (1, 1), *if there exist* $m \times r$ *matrices* α_s *of full column rank,* $0 < r < m$, $s = 1,\ldots,4$, *such that at least one of the components of the VQ process of* $\sum_{s=1}^{4} D_{st}\alpha_s' x_t$ *is stationary.*

Notice that $\{x_t\}$ is allowed to be $I(1)$, seasonally integrated or periodically integrated. If a process is fully cointegrated, then in each of the seasons there is adjustment towards the long-run relationship $\alpha_s' x_t$. If it is partially cointegrated, then in some of the quarters there will be no adjustment.

In this paper, we shall consider at most 1 cointegrating relationship per season, so $r = 0$ or $r = 1$. Moreover, we shall assume that the cointegrating vectors α_s can be normalized with respect to the first components of x_t, i.e., $\alpha_s' x_t = y_t - \theta_s' z_t$. This naturally leads to (6), although the system needs to be completed with a model for z_t. In the next section we shall consider statistical inference on periodic cointegration in a slight generalization of (6) under certain conditions on the explanatory variables z_t.

III. Statistical Inference on Periodic Cointegration

In this section we restate the periodic error correction model, together with some assumptions. Then we give a class of Wald tests for periodic cointegration, and derive their asymptotic properties. Next, we analyze inference on the long-run parameters and adjustment coefficients. Finally, we bring together the various estimators and test statistics in an empirical modelling procedure, summarizing the implications of the results of this section for practical inference.

A. *The Model and Assumptions*

Consider the periodic error correction model for a quarterly time series y_t, conditional upon a k-vector of quarterly time series z_t:

$$\Delta_4 y_t = \lambda_s(y_{t-4} - \theta_s' z_{t-4}) + \sum_{i=1}^{p} \gamma_i \Delta_4 y_{t-i} + \sum_{i=0}^{p} \beta_i' \Delta_4 z_{t-i} + \epsilon_t,$$
$$t = 1,\ldots,n, \quad s = 1,\ldots,4. \quad (7)$$

Here $\{\epsilon_t\}$ is a white noise process with variance σ^2, γ_i, $i = 1,\ldots,p$, are scalar parameters whereas β_i, $i = 0,\ldots,p$, are $k \times 1$ parameter vectors, and finally λ_s and θ_s are the adjustment parameter and the vector of long-run parameters in season s, respectively. Notice that in (7), no deterministic regressors such as seasonal dummy variables or a linear trend term are included. Below we shall discuss such extensions in detail.

For the asymptotic results to be derived here, we shall require some assumptions on the disturbances and the conditioning variables. First, we assume that the fourth difference of each component of z_t has an unconditional mean that depends upon the season only, i.e.,

$$\Delta_4 z_{jt} = \sum_{s=1}^{4} D_{st}\mu_{js} + u_{jt},$$
$$j = 1,\ldots,k, \quad t = 1,\ldots,n, \quad (8)$$

where $\{u_{jt}\}$ are zero-mean processes. One can expect the restrictions $\mu_{j1} = \mu_{j2} = \mu_{j3} = \mu_{j4}$, $j = 1,\ldots,k$, to hold in practice, since otherwise the seasons will diverge as $t \to \infty$. Let $u_{0t} = \epsilon_t$ and $u_t = (u_{0t}, u_{1t},\ldots,u_{kt})'$. We shall consider two different VQ processes, with its components ordered by variable or by season. Thus, let $U_T = (U_{0T}', U_{1T}',\ldots,U_{kT}')'$ where U_{jT} is the VQ process of u_{jt}, and alternatively define $\tilde{U}_T = (\tilde{U}_{1T}',\ldots,\tilde{U}_{4T}')'$, where $\tilde{U}_{sT}$ stacks the s^{th} components of U_{jT}, $j = 0,\ldots,k$. Notice that $\tilde{U}_T = K_{4m}U_T$, where $m = k + 1$ and K_{4m} is the commutation matrix (cf. Magnus and Neudecker, 1988). Define the long-run covariance matrices

$$\Omega = \sum_{i=-\infty}^{\infty} E[U_T U'_{T-i}],$$

$$\tilde{\Omega} = \sum_{i=-\infty}^{\infty} E\left[\tilde{U}_T \tilde{U}'_{T-i}\right] = K_{4m}\Omega K'_{4m}. \quad (9)$$

The matrix Ω consists of variable-blocks Ω_{ij}, $i, j = 0, \ldots, k$, whereas $\tilde{\Omega}$ contains season-blocks $\tilde{\Omega}_{rs}$, $r, s = 1, \ldots, 4$.

ASSUMPTION 1: *The VQ process* $\{U_T\}$ *satisfies the following conditions*:

- (*i*) $\{U_T\}$ *is stationary and obeys a multivariate invariance principle, i.e.,* $N^{-1/2}\sum_{T=1}^{[rN]} U_T \Rightarrow B(r)$, *where* $[rN]$ *is the integer part of rN, the symbol* "$\Rightarrow$" *denotes weak convergence, and* $B(r)$ *is a* $(4m)$*-vector Brownian motion process with covariance matrix* Ω.
- (*ii*) $\{U_{0T}\}$ *is a martingale difference sequence relative to the sequence of σ-fields generated by* $\{U_{1T}, \ldots, U_{kT}, U_{T-i}, i = 1, 2, \ldots\}$, *with covariance matrix* $\Omega_{00} = \sigma^2 I_4$, *so that for* $j = 1, \ldots, k$, $\sum_{i=0}^{\infty} E[U_{0T}U'_{j,T-i}] = 0$.
- (*iii*) $\tilde{\Omega}_{ss} > 0$ (*positive definite*), $s = 1, \ldots, 4$.

By stationarity of $\{U_T\}$ we exclude explosive or $I(d)$ processes with $d > 1$. Condition (*ii*) simply restates that (7) represents a conditional model of y_t given z_t and the past, with homoscedastic martingale difference errors. Finally, (*iii*) states that in each season, the cumulative sum of u_t is not cointegrated. By this assumption two cases are excluded. Firstly, cointegrating relationships between the components of z_t only are assumed not to exist. Moreover, the possibility is excluded that all λ_s in (7) equal zero and yet y_t and z_t are cointegrated due to error correcting behavior of z_t, since that would imply that the cumulative sum of $u_{0t} = \epsilon_t$ is cointegrated with z_t. Thus, under Assumption 1 the parameters λ_s determine whether or not y_t and z_t are periodically cointegrated, which is the basis of the Wald tests proposed below. If $\lambda_s < 0$ in some season s, then y_t and z_t are partially cointegrated in that season; if all $\lambda_s < 0$, then full periodic cointegration holds. Finally, observe that Assumption 1 does not require Ω or $\tilde{\Omega}$ to be positive definite; in fact, when z_{jt} is periodically integrated, then the corresponding block Ω_{jj} has rank 1.

ASSUMPTION 2: *The VQ process of* $\Delta_4 y_t$ *is stationary.*

This entails that the characteristic equation of the VQ representation of (7), given in (A1) in appendix 1, has at most 4 roots equal to one and all other roots outside the unit circle, so that explosive processes (as well as, e.g., $I(2)$ processes) are excluded.

ASSUMPTION 3 (*Long-run weak exogeneity*): $\Omega_{0j} = 0$, $j = 1, \ldots, k$.

This assumption will be required only for mixed Gaussian inference on the cointegration parameters, see Theorem 2; it does not play any role in testing for cointegration. The assumption implies that the Brownian motions corresponding to z_t and the cumulative sum of ϵ_t are independent. This requires that in a model for z_t, the coefficients of the error correction terms $(y_{t-4} - \theta_s' z_{t-4})$ are all equal to zero. This is related to the notion of weak exogeneity, see Engle et al. (1983). For a discussion of the role of weak exogeneity in cointegrated systems, see Johansen (1992), who shows that the assumption implies that inference on the cointegration parameters in conditional models is efficient and that hypothesis tests have an asymptotic χ^2 distribution. In a recent paper, Dolado (1992) argues that weak exogeneity is too strong an assumption for these results, and that a zero long-run correlation between the errors and the conditioning variables (called long-run weak exogeneity) suffices. Boswijk (1992) proposes a Lagrange-multiplier test for (long-run) weak exogeneity, which involves a simple variable addition test statistic for the estimated error correction terms in a model for z_t. This test can be straightforwardly extended to the present context.

B. *Testing for Periodic Cointegration*

As discussed in the previous subsection, periodic cointegration requires the adjustment parameters λ_s to be strictly smaller than zero. In this section we propose a class of Wald tests for the null hypothesis of no cointegration against the alternative of periodic cointegration. Two versions of this test are considered: the $Wald_s$ statistic for the periodic cointegration in season s, and the joint *Wald* statistic. Both versions are extensions of Boswijk's (1994) cointegration test, which has been developed for non-seasonal data.

In order to obtain expressions for the test statistics, we rewrite the model (7) more concisely. Recall that $x_t = (y_t, z_t')'$, and that D_{st}, $s = 1, \ldots, 4$, denote seasonal dummy variables. Next, define $\delta_s = (\delta_{1s}, \delta_{2s}')' = (\lambda_s, -\lambda_s \theta_s')'$, $s = 1, \ldots, 4$. Finally, let w_t denote the vector of differenced explanatory variables in (7), with π as its coefficient vector. Then (7) becomes

$$\Delta_4 y_t = \sum_{s=1}^{4} \delta_s' D_{st} x_{t-4} + \pi' w_t + \epsilon_t, \qquad t = 1, \ldots, n. \quad (10)$$

Notice that from the definition of δ_s, it follows that $\lambda_s = 0$ implies $\delta_s = 0$. Thus the null and alternative hypotheses for the $Wald_s$ test are, for some particular s

$$H_{0s}: \delta_s = 0, \qquad H_{1s}: \delta_s \neq 0, \quad (11)$$

whereas for the *Wald* test, they are

$$H_0: \delta = (\delta_1', \ldots, \delta_4')' = 0, \qquad H_1: \delta \neq 0. \quad (12)$$

Notice that the $Wald_s$ statistics are designed to test for partial periodic cointegration, whereas the *Wald* statistic corresponds to the full periodic cointegration hypothesis. Nevertheless, both tests will have power against both alternative hypotheses.

Let $\hat{\delta}_s$ denote the ordinary least-squares (OLS) estimator of δ_s, and let $\hat{V}[\hat{\delta}_s]$ denote the OLS covariance matrix estimator. Similarly, define $\hat{\delta} = (\hat{\delta}_1', \ldots, \hat{\delta}_4')'$ and its estimated covariance matrix $\hat{V}[\hat{\delta}]$. Then the Wald statistics are given by

$$Wald_s = \hat{\delta}_s' \left(\hat{V}[\hat{\delta}_s] \right)^{-1} \hat{\delta}_s = (n - l)(RSS_{0s} - RSS_1)/RSS_1, \quad (13)$$

$$Wald = \hat{\delta}' \left(\hat{V}[\hat{\delta}] \right)^{-1} \hat{\delta} = (n - l)(RSS_0 - RSS_1)/RSS_1, \quad (14)$$

where l is the number of estimated parameters in (10), and where RSS_{0s}, RSS_0 and RSS_1 denote the OLS residual sum of squares under H_{0s}, H_0 and H_1, respectively.

The model (7) contains no deterministic regressors. However, if the long-run relationships require an intercept, then seasonal dummies should be added to the error correction model, i.e., the tests should be based on

$$\Delta_4 y_t = \sum_{s=1}^{4} \mu_{0s} D_{st} + \sum_{s=1}^{4} \delta_s' D_{st} x_{t-4} + \pi' w_t + \epsilon_t, \qquad t = 1, \ldots, n. \quad (15)$$

Finally, if the variables contain a drift, i.e., $\mu = (\mu_0', \mu_1', \ldots, \mu_k') \neq 0$, where $\mu_j = (\mu_{j1}, \ldots, \mu_{j4})'$, $j = 0, \ldots, k$, then this will change the asymptotic distributions of the Wald statistics, as follows from the general results by Park and Phillips (1988, 1989) and Sims et al. (1990). Thus, in order to obtain distributions that are invariant to the presence of drifts, a set of linear trend terms should be added to the regressors, leading to

$$\Delta_4 y_t = \sum_{s=1}^{4} (\mu_{0s} D_{st} + \tau_s D_{st} t) + \sum_{s=1}^{4} \delta_s' D_{st} x_{t-4} + \pi' w_t + \epsilon_t, \qquad t = 1, \ldots, n. \quad (16)$$

This also allows for a trend term to appear in the long-run relationships.

In Theorem 1 we state the asymptotic properties of the $Wald_s$ and *Wald* statistic. Unless indicated otherwise, all integrals in this section are taken from 0 to 1 and with respect to Lebesgue measure (e.g., $\int_0^1 B(r)\,dr$ is denoted by $\int B$). Moreover, we shall use the following functional:

$$h(W, U) = \int dWU' \left[\int UU' \right]^{-1} \int UdW', \quad (17)$$

where $W(r)$ is a scalar standard Brownian motion process and $U(r)$ is a vector process on $[0, 1]$ satisfying $\int UU' > 0$. The squared multiple correlation coefficient between $W(1)$ and $U(1)$ is denoted by $\rho^2(W, U)$. Finally, we use U^* and U^{**} to denote "demeaned" and "detrended" processes, respectively (cf. Park and Phillips, 1988, pp. 474–475):

$$U^*(r) = U(r) - \int U,$$

$$U^{**}(r) = U^*(r) - \left(r - \tfrac{1}{2}\right) \left[\int \left(r - \tfrac{1}{2}\right)^2 \right]^{-1} \int \left(r - \tfrac{1}{2}\right) U^*. \quad (18)$$

THEOREM 1: *Let y_t be generated by* (7), *let Assumptions* 1 *and* 2 *hold, and assume* $\mu = 0$. *Then,* *as $n \to \infty$, and under H_{0s} and H_0, respectively,*

$$Wald_s \Rightarrow h(W_s, U_s), \qquad Wald \Rightarrow \sum_{s=1}^{4} h(W_s, U_s), \tag{19}$$

where $W_s(r)$, $s = 1, \ldots, 4$, are independent standard Brownian motion processes, whereas $U_s(r)$ are standard m-vector Brownian motion processes, satisfying:

(*i*) *if either $\lambda_q < 0$, $q \neq s$, or $\gamma_i = 0$ and $\beta_i = 0$ for $i \bmod 4 \neq 0$, then $\rho^2(W_s, U_s) = 1$;*
(*ii*) *if, in addition, $\tilde{\Omega}_{sq} = 0$, then U_s and U_q are independent.*

If the test statistics are based on (15), *and $\mu = 0$, then U_s in* (19) *should be replaced by U_s^*; if they are based on* (16), *then U_s should be replaced by U_s^{**} (regardless of the value of μ).*

Under H_{1s}, and as $n \to \infty$, $Wald_s \to \infty$ and $Wald \to \infty$, so that the tests are consistent.

Proofs are given in appendix 1. Theorem 1 implies that the asymptotic null distributions of the $Wald_s$ and $Wald$ statistics depend upon a number of nuisance parameters, viz. the correlation between W_s and U_s, and the correlation between U_s and U_q, $s \neq q$.

Consider first the $Wald_s$ statistic. The possible correlation between U_q and U_s, $q \neq s$, of course leaves the asymptotic distribution of this statistic unaffected. Moreover, since the null hypothesis H_{0s} does not restrict λ_q, $q \neq s$, the condition $\lambda_q < 0$ is *generically* satisfied. Thus, generically we have $\rho^2(W_s, U_s) = 1$, in which case we may define $W_s = U_{1s}$ without loss of generality. This is also the case if y_t is only related to z_t, $x_{t-4}, x_{t-8}, \ldots$; then $\rho^2(W_s, U_s) = 1$ regardless of λ_q, $q \neq s$. The distribution of $h(U_{1s}, U_s)$ is identical to the one obtained in Boswijk (1994) for the non-seasonal case, where simulated critical values are tabulated. In appendix 2, table A1, these are reiterated, based on 50,000 replications. Finally, if $\rho^2(W_s, U_s) < 1$, then it can be shown that the distribution of $h(W_s, U_s)$ is a mixture of a $\chi^2(m)$ distribution and the distribution of $h(U_{1s}, U_s)$. Since this distribution is more concentrated towards zero, the actual significance level will be smaller than the nominal one if critical values from table A1 are used, so that the test size (maximum significance level) can be controlled.

For the joint $Wald$ test, Monte Carlo simulation of the distribution of $\Sigma_{s=1}^4 h(W_s, U_s)$ reveals that it is invariant to possible correlations between U_s and U_q, $q \neq s$. This is also confirmed in the finite sample Monte Carlo study in section IV. Thus, since W_s is independent of W_q, we may take the four terms $h(W_s, U_s)$ to be independent. However, since the null hypothesis now entails that all $\lambda_s = 0$, the distribution of each of the terms will depend upon $\rho^2(W_s, U_s)$, which only equals 1 if the differenced variables in (7) only appear with their lags being a multiple of 4. In table A2, the asymptotic critical values are given for this case, i.e., for $\Sigma_{s=1}^4 h(U_{1s}, U_s)$. As indicated above, the appropriate critical values for the other case will be smaller, so that the size is controllable.

C. Inference on Long-run Parameters and Testing for Periodicity

If the tests proposed in the previous subsection reject the null hypothesis of no cointegration, then the next step is to estimate the cointegration parameters and adjustment coefficients, and to test hypotheses on these parameters. For that purpose, we analyze the asymptotic properties of such estimators and test statistics here. In particular, we consider the hypothesis that some or all parameters are constant over the seasons. Throughout this subsection, we assume that y_t and z_t are fully periodically cointegrated. Results for the partial cointegration case can be derived analogously.

From the definition of δ_s, we have that $\lambda_s = \delta_{1s}$, and $\theta_s = -\delta_{2s}/\delta_{1s}$. This suggests the following estimators of the adjustment coefficients and long-run parameters:

$$\hat{\lambda}_s = \hat{\delta}_{1s}, \qquad \hat{\theta}_s = -\hat{\delta}_{2s}/\hat{\delta}_{1s}, \qquad s = 1, \ldots, 4. \tag{20}$$

These estimators are closely related to Stock's (1987) nonlinear least-squares estimator for the nonperiodic case, except that he excludes the contemporaneous Δz_{jt} regressors in (7). If z_t is weakly exogenous for the parameters of (7), i.e., if the parameters of a model for z_t given the past are variation independent of the parameters of (7), and if in addition the errors $\{\epsilon_t\}$ are Gaussian, then (20) gives the maximum likelihood estimates of λ_s and θ_s. Define the parameter vector

$\varphi = (\lambda', \theta')'$, with $\lambda = (\lambda_1, \dots, \lambda_4)'$ and $\theta = (\theta_1', \dots, \theta_4')'$, and similarly $\hat{\varphi}$. The Jacobian matrix or derivative of the transformation $\varphi(\delta)$ is

$$D = \frac{\partial \varphi}{\partial \delta'} = \begin{bmatrix} I_4 \otimes e_1' \\ J \end{bmatrix}, \quad J = \text{diag}(J_1, \dots, J_4), \tag{21}$$

where e_1 is the first unit m-vector, and where

$$\begin{aligned} J_s = \frac{\partial \theta_s}{\partial \delta_s'} &= \left[\delta_{1s}^{-2} \delta_{2s} \ \vdots \ -\delta_{1s}^{-1} I_k \right] \\ &= -\frac{1}{\lambda_s} \left[\theta_s \ \vdots \ I_k \right], \qquad s = 1, \dots, 4. \end{aligned} \tag{22}$$

where $\Sigma_v > 0$ is some 4×4 matrix, defined in appendix 1. *If and only if, in addition, Assumption* 3 *holds, then $\hat{\varphi}$ is asymptotically mixed Gaussian, i.e., $\Upsilon_N(\hat{\varphi} - \varphi) \Rightarrow \int_{G>0} N(0, G)\, dP(G)$.*

In order to state the asymptotic properties of $\hat{\varphi}$, define the scaling matrix $\Upsilon_N = \text{diag}(\sqrt{N} \cdot I_4, N \cdot I_{4k})$. Moreover, express the $(4m)$-vector Brownian motion process $B(r)$ as $K_{m4}(B_1(r)', \dots, B_4(r)')'$, where $B_s(r)$ is an m-vector Brownian motion with covariance matrix $\tilde{\Omega}_{ss}$, corresponding to the partial sum of $\tilde{U}_{sT}$ (cf. Assumption 1). Finally, partition $B_s(r)$ as $(B_{0s}(r), B_{1s}(r)')'$. In Theorem 2 the asymptotic properties of $\hat{\varphi}$ and $\hat{V}[\hat{\varphi}]$ are given for the case where no seasonal intercepts or trends are included in (7), and $\mu = 0$. Generalizations in this direction, i.e., for the case where $\mu \neq 0$, can be dealt with analogously, and do not lead to fundamentally different results.

THEOREM 2: *Let y_t be generated by* (7) *with $\lambda_s < 0$, $s = 1, \dots, 4$, let $\mu = 0$, and let Assumptions* 1 *and* 2 *hold. Then, as $n \to \infty$,*

$$\Upsilon_N(\hat{\varphi} - \varphi) \Rightarrow \begin{pmatrix} N(0, \sigma^2 \Sigma_v^{-1}) \\ \text{vec}\left\{ \left[\int B_{1s} B_{1s}' \right]^{-1} \int B_{1s}\, dB_{0s}/\lambda_s,\ s = 1, \dots, 4 \right\} \end{pmatrix}, \tag{23}$$

$$\Upsilon_N \hat{V}[\hat{\varphi}] \Upsilon_N \Rightarrow G = \sigma^2 \begin{bmatrix} \Sigma_v & 0 \\ 0 & \text{diag}\left\{ \lambda_s^2 \int B_{1s} B_{1s}',\ s = 1, \dots, 4 \right\} \end{bmatrix}^{-1}, \tag{24}$$

where $\Sigma_v > 0$ is some 4×4 matrix, defined in appendix 1. *If and only if, in addition, Assumption* 3 *holds, then $\hat{\varphi}$ is asymptotically mixed Gaussian, i.e., $\Upsilon_N(\hat{\varphi} - \varphi) \Rightarrow \int_{G>0} N(0, G)\, dP(G)$.*

Theorem 2 implies that, regardless of any exogeneity assumptions, the long-run parameter estimators $\hat{\theta}_s$ are superconsistent just as in the non-periodic case, cf. Stock (1987). However, Assumption 3 is required for the estimators to have an asymptotic mixed Gaussian distribution (and to be asymptotically efficient); we have used $\int_{G>0} N(0, G)\, dP(G)$ to denote a distribution which is normal conditional upon G in (24). The most relevant consequence of this result is summarized in Corollary 1. Consider the (generalized) F-statistic

$$F = \frac{1}{h}(R\hat{\varphi} - r)' \left[R\hat{V}[\hat{\varphi}]R' \right]^{-1} (R\hat{\varphi} - r) \tag{25}$$

for the null hypothesis

$$H_0\colon R\varphi = R_1\lambda + R_2\theta = r, \tag{26}$$

where $R = [R_1 \vdots R_2]$ is a known $h \times 4m$ matrix of full row rank and r is a known h-vector.

COROLLARY 1: *Make the assumptions of Theorem* 2. *If and only if either Assumption* 3 *or the condition* $\text{rank}(R_1) = h$ *(or both) hold, then $hF \Rightarrow \chi^2(h)$ under H_0 and as $n \to \infty$.*

The condition $\text{rank}(R_1) = h$ implies that the asymptotic distribution of F is determined by the distribution of $\hat{\lambda}$, which is normal whether or not Assumption 3 holds. If this is not the case, and in particular if the restrictions concern θ only, then long-run weak exogeneity is a necessary requirement for the conventional χ^2 null distribution (or the F-distribution for the F-statistic) to apply. If Assumption 3 does not hold, then the asymptotic distribution of F will depend upon nuisance parameters, cf. Park and Phillips (1988, 1989) and

Sims et al. (1990). Without proof, we state that the same results will apply to a likelihood ratio, Lagrange multiplier or Wald statistic for H_0 in (26); this can be proved using Taylor series expansions as usual.

A class of hypotheses which are of particular interest in the present set-up concerns the parameter variation. We distinguish three null hypotheses on the parameters λ_s and θ_s in (7), and on δ_s in (10):

$$\begin{aligned} H_{0\lambda}&: \lambda_s = \lambda_1, \quad s = 2,3,4; \\ H_{0\theta}&: \theta_s = \theta_1, \quad s = 2,3,4; \\ H_{0\delta}&: \delta_s = \delta_1, \quad s = 2,3,4. \end{aligned} \tag{27}$$

Notice that $H_{0\delta} = H_{0\lambda} \cap H_{0\theta}$. As the alternative hypothesis we may consider simply $\delta_s \neq \delta_1$, $s = 2,3,4$, but also more restricted hypotheses. For example, we may wish to test constancy of θ_s under the maintained hypothesis of constant λ_s, which is a test for $H_{0\delta}$ against $H_{0\lambda}$. Each of these hypotheses may be tested using the generalized F-statistic in (25). For $H_{0\lambda}$ and $H_{0\delta}$ these will correspond to the classical F-statistic, since both the model and the hypotheses are linear in λ and δ. However, for the hypothesis $H_{0\theta}$, we might instead of (25) use the likelihood ratio-based F-statistic $F_{LR} = [(n - l)/h](RSS_1 - RSS_0)/RSS_1$, where l is the number of estimated parameters in (10), where RSS_1 is the unrestricted OLS residual sum of squares, whereas RSS_0 is the restricted residual sum of squares, obtained from non-linear least-squares estimation of

$$\Delta_4 y_t = \sum_{s=1}^{4} \lambda_s D_{st}(y_{t-4} - \theta_1' z_{t-4}) + \pi' w_t + \epsilon_t, \quad t = 1, \ldots, n. \tag{28}$$

Corollary 1 implies that the F-statistic for $H_{0\lambda}$ will always have an asymptotic F-distribution under the null. This is related to the fact that under the cointegration hypothesis, the adjustment parameters λ_s are coefficients of stationary variables (cf. Sims et al., 1990). The test statistics for the other two periodicity hypotheses, however, require weak exogeneity.

As discussed at the beginning of this section, a Lagrange-multiplier statistic for weak exogeneity can be obtained as follows. First, construct the 4×1 vector $\{\hat{v}_t,\ t = 1, \ldots, n\}$ of estimated disequilibrium errors, so $\hat{v}_{st} = D_{st}(y_t - \hat{\theta}_s' z_t)$, $s = 1, \ldots, 4$. Next, test the hypothesis $\kappa = 0$ in the suitably selected multivariate model

$$A(L)\Delta_4 z_t = B(L)\Delta_4 y_{t-1} + \kappa' \hat{v}_{t-4} + \eta_t, \quad t = 1, \ldots, n, \tag{29}$$

where $A(L)$ and $B(L)$ are matrix lag polynomials with $A(0) = I_k$, and $\{\eta_t\}$ is a vector white noise process. Under the null hypothesis, (29) implies that z_t is seasonally integrated; if some of the components of z_t are periodically integrated, the model should be changed accordingly. Since $\hat{\theta}_s$ is estimated super-consistently and v_t is stationary, a Wald, likelihood ratio or Lagrange-multiplier statistic for $\kappa = 0$ has an asymptotic χ^2 null distribution, see Boswijk (1992).

If the hypothesis is rejected, a number of different approaches to obtain efficient estimators and χ^2 hypothesis tests may be found in the (non-periodic) cointegration literature. Firstly, we may start out with a multivariate periodic model for $x_t = (y_t, z_t')'$ and test for cointegration in this system using Johansen's (1991) approach. For bivariate x_t, this is analyzed in Franses (1995a). However, as mentioned above, such models can easily be overparametrized.

A second approach that deals with this problem is to treat the short-run dynamics nonparametrically. For example, Phillips and Hansen's (1990) fully modified estimator could be extended to the present context. More easily, we could follow the approach of Stock and Watson (1993) inter alia, to include both lags and *leads* of the differenced conditioning variables in the regression equation. In our case, this would lead to estimating θ_s in

$$\Delta_4 y_t = \sum_{s=1}^{4} \lambda_s D_{st}(y_{t-4} - \theta_s' z_{t-4}) + \sum_{i=1}^{p} \gamma_i \Delta_4 y_{t-i} + \sum_{i=-p}^{p} \beta_i' \Delta_4 z_{t-i} + \epsilon_t, \quad t=1, \ldots, n. \tag{30}$$

An explicit asymptotic analysis of such estimators and related test statistics is beyond the scope of this paper, but we can expect the least-squares estimators from (30) to be asymptotically efficient and mixed normal, provided that p goes to infinity as a suitable function of n. An elegant explanation of this result using likelihood factorizations is given by Stock and Watson (1993, p. 785).

A final approach may be to parametrize the endogeneity, i.e., formulate a simultaneous model for those variables which may not be considered weakly exogenous. This is similar in spirit to the first approach based on VARs, but by imposing more identifying information the dimensionality problem may be reduced. A development of this approach is left for future research.

D. *An Empirical Modelling Procedure*

The foregoing results suggest an empirical modelling procedure consisting of four different steps. First, the long-run relationship of interest is formulated, and the variables that enter such a relationship are collected. In order to (partially) check the validity of Assumptions 1 and 2, the univariate properties of these variables are analyzed, in particular with respect to the presence of stochastic and deterministic trends and periodic behavior (cf. section II, and Boswijk and Franses, 1994). The choice of the explanatory variables should be such that one is reasonably confident that there are no cointegrating relationships between them. If there were such relationships, then part (*iii*) of Assumption 1 would be violated, and the cointegrating relationship of interest (i.e., between the dependent and explanatory variables) would not be identified.

In the second step, an initial specification of the conditional error correction model is formulated, estimated, and tested. If diagnostics reveal no misspecification, then the $Wald_s$ and *Wald* cointegration test statistics are computed from the estimated ECM, if necessary augmented with periodic trend terms. Comparing these statistics with critical values from tables A1 and A2, respectively, one then tests for partial or full periodic cointegration.

If the hypothesis of no cointegration is rejected, then one can proceed with estimation of and inference on the long-run parameters and adjustment coefficients. In particular, the periodic parameter variation is tested, since this may lead to a more parsimonious model. The asymptotic theory implies that F-statistics on the long-run parameters may only be compared with critical values from F-tables if the explanatory variables are weakly exogenous (Assumption 3). Thus, in step 4, we propose to test for weak exogeneity using a variable addition test for the estimated disequilibrium errors in a model for the explanatory variables. If weak exogeneity is rejected, one has to consider alternative estimators of the long-run parameters, which were discussed in the previous subsection.

IV. A Monte Carlo Experiment

In this section we use a Monte Carlo experiment to investigate the finite sample performance of the tests proposed in section III. The purpose of this experiment is quite limited: therefore, we do not compare our approach to seasonal cointegration, or study the effect of dynamic misspecification on the tests (e.g., by assuming a moving average process with a near-unit root; see Boswijk and Franses, 1992). We use the following data-generating process (DGP) for the bivariate time series $(y_t, z_t)'$:

$$\begin{aligned}\Delta_4 y_t &= \sum_{s=1}^{4} D_{st}\lambda_s(y_{t-4} - \theta_s z_{t-4}) \\ &\quad + \gamma\Delta_4 y_{t-1} + \beta\Delta_4 z_t + \epsilon_t, \\ \Delta_q z_t &= \sum_{s=1}^{4} D_{st}\kappa_s(y_{t-4} - \theta_s z_{t-4}) + \eta_t, \\ &\qquad t = 1, \ldots, n, \quad (31)\end{aligned}$$

where $\{(\epsilon_t, \eta_t)'\}$ are i.i.d. $N(0, I_2)$, and where $q \in \{1, 4\}$ determines the appropriate differencing filter for z_t. In all cases, we take $\beta = 0.5$ and $n \in \{100, 200\}$, which corresponds to either 25 or 50 years of quarterly observations. For the remaining parameters, we consider the following cases:

A: $\lambda_s = \kappa_s = 0, \forall s; \gamma = 0; q = 4$
B: $\lambda_s = \kappa_s = 0, \forall s; \gamma = 0; q = 1$
C: $\lambda_s = \kappa_s = 0, \forall s; \gamma = 0.3; q = 4$
D: $\lambda_s = -0.5, \kappa_s = 0, \theta_s = 1, \forall s;$ $\gamma = 0; q = 4$
E: $\lambda_s = -0.5, \kappa_s = 0.3, \theta_s = 1, \forall s;$ $\gamma = 0; q = 4$
F: $\lambda_s = -0.5, \kappa_s = 0, \forall s;$ $\theta = (0.8, 1.0, 1.2, 1.0)'; \gamma = 0; q = 4$
G: $\lambda = -(0.2, 0.4, 0.6, 0.8)';$ $\kappa_s = 0, \theta_s = 1, \forall s; \gamma = 0; q = 4.$

The results are based on 10,000 replications. In each replication only the first equation of (31) is estimated, augmented with 4 seasonal dummies. From this we compute the $Wald_s$ and *Wald* test statistics, as well as F-statistics for $H_{0\lambda}$, $H_{0\theta}$ and $H_{0\delta}$, see (27). Cases A–C are used to assess the

TABLE 1.—REJECTION FREQUENCIES OF THE $Wald_s$ AND $Wald$ STATISTICS

Case	n	$Wald_1$	$Wald_2$	$Wald_3$	$Wald_4$	$Wald$
			Size			
A	100	0.0531	0.0486	0.0529	0.0516	0.0621
	200	0.0529	0.0484	0.0507	0.0523	0.0563
B	100	0.0535	0.0512	0.0541	0.0524	0.0633
	200	0.0522	0.0500	0.0516	0.0510	0.0598
C	100	0.0505	0.0450	0.0480	0.0471	0.0500
	200	0.0447	0.0419	0.0434	0.0474	0.0409
			Power			
D	100	0.5711	0.5657	0.5674	0.5688	0.9821
	200	0.9636	0.9622	0.9638	0.9641	1.0000
E	100	0.4554	0.4556	0.4564	0.4637	0.8794
	200	0.8916	0.8888	0.8919	0.8896	0.9999
F	100	0.4926	0.5666	0.6710	0.7668	0.9927
	200	0.9324	0.9619	0.9837	0.9945	1.0000
G	100	0.1289	0.3986	0.7231	0.9101	0.9885
	200	0.3405	0.8683	0.9919	0.9998	1.0000

size of the cointegration tests. In case B, z_t is a random walk, which implies that the Brownian motions U_s in Theorem 1 will be dependent. In the previous section it was suggested that the asymptotic distribution is invariant to this dependence. In case C there is short-run dynamics in $\Delta_4 y_t$, which by Theorem 1 should lead to lower significance levels of the $Wald_s$ and $Wald$ test statistics, since condition (i) is not satisfied. The remaining cases D–G are used to study the power of the cointegration tests, and the size and power of the various tests for variation in the cointegration parameters (F) and in the adjustment parameters (G). E is the only case where weak exogeneity is violated, so that we may expect a size distortion of the tests for seasonal parameter variation.

The rejection frequencies at a nominal 5% level are presented in table 1 for the $Wald_s$ and $Wald$ periodic cointegration test statistics, and in table 2 for the periodicity F-test statistics.

Starting with the Wald statistics, we observe that in cases A and B, the rejection frequencies are quite close to the nominal size, whereas in case C, they are somewhat smaller. This is in agreement with Theorem 1, which states that in the absence of short-run dynamics, the asymptotic null distribution of the test statistics is free of nuisance parameters, whereas if $\gamma \neq 0$, the distribution is more concentrated towards zero. For the remaining cases D to G, we observe that the power of the tests increases (not surprisingly) with the sample size n, with the absolute value of the error correction parameter λ_s, and also slightly with the long-run parameter θ_s. Moreover, it can be seen that the power of the joint $Wald$ test is much larger than the individual $Wald_s$ tests. However, this is probably caused by the fact that we consider only fully cointegrated DGP's; for partially cointegrated systems, the joint $Wald$ statistic can be expected to lose power relative to the individual tests.

TABLE 2.—REJECTION FREQUENCIES OF THE PERIODICITY F-STATISTICS

Case	n	$F(H_{0\lambda})$	$F(H_{0\theta})$	$F(H_{0\delta})$
D	100	0.0404	0.0762	0.0652
	200	0.0464	0.0688	0.0626
E	100	0.0398	0.0771	0.0701
	200	0.0475	0.0787	0.0725
F	100	0.0422	0.3525	0.3175
	200	0.0488	0.8407	0.8065
G	100	0.4196	0.0923	0.4367
	200	0.8566	0.0849	0.8378

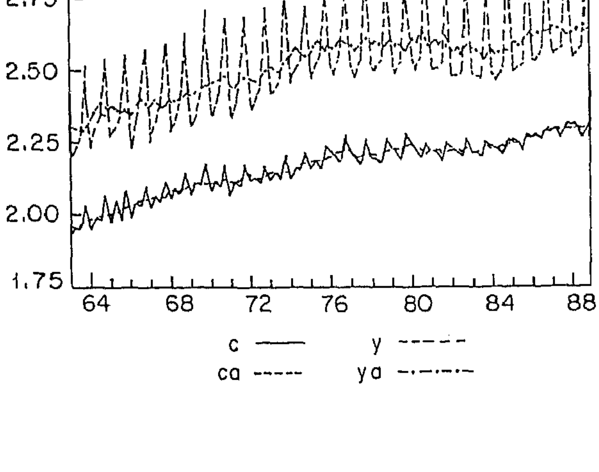
Next, consider the results for the periodicity tests in table 2. We observe that, whereas in cases D and E the actual size of the F-test for constancy of the error correction parameters λ_s is quite close to its nominal value, the tests for constancy of θ_s and for overall non-periodicity suffer from size distortions if $n = 100$. If n gets larger, the rejection frequencies appear to converge to 5% for case D, but the distortions persist in case E. This can be explained by Corollary 1 and the violation of Assumption 3. The power of the tests increases with the sample size as expected.

In summary, this Monte Carlo experiment indicates that the periodic cointegration tests perform quite well in terms of size and power properties. The effect of introducing short-run dynamics ($\gamma \neq 0$) is rather small, and in agreement with Theorem 1. The F-tests for periodic parameter variation have reasonable power properties. The tests that involve the long-run parameters appear to suffer from modest size distortions, which however vanish as the sample size increases, provided that the weak exogeneity assumption holds.

V. An Application: Consumption and Income in Sweden

In Osborn (1988) it is shown that Hall's (1978) version of the life cycle–permanent income hypothesis, augmented with periodic preferences, implies a periodically integrated AR(1) process instead of Hall's random walk process for aggregate non-durables consumption. However, the same paper demonstrates that in the United Kingdom, a lagged income variable contributes significantly to the explanation of consumption, indicating a violation of the assumptions underlying the life cycle–permanent income hypothesis. Given the periodicity in the consumption series, these results may suggest consideration of a periodic error correction model of consumption given income. In this section, we analyze the quarterly unadjusted time series of real per capita non-durables consumption and real per capita disposable income in Sweden, over the period 1963.1–1988.4. The natural logarithms of the data, denoted c_t and y_t, and their seasonally adjusted series, denoted ca_t and ya_t, are displayed in figure 1.

FIGURE 1.—LOG OF SWEDISH REAL PER CAPITA NON-DURABLES CONSUMPTION (c_t) AND DISPOSABLE INCOME (y_t), AND THEIR SEASONALLY ADJUSTED SERIES (ca_t AND ya_t), 1963.1–1988.4



We start with a univariate analysis of the two series, then we analyze a periodic error correction model for c_t given y_t, and finally we repeat this analysis for seasonally adjusted data. The latter investigation is motivated by the conjecture formulated in Franses (1995b) and Ghysels and Hall (1992) that, theoretically, linear seasonal adjustment filters do not entirely remove the periodicity in a time series.

All empirical models presented below are tested for possible misspecification, viz. first- and fourth-order serial correlation, first-order periodic serial correlation (see Franses (1993)), first- and fourth-order ARCH effects, and non-normality. In order to save space, we only report whether the diagnostics indicate any misspecification.

A. *Univariate Time Series Analysis*

To select an appropriate univariate model for consumption and income, we start with a general PAR(6) model, which is subsequently simplified according to the outcome of F-statistics for parameter restrictions and diagnostic checks. For c_t, the F-statistic for periodic parameter variation in the initial model equals 2.525, which is significant at the 5% level. The PAR(6) model may be reduced to the following periodic subset model:

$$c_t = \alpha_s + \varphi_{1s} c_{t-1} + \varphi_{4s} c_{t-4} + \epsilon_t, \quad t = 1, \ldots, n, \quad s = 1, \ldots, 4, \quad (32)$$

denoted PAR(1, 4). None of the diagnostics indicate misspecification of this model, and the null hypothesis of a non-periodic AR process can again be rejected. To test for periodic integration, we rewrite (32) in its VQ representation (cf. section III), and check whether the characteristic equation has a unit root. It can be easily checked that this is the case if $\prod_{s=1}^{4}(1 - \varphi_{4s}) = \prod_{s=1}^{4} \varphi_{1s}$, which can be tested with the Boswijk and Franses (1995) test statistic.

For the estimated model, the roots of the characteristic equation are 1.068, 4.104, and 2.167 $\pm$ 1.407i. Although these are obviously all outside the unit circle, the first root appears to be quite close to unity. The Boswijk-Franses test statistic for a unit root equals -2.543, which should be compared with critical values of Fuller's (1976) $\hat{\tau}_{\mu}$ statistic. If a linear trend term with four periodic coefficients is added to (32), to allow for periodic trend-stationarity under the alternative, the Boswijk-Franses statistic becomes -2.766; this statistic should be compared with the critical values of Fuller's $\hat{\tau}_{\tau}$ statistic. Hence the null hypothesis of a unit root cannot be rejected at the 5% or 10% level, either with or without linear trends included, and the Swedish consumption series appears to be well described by a periodically integrated AR model.

For the disposable income series, we start again with a PAR(6) model. The diagnostics do not indicate any misspecification of this model. However, the periodicity F-statistic equals 0.663, which implies that a non-periodic AR model will suffice to describe y_t. Therefore, we use HEGY's method to test for possible unit roots in this series. We start with an AR(6) model, which can be rewritten as (2), augmented with two lagged differences ($\Delta_4 y_{t-1}$ and $\Delta_4 y_{t-2}$) and 4 seasonal dummies (and possibly a linear trend). If a linear trend is omitted, the HEGY test statistics are $t(\pi = 1) = -1.550$, $t(\pi_2 = 0) = -2.588$, and $F(\pi_3 = \pi_4 = 0) = 5.671$; with a trend added, they become $t(\pi = 1) = -1.491$, $t(\pi_2 = 0) = -2.601$, and $F(\pi_3 = \pi_4 = 0) = 5.553$. Upon comparison with the critical values from HEGY, none of the unit root statistics is significant at the 5% level, whether or not a linear trend term is included; the F-statistics, however, are significant at the 10% level. Rejection of the hypothesis $\pi_3 = \pi_4 = 0$ would imply that the $\Delta_2 = (1 - L^2)$ filter is required to achieve stationarity. However, because the evidence is not very strong, we arrive at an AR(2) model for $\Delta_4 y_t$. Subsequent testing shows that the linear trend term can be deleted, and the four dummies may be replaced by a single intercept. The diagnostics for this final model reveal no misspecification. In summary, the outcome of the univariate analysis is that c_t is periodically integrated, and that y_t is seasonally integrated.

B. *A Conditional Error Correction Model*

At first sight, the results of the univariate analysis leave little room for possible cointegrating relationships between c_t and y_t: periodic integration of c_t implies that all seasons of this series have a common stochastic trend, whereas the quarters of the seasonally integrated y_t have four different trends. Thus, if these results are taken literally, cointegration in at most one quarter seems possible. Notice, however, that there is some indication of rejection of the $F(\pi_3 = \pi_4 = 0)$ test, which indicates that the quarters of y_t can have two common trends (cf. Franses (1994)).

TABLE 3.—PERIODIC COINTEGRATION *Wald* TEST STATISTICS

	4 Dummies	4 Dummies, 4 Linear Trends
$Wald_1$	2.253	12.612
$Wald_2$	8.433	14.786
$Wald_3$	0.744	3.282
$Wald_4$	8.343	8.807
$Wald$	19.280	38.469

Therefore, we now proceed to investigate possible cointegrating relationships in a periodic error correction model, where both c_t and y_t are transformed to (periodic) stationarity by the Δ_4 filter. After some experimenting with different lag lengths, we choose the following initial specification

$$\Delta_4 c_t = \mu_s + \beta\Delta_4 y_t + \gamma\Delta_4 c_{t-1} + \lambda_s(c_{t-4} - \theta_s y_{t-4}) + \epsilon_t, \quad t = 1, \ldots, n, \quad s = 1, \ldots, 4, \tag{33}$$

which does not appear to be misspecified. We shall report on the estimated parameter values only in a simplified model below. Table 3 reports the $Wald_s$ and $Wald$ cointegration statistics in (33), as well as in the same model augmented with four periodic linear trends. The latter is required to obtain an asymptotically similar test in case the series are integrated with drift, see section III.

The results for the model without trends indicate no cointegration at the 5% or 10% level in any of the quarters; the $Wald_2$ and $Wald_4$ statistics would be significant at the (rather generous) 20% level. If the trend terms are added, the results are slightly more promising: the statistics for the first and second quarter are significant at the 10% and 5% level, respectively, and the overall *Wald* statistic is quite close to its 10% critical value. However, the significance of $Wald_1$ is reflected in a highly significant trend term in this quarter, which may be an indication of a missing variable. On the basis of these tests, we tentatively proceed with a model with error correction terms only in the second and fourth quarter:

$$\begin{aligned}\Delta_4 c_t = {} & \underset{(0.003)}{0.008} + \underset{(0.055)}{0.213}\,\Delta_4 y_t + \underset{(0.090)}{0.224}\,\Delta_4 c_{t-1} \\ & - \underset{(0.109)}{0.311}\, D_{2t}\Big(c_{t-4} - \underset{(0.261)}{0.049} - \underset{(0.107)}{0.868}\, y_{t-4}\Big) \\ & - \underset{(0.160)}{0.448}\, D_{4t}\Big(c_{t-4} - \underset{(0.170)}{0.517} - \underset{(0.062)}{0.614}\, y_{t-4}\Big) \\ & + \hat{\varepsilon}_t, \qquad (34)\end{aligned}$$

where $\hat{\sigma} = 0.015$, and no misspecification is indicated by the diagnostics. The $F(1,90)$ statistic for equality of error correction coefficients equals 0.507, the statistic for $\theta_2 = \theta_4$ is $F(1,90) = 2.663$, and the joint test statistic equals $F(2,90) = 2.253$. Hence we cannot reject the null hypothesis of equal adjustment parameters and long-run elasticities in the second and fourth quarter. An F-statistic for equality of all parameters of the error correction terms (including the long-run intercept) yields $F(3,90) = 3.938$, so that this hypothesis is rejected. Hence our final model is

$$\begin{aligned}\Delta_4 c_t = {} & \underset{(0.002)}{0.008} + \underset{(0.055)}{0.196}\,\Delta_4 y_t + \underset{(0.091)}{0.238}\,\Delta_4 c_{t-1} \\ & - \underset{(0.080)}{0.260}\Big\{ D_{2t}\Big(c_{t-4} - \underset{(0.210)}{0.489} - \underset{(0.086)}{0.687}\, y_{t-4}\Big) \\ & \qquad + D_{4t}\Big(c_{t-4} - \underset{(0.236)}{0.313} - 0.687 y_{t-4}\Big)\Big\} \\ & + \hat{\varepsilon}_t. \qquad (35)\end{aligned}$$

To test for weak exogeneity of y_t for the long-run parameters, we extract the error correction variables from (34) and add these to the univariate AR(2) model for $\Delta_4 y_t$. An F-test statistic for their joint significance is $F(2,93) = 2.985$; because this statistic is only slightly smaller than the 5% critical value, the validity of the weak exogeneity assumption may be doubtful. This implies that more efficient estimation of the cointegration parameters is possible in a joint model of c_t and y_t.

Because the series are in natural logarithms, the model implies the following two long-run targets for the original series (indicated by capitals):

$$\begin{aligned} C &= 1.631 Y^{0.687} \quad (s = 2), \\ C &= 1.368 Y^{0.687} \quad (s = 4). \qquad (36)\end{aligned}$$

Thus, although the long-run elasticity is equal in both quarters, a change in disposable income will have a smaller long-run effect on non-durables consumption in the fourth quarter. The speed of adjustment towards these targets is equal in these quarters; in the first and third quarter, the speed of adjustment is restricted to zero.

C. *Seasonally Adjusted Time Series*

A natural question that can be raised concerns the effects of analyzing seasonally adjusted data instead of the original data. In the literature there is some indication that seasonal adjustment filters may affect unit root inference (see Ghysels and Perron (1993) *inter alia*) in the sense that test statistics are biased towards non-rejection. Further, there is evidence that periodicity in the autoregressive model for the original data is not completely removed by seasonal adjustment (see Franses (1995b)) which is due to the fact that linear filters handle the observations in the same way throughout the year.

The data on Swedish consumption and income, given in the appendix, are seasonally adjusted along the lines described in Ooms (1994, section 4.5). It consists of applying a linearization of the Census-X11 filter to the time series, extended with 28 post-sample forecasts and 28 pre-sample backcasts. These forecasts are constructed using a nonperiodic and unrestricted (i.e., with no unit roots imposed) AR model, containing seasonal dummies and a linear trend. Of course, when there is significant periodicity in the data, the non-periodic autoregressive order required to obtain white noise errors can be very large (see Osborn (1991)). The adjusted data, denoted by ca_t and ya_t, are displayed in figure 1.

We start again with a univariate analysis of the two series. For the sake of completeness, we start again with periodic AR models. For ca_t, a PAR(3) model does not suffer from misspecification according to the diagnostics. The test statistic for the null hypothesis of no periodicity equals 0.587; hence periodicity cannot be detected any more in the adjusted time series. The augmented Dickey-Fuller statistics for the emerging AR(3) model equal $\hat{\tau}_\mu = -2.284$ and $\hat{\tau}_\tau = -2.324$, so that we end up with an AR(2) model for $\Delta_1 ca_t$. Note that seasonally adjusting a $PI(1)$ series here leads to an $I(1)$ series, as is also found in the simulations reported in Franses (1995b). For ya_t, we start with a periodic AR(4) model. Here the periodicity F-test is 0.411, and in the resulting AR(4) model, the ADF statistics are $\hat{\tau}_\mu = -1.531$ and

$\hat{\tau}_\tau = -0.078$, so that we arrive at an AR(3) model for $\Delta_1 ya_t$.

Applying Boswijk's (1992) approach to test for cointegration in a conditional error correction model for the adjusted series, we start with a nonperiodic version of (33):

$$\Delta_1 ca_t = \underset{(0.030)}{0.036} - \underset{(0.090)}{0.375}\,\Delta_1 ca_{t-1} + \underset{(0.051)}{0.164}\,\Delta_1 ya_t - \underset{(0.055)}{0.111}\left(ca_{t-1} - \underset{(0.128)}{0.746}\,ya_{t-1}\right) + \hat{\varepsilon}_t, \tag{37}$$

with $\hat{\sigma} = 0.011$. A variable addition test for $\{D_{st}, D_{st}ca_{t-1}, D_{st}ya_{t-1},\ s = 1, 2, 3\}$, equals $F(9, 90) = 0.414$, so that there is no indication of periodic effects in the long-run and adjustment parameters. The Wald statistic for cointegration in (37) equals 5.854; this should be compared with critical values from table A1(b), so that the null hypothesis of no cointegration cannot be rejected. Although there is slightly more evidence of cointegration if a linear trend term is added to the right-hand side of (37), it is still not significant: the Wald statistic equals 10.792.

In summary, the analysis with seasonally adjusted time series shows that linear adjustment filters may affect unit root and cointegration inference indeed, in the sense that the (already rather weak) cointegrating relationships in the original data are obscured. Moreover, we find that in this case the underlying periodicity is removed. Hence, although seasonal adjustment may be useful for some purposes, it appears to generate misleading inferences for possibly periodically cointegrated time series.

VI. Concluding Remarks

In this paper we have analyzed periodic cointegration and error correction. Periodic error correction models appear to be quite useful for the analysis of non-stationary seasonal data, because they allow, e.g., preferences, constraints and adjustment costs, to vary over the seasons. This is illustrated by the empirical model for Swedish consumption, where adjustment towards equilibrium only takes place in the second and fourth quarter. Such behavior may have important implications, not only for modelling and forecasting seasonal economic series, but also for policy analysis.

The modelling procedure that we have proposed comprises a class of Wald tests for cointegration, estimators of cointegration parameters and adjustment coefficients, and tests for hypotheses on these parameters, in particular with respect to their periodic variation. For the estimators and hypothesis tests, the current approach requires some exogeneity assumption. To relax this assumption, the possible simultaneous error correcting behavior has to be accommodated, which is the subject of current research.

A possible disadvantage of periodic models in general is that they require a rather large number of parameters. For this reason we have proposed to analyze the conditional error correction model rather than a general VAR model for the VQ process of the vector time series. Further, in our model, the error correction term appears only with a lag of four, and the short-run dynamics are restricted to be constant over the seasons, which reduces the parameter dimensionality problem. Moreover, we have stressed testing for periodicity. With these tests, one can check whether the parameter variation is significant, and if not, one may want to consider a more parsimonious model with constant parameters.

APPENDIX 1

Proofs

This appendix contains the proofs of Theorems 1 and 2, and of Corollary 1. We shall make use of the following lemma:

LEMMA A1: *Let y_t be generated by (7), augmented by four seasonal intercepts μ_{0s}, $s = 1, \ldots, 4$, and let Assumptions 1 and 2 hold. Define $\{X_T = (Y_T', Z_T')'\}$, the VQ process of $\{x_t\}$. Moreover, let $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_4)$, $\Theta = [\Theta_1 \vdots \cdots \vdots \Theta_k]'$ with $\Theta_j = \mathrm{diag}(\theta_{j1}, \ldots, \theta_{j4})$; if $\lambda_s = 0$, then $\theta_{js} = 0$. Let Γ_i and B_i, $i = 1, \ldots, P = [(p-1)/4] + 1$, denote matrices of order 4×4 and $4 \times 4k$, respectively, such that*

$$\Gamma_0 \Delta Y_T = \mu_0 + B_0 \Delta Z_T + \sum_{i=1}^{P} (\Gamma_i \Delta Y_{T-i} + B_i \Delta Z_{T-i}) + \Lambda Y_{T-1} - \Lambda\Theta' Z_{T-1} + U_{0T}, \tag{A1}$$

$T = 1, \ldots, N$, is the VQ representation of (7). Finally, let $\Gamma = \Gamma_0 - \Sigma_{i=1}^{P}\Gamma_i$ and $B = -\Sigma_{i=0}^{P} B_i$. Then $\{X_T\}$ has the following representation:

$$X_T = X_0 + C\mu T + C\sum_{j=1}^{T} U_j + E_T, \tag{A2}$$

where $E_T = (E_{1T}', 0)'$ with $\{E_{1T}\}$ a 4×1 mean-zero stationary *vector process, and*

$$C = \begin{bmatrix} C_{11} & C_{12} \\ 0 & I_{4k} \end{bmatrix}. \tag{A3}$$

Let r denote the number of non-zero λ_s parameters. If $r = 4$, then $C_{11} = 0$ and $C_{12} = \Theta'$; if $r < 4$, then $C_{11} = I_0(I_0'\Gamma I_0)^{-1}I_0'$ and $C_{12} = -I_0(I_0'\Gamma I_0)^{-1}I_0'(\Gamma + B) + \Theta'$, where I_0 is a $(4-r) \times 4$ matrix containing those columns of the identity matrix I_4 that correspond to the zero λ_s parameters.

Proof of Lemma A1: Define the $4m \times 4m$ matrices

$$\Phi_0 = \begin{bmatrix} \Gamma_0 & -B_0 \\ 0 & I_{4k} \end{bmatrix}, \quad \Phi_i = \begin{bmatrix} \Gamma_i & B_i \\ 0 & I_{4k} \end{bmatrix}, \quad i = 1, \ldots, P,$$

$$\Phi = \Phi_0 - \sum_{i=1}^{P} \Phi_i = \begin{bmatrix} \Gamma & B \\ 0 & I_{4k} \end{bmatrix}. \tag{A4}$$

Next, let the $4m \times r$ matrices α and β be defined such that α contains the non-zero columns of $[\Lambda \vdots 0]'$ and β' contains the corresponding rows of $[I_4 \vdots -\Theta']$; if $r = 0$, then both α and β are void. The relationship between X_T and U_T can now be derived from the multivariate model

$$\Phi(L)\Delta X_T = \left(\Phi_0 - \sum_{i=1}^{P} \Phi_i L^i\right)\Delta X_T = \mu + \alpha\beta' X_{T-1} + U_T, \quad T = 1, \ldots, N. \tag{A5}$$

This is a VAR model in ECM format, although the errors U_T are merely stationary, not white noise. Assumption 2 now guarantees that Johansen's version of the Granger representation theorem (cf. Johansen (1991), Theorem 4.1) applies. Let $\alpha_\perp$ and $\beta_\perp$ denote $(4m - r) \times 4m$ matrices of rank $(4m - r)$ such that $\alpha'\alpha_\perp = 0$ and $\beta'\beta_\perp = 0$. If $r = 4$, then these are given by $\alpha_\perp = [0 \vdots I_{4k}]'$ and $\beta_\perp = [\Theta \vdots I_{4k}]'$. If $r < 4$, then

$$\alpha_\perp = \begin{bmatrix} I_0 & 0 \\ 0 & I_{4k} \end{bmatrix}, \quad \beta_\perp = \begin{bmatrix} I_0 & \Theta' \\ 0 & I_{4k} \end{bmatrix}. \tag{A6}$$

Johansen's (1991) Theorem 4.1 states that (A5) implies (A2), with E_T a stationary moving average of U_T, and $C = \beta_\perp(\alpha_\perp' \Phi\beta_\perp)^{-1}\alpha_\perp'$. Substitution of the definitions of $\alpha_\perp$, $\beta_\perp$ and Φ yields (A3). Finally, from the definition of X_T, U_T and C it is obvious that the part of E_T corresponding to Z_T vanishes, i.e., $E_T = (E_{1T}', 0)'$. □

Proof of Theorem 1: Let X_s denote the $n \times m$ matrix of observations on $D_{st}x_{t-4}$, let $X = [X_1 \vdots \cdots \vdots X_4]$, and let X_{-s} denote the $n \times 3m$ matrix containing X_r, $r \neq s$. Next, let $W = [w_1, \ldots, w_n]'$ and $\epsilon = (\epsilon_1, \ldots, \epsilon_n)'$, and define the familiar projection matrix $M(X) = I_n - X(X'X)^{-1}X'$. For the $Wald_s$ statistic, we have under H_{0s},

$$\hat{\delta}_s = \left[X_s' M(W \vdots X_{-s}) X_s\right]^{-1} X_s' M(W \vdots X_{-s})\epsilon,$$
$$\hat{V}\left[\hat{\delta}_s\right] = \hat{\sigma}^2\left[X_s' M(W \vdots X_{-s}) X_s\right]^{-1}. \tag{A7}$$

Because $X_s' X_{-s} = 0$, we have

$$X_s' M(W \vdots X_{-s}) X_s = X_s' X_s - X_s' W\left[W' M(X_{-s}) W\right]^{-1} W' X_s,$$
$$X_s' M(W \vdots X_{-s})\epsilon = X_s'\epsilon - X_s' W\left[W' M(X_{-s}) W\right]^{-1} W' M(X_{-s})\epsilon. \tag{A8}$$

Now $\mu = 0$ implies that w_t contains mean-zero, stationary components, uncorrelated with ϵ_t. Moreover, X_{-s} contains $I(1)$ components, possibly with some stationary linear combinations (if some $\lambda_q < 0$, $q \neq s$). Thus the convergence rates of Park and Phillips' (1989) Lemma 2.1 imply

$$\left(N^{-1}X_s'W, \quad N^{-1}W'M(X_{-s})W, \quad N^{-1/2}W'M(X_{-s})\epsilon\right) = O_p(1). \tag{A9}$$

Therefore, by Assumption 1, Lemma A1 and the continuous mapping theorem, we have

$$N^{-2}X_s' M(W \vdots X_{-s}) X_s = N^{-2}X_s' X_s + o_p(1) \Rightarrow \tilde{C}_s' \int BB'\tilde{C}_s,$$
$$N^{-1}X_s' M(W \vdots X_{-s})\epsilon = N^{-1}X_s'\epsilon + o_p(1) \Rightarrow \tilde{C}_s' \int B\, dB_{0s}. \tag{A10}$$

Here $\tilde{C}_s$, $s = 1, \ldots, 4$, are $4m \times m$ matrices such that $\tilde{C} = [\tilde{C}_1 \vdots \cdots \vdots \tilde{C}_4]' = K_{4m}C$. Notice that

$$\tilde{C}_s' = \begin{bmatrix} C_{11,s} & C_{12,s} \\ 0 & I_k \otimes e_s' \end{bmatrix}, \tag{A11}$$

where $C_{11,s}$ and $C_{12,s}$ are the s^{th} rows of C_{11} and C_{12}, respectively, whereas e_s is the s^{th} 4×1 unit vector. The orders of convergence imply that $\hat{\delta}_s$ is super-consistent; consistency of $\hat{\sigma}^2$ shall not be proved here explicitly but follows from the general results by Park and Phillips (1988, 1989). Lemma 1 implies that $C_{11,s} \neq 0$; together with Assumption 1 *(iii)*, this implies that $\tilde{C}_s'\Omega\tilde{C}_s > 0$. Define $U_s(r) = (\tilde{C}_s'\Omega\tilde{C}_s)^{-1/2}\tilde{C}_s' B(r)$, an m-vector standard Brownian motion process, and $W_s(r) = B_{0s}(r)/\sigma$, a scalar standard Brownian motion, then we have

$$Wald_s \Rightarrow \frac{1}{\sigma^2}\int dB_{0s} B'\tilde{C}_s\left[\tilde{C}_s'\int BB'\tilde{C}_s\right]^{-1}\tilde{C}_s'\int B\, dB_{0s} = h(W_s, U_s). \tag{A12}$$

The results for the case where seasonal dummies and trends are included can be proved analogously.

The correlation between W_s and U_s equals 1 if W_s is a linear combination of U_s. This is the case if and only if for some m-vector $(c_{1s}, c_{2s}')'$, $C_{11,s} = c_{1s}e_s'$ and $C_{12,s} = c_{2s} \otimes e_s'$; then $C_s'B$ is a non-singular linear transformation of B_s. From the definition of C_{11} and C_{12} it can be checked this is the case, either if $\Gamma = \gamma I_4$ and $B = (b' \otimes I_4)$ for some scalar γ and k-vector b, or if $I_0 = e_s$, i.e., if λ_s is the only zero error correction parameter.

For the joint *Wald* statistic, we have under H_0,

$$Wald = \frac{1}{\hat{\sigma}^2}\epsilon' M(W) X\left[X' M(W) X\right]^{-1} X' M(W)\epsilon. \tag{A13}$$

Because $N^{-1}X'M(W)\epsilon = N^{-1}\operatorname{vec}\{X_s'\epsilon\}_1^4 + o_p(1)$ and $N^{-2}X'M(W)X = N^{-2}\operatorname{diag}\{X_s'X_s\}_1^4 + o_p(1)$, it follows from the derivations above that under H_0, $Wald = \sum_{s=1}^4 Wald_s + o_p(1)$, which leads to the required result. Independence of U_s and U_q under the stated conditions follows from the fact that under those conditions, U_s is a non-singular linear transformationof B_s.

To prove consistency, we note that except for the degrees of freedom correction, the $Wald_s$ statistic is equal to the squared

t-statistic for $\lambda_s = 0$ in

$$\Delta_4 y_t = \lambda_s \hat{v}_{s,t-4} + \sum_{r \neq s}^{4} \delta_r' D_{rt} x_{t-4} + \pi' w_t + \epsilon_t, \quad t = 1, \ldots, n, \quad \text{(A14)}$$

where $\hat{v}_{st} = D_{st}(y_t - \hat{\theta}_s' z_t)$. This can be checked by expressing both statistics in RSS_{0s} and RSS_1, cf. (13). In Theorem 2 it is proved that if $\lambda_s < 0$, $\hat{\theta}_s$ is super-consistent, which implies that the t-ratio in (A14) is asymptotically equivalent to the t-ratio in the same equation, but with $\hat{v}_{st}$ replaced by v_{st}; since this is a stationary process and $\lambda_s < 0$, the t-ratio will diverge to minus infinity as $n \to \infty$, and hence $Wald_s \to \infty$. From this, divergence of $Wald$ can be deduced as well. □

Proof of Theorem 2: Express the least-squares estimator $\hat{\delta}$ as

$$\begin{aligned}\hat{\delta} &= \delta + [X'M(W)X]^{-1} X'M(W)\epsilon \\ &= \delta + D^{-1}\Upsilon_N^{-1}[\Upsilon_N^{-1}(D^{-1})'X'M(W)XD^{-1}\Upsilon_N^{-1}]^{-1} \\ &\quad \times \Upsilon_N^{-1}(D^{-1})'X'M(W)\epsilon, \end{aligned} \quad \text{(A15)}$$

where

$$\begin{aligned} D^{-1} &= \frac{\partial \delta}{\partial \varphi'} \\ &= [\mathrm{diag}(\lambda_1^{-1}\delta_1, \ldots, \lambda_4^{-1}\delta_4) \vdots \mathrm{diag}(\lambda_1 F, \ldots, \lambda_4 F)], \\ F &= \begin{bmatrix} 0 \\ I_k \end{bmatrix}. \end{aligned} \quad \text{(A16)}$$

Hence $XD^{-1} = [V \vdots \lambda_1 Z_1 \vdots \cdots \vdots \lambda_4 Z_4]$, where V is a matrix of observations on $v_{s,t-4} = D_{st}(y_{t-4} - \theta_s' z_{t-4})$, of order $n \times 4$, and Z_s is an $n \times k$ matrix of observations on $D_{st} z_{t-4}$, $s = 1, \ldots, 4$. Notice that $V'V$ is a diagonal matrix, and similarly $Z_s' Z_r = 0$, $s \neq r$. Assumption 1 and the continuous mapping theorem imply that

$$\begin{aligned} N^{-2} Z_s' M(W) Z_s &= N^{-2} Z_s' Z_s + o_p(1) \Rightarrow \int B_{1s} B_{1s}', \\ N^{-1} Z_s' M(W)\epsilon &= N^{-1} Z_s'\epsilon + o_p(1) \Rightarrow \int B_{1s}\, dB_{0s}, \\ &\qquad s = 1, \ldots, 4, \end{aligned} \quad \text{(A17)}$$

where the $o_p(1)$ terms follow the convergence rates of Lemma 2.1 of Park and Phillips (1989), and the fact that w_t and ϵ_t are mean-zero, mutually uncorrelated stationary processes. Define $\Sigma_v \equiv \mathrm{Plim}\, N^{-1} V'M(W)V$, and notice that although $V'V$ is diagonal, the correction with respect to W implies that Σ_v is not necessarily diagonal. By Assumption 1 (*ii*), the fact that both V and W contain stationary processes, a martingale difference central limit theorem implies that $N^{-1/2} V'M(W)\epsilon \Rightarrow N(0, \sigma^2 \Sigma_v)$. Finally, by an appeal to Park and Phillips' (1989) Lemma 2.1 it can be shown that $N^{-3/2} Z_s' M(W) V = o_p(1)$, $s = 1, \ldots, 4$. Summarizing these results, we have

$$\Upsilon_N^{-1}(D^{-1})'X'M(W)XD^{-1}\Upsilon_N^{-1} \Rightarrow \begin{bmatrix} \Sigma_v & 0 \\ 0 & \mathrm{diag}\left\{\lambda_s^2 \int B_{1s} B_{1s}'\right\} \end{bmatrix}, \quad \text{(A18)}$$

$$\Upsilon_N^{-1}(D^{-1})'X'M(W)\epsilon \Rightarrow \begin{bmatrix} N(0, \sigma^2 \Sigma_v) \\ \mathrm{vec}\left\{\int \lambda_s B_{1s}\, dB_{0s}\right\} \end{bmatrix}. \quad \text{(A19)}$$

Since $D^{-1}\Upsilon_N^{-1} \to 0$, (A15), (A18) and (A19) together imply that $\hat{\delta}$ is consistent. Define $\bar{D}$ as D in (21)–(22), with J_s replaced by $\bar{J}_s = -\hat{\lambda}_s^{-1}[\theta_s \vdots I_k]$. Notice that $\bar{J}_s(\hat{\delta}_s - \delta_s) = (\hat{\theta}_s - \theta_s)$, so that $\Upsilon_N(\hat{\varphi} - \varphi) = \Upsilon_N \bar{D}(\hat{\delta} - \delta)$. Because $\mathrm{Plim}\, \Upsilon_N \bar{D} D^{-1} \Upsilon_N^{-1} = I_{4m}$, we obtain (23). Express the standardized estimated covariance matrix of $\hat{\varphi}$ as

$$\begin{aligned} \Upsilon \hat{V}p[\hat{\varphi}]\Upsilon_N &= \hat{\sigma}^2 \Upsilon_N \hat{D} D^{-1} \Upsilon_N^{-1} \\ &\quad \times [\Upsilon_N^{-1}(D^{-1})'X'M(W)XD^{-1}\Upsilon_N^{-1}]^{-1} \\ &\quad \times \Upsilon_N^{-1}(D^{-1})'\hat{D}'\Upsilon_N. \end{aligned} \quad \text{(A20)}$$

Because $\mathrm{Plim}\, \Upsilon_N \hat{D} D^{-1} \Upsilon_N^{-1} = I_{4m}$, and because consistency of $\hat{\delta}$ implies that $\mathrm{Plim}\, \hat{\sigma}^2 = \sigma^2$, (A20) together with (A18) implies (24).

If and only if B_{1s} and B_{0s} are independent, we have that

$$\left[\int B_{1s} B_{1s}'\right]^{-1} \int B_{1s}\, dB_{0s}/\lambda_s | B_{1s} \sim N\left(0, \sigma^2 \left[\lambda_s^2 \int B_{1s} B_{1s}'\right]^{-1}\right), \quad s = 1, \ldots, 4. \quad \text{(A21)}$$

Hence, only if Assumption 3 holds, then (A21) holds for all s, which gives the required result. □

Proof of Corollary 1: Express $R\hat{\varphi} - r$ under H_0 as $R(\hat{\varphi} - \varphi) = R_1(\hat{\lambda} - \lambda) + R_2(\hat{\theta} - \theta)$. Consider first the case where rank $(R_1) = h$. Then we have, under H_0,

$$\begin{aligned} \sqrt{N} R_1(\hat{\lambda} - \lambda) + \sqrt{N} R_2(\hat{\theta} - \theta) &= \sqrt{N} R_1(\hat{\lambda} - \lambda) \\ &\quad + o_p(1) \Rightarrow N(0, \sigma^2 R_1 \Sigma_v^{-1} R_1'), \end{aligned} \quad \text{(A22)}$$

and similarly, because of (24) and the definition of Υ_N,

$$NR\hat{V}[\hat{\varphi}]R' = NR_1 \hat{V}[\hat{\lambda}]R_1' + o_p(1) \xrightarrow{P} \sigma^2 R_1 \Sigma_v R_1'. \quad \text{(A23)}$$

Since $\mathrm{rank}(R_1) = h$ implies that the limit in (A23) is invertible, we obtain $hF \Rightarrow \chi^2(h)$, whether or not Assumption 3 holds.

If $\mathrm{rank}(R_1) = q < h$, then (A22) and (A23) still hold, but because the asymptotic covariance matrix in (A22) is no longer invertible, we cannot deduce a χ^2 null distribution from the asymptotics of $\hat{\lambda}$ only. Define the $h \times h$ orthogonal matrix $H = [H_1 \vdots H_2]$, where H_1 is an $h \times q$ matrix such that $H_1' R_1$ is of full row rank q, and where H_2 is of order $h \times (h - q)$ and satisfies $H_2' R_1 = 0$; if $R_1 = 0$, then H_1 is void and $H_2 = I_h$. Notice that a full row rank of R implies that $\mathrm{rank}(H_2' R_2) = h - q$. Next, define the $h \times h$ scaling matrix $\Upsilon_N^* = \mathrm{diag}(\sqrt{N} \cdot I_q, N \cdot I_{h-q})$, or $\Upsilon_N^* = NI_h$ if $R_1 = 0$. The F-test statistic can be expressed under H_0 as

$$\begin{aligned} F &= \frac{1}{h}[\Upsilon_N^* H'R(\hat{\varphi} - \varphi)]'[\Upsilon_N^* H'R\hat{V}[\hat{\varphi}]R'H\Upsilon_N^*]^{-1} \\ &\quad \times [\Upsilon_N^* H'R(\hat{\varphi} - \varphi)]. \end{aligned} \quad \text{(A24)}$$

Now the orders of convergence in Theorem 2 imply

$$\Upsilon_N^* H'R(\hat{\varphi} - \varphi) = \begin{bmatrix} \sqrt{N} H_1' R_1(\hat{\lambda} - \lambda) + o_p(1) \\ NH_2' R_2(\hat{\theta} - \theta) \end{bmatrix}, \quad \text{(A25)}$$

and

$$\Upsilon_N^* H'\hat{V}[\hat{\varphi}]R'H\Upsilon_N^* = \begin{bmatrix} NH_1' R_1 \hat{V}[\hat{\lambda}]R_1' H_1 + o_p(1) & o_p(1) \\ o_p(1) & N^2 H_2' R_2 \hat{V}[\hat{\theta}]R_2' H_2 \end{bmatrix}. \quad \text{(A26)}$$

Thus asymptotically, the F-statistic equals the sum of a statistic for $H_1'R_1\lambda = H_1'r_1$, and a statistic for $H_2'R_2\theta = H_2'r_2$, where $r = (r_1', r_2')'$ in an obvious partition. The first term will always have an asymptotic $\chi^2(q)$ distribution under the null, but the second term requires mixed normality of $\hat{\theta}$, and hence Assumption 3, to have an asymptotic $\chi^2(h - q)$ distribution under the null (independent of the first term). □

APPENDIX 2

Critical Values

TABLE A1.—ASYMPTOTIC CRITICAL VALUES FOR THE *Wald*$_s$ TEST

k	20%	10%	5%	2.5%	1%
		(a) No Constant or Trend			
1	4.80	6.48	8.10	9.66	11.60
2	7.40	9.38	11.18	12.99	15.12
3	9.87	12.10	14.20	16.09	18.64
4	12.21	14.72	16.97	19.08	21.72
5	14.55	17.22	19.72	21.98	24.90
		(b) Constant, no Trend			
1	7.49	9.50	11.36	13.10	15.25
2	9.92	12.18	14.24	16.17	18.64
3	12.29	14.79	16.99	19.09	21.81
4	14.63	17.29	19.74	21.95	24.86
5	16.86	19.82	22.33	24.74	27.82
		(c) Constant and Trend			
1	10.13	12.38	14.39	16.33	18.71
2	12.45	14.89	17.11	19.23	21.78
3	14.78	17.39	19.78	22.00	24.84
4	17.03	19.86	22.43	24.78	27.89
5	19.25	22.31	24.95	27.48	30.61

Note: The quantiles are obtained via Monte Carlo simulation with 50,000 replications, where Brownian motions are approximated by Gaussian random walks of 500 observations; k denotes the number of exogeneous variables.

TABLE A2.—ASYMPTOTIC CRITICAL VALUES FOR THE *Wald* TEST

k	20%	10%	5%	2.5%	1%
		(a) No Constant or Trend			
1	16.17	19.09	21.65	24.00	26.99
2	25.26	28.73	31.75	34.60	37.88
3	34.02	38.03	41.50	44.73	48.79
4	42.77	47.20	51.13	54.74	58.71
5	51.35	56.15	60.41	64.21	68.41
		(b) Constant, No Trend			
1	25.34	28.75	31.82	34.58	37.97
2	34.13	38.07	41.51	44.74	48.61
3	42.85	47.22	51.06	54.56	58.88
4	51.29	56.22	60.45	64.13	68.80
5	59.78	64.99	69.42	73.35	78.15
		(c) Constant and Trend			
1	35.00	38.97	42.49	45.89	49.43
2	43.50	47.92	51.73	55.21	59.25
3	51.93	56.72	60.78	64.39	68.82
4	60.21	65.48	69.87	73.68	78.43
5	68.51	74.02	78.53	82.85	88.05

Note: See table A1.

REFERENCES

Birchenhall, Chris R., R. C. Bladen-Hovell, Alice P. L. Chui, Denise R. Osborn and Jeremy P. Smith, "A Seasonal Model of Consumption," *Economic Journal* 99 (1989), 837–843.

Boswijk, H. Peter, *Cointegration, Identification and Exogeneity. Inference in Structural Error Correction Models* (Amsterdam: Thesis/Tinbergen Institute, 1992).

_____, "Testing for an Unstable Root in Conditional and Structural Error Correction Models," *Journal of Econometrics* (1994), 37–60.

Boswijk, H. Peter and Philip Hans Franses, "Dynamic Specification and Cointegration," *Oxford Bulletin of Economics and Statistics* 54 (1992), 369–381.

_____, "Unit Roots in Periodic Autoregressions," Tinbergen Institute Discussion Paper TI 94-4 (1994).

_____, "Testing for Periodic Integration," *Economics Letters* 48 (1995), 241–248.

Box, George E. P. and Gwilym M. Jenkins, *Time Series Analysis, Forecasting and Control* (San Francisco: Holden-Day, 1976).

Dolado, Juan J., "A Note on Weak Exogeneity in VAR Cointegrated Models," *Economics Letters* 38 (1992), 139–143.

Engle, Robert F., and Clive W. J. Granger, "Cointegration and Error Correction: Representation, Estimation and Testing," *Econometrica* 55 (1987), 251–276.

Engle, Robert F., Clive W. J. Granger, Sven Hylleberg, and Hahn S. Lee, "Seasonal Cointegration: The Japanese Consumption Function," *Journal of Econometrics* 55 (1993), 275–298.

Engle, Robert F., David F. Hendry, and Jean-François Richard, "Exogeneity," *Econometrica* 51 (1983), 277–304.

Franses, Philip Hans, "Periodically Integrated Subset Autoregressions for Dutch Industrial Production and Money Stock," *Journal of Forecasting* 12 (1993), 601–613.

_____, "A Multivariate Approach to Modelling Univariate Seasonal Time Series," *Journal of Econometrics* (1994) 133–151.

_____, "A Vector of Quarters Representation for Bivariate Time Series," *Econometric Reviews* 14 (1995a), 55–63.

_____, "The Effects of Seasonally Adjusting a Periodic Autoregressive Process," *Computational Statistics and Data Analysis* 19 (1995b), 683–704.

Franses, Philip Hans, and Teun Kloek, "A Periodic Cointegration Model of Consumption in Austria and Japan," Econometric Institute Report 9172, Erasmus University Rotterdam (1991).

Fuller, Wayne A., *Introduction to Statistical Time Series* (New York: John Wiley, 1976).

Ghysels, Eric, and Alistair Hall, "Testing for Periodicity in Some Linear Macroeconomic Models," Discussion paper, C.R.D.E., University of Montreal (1992).

Ghysels, Eric, and Pierre Perron, "The Effect of Seasonal Adjustment Filters on Tests for a Unit Root," *Journal of Econometrics* 55 (1993), 57–98.

Hall, Robert E., "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis," *Journal of Political Economy* 86 (1978), 971–987.

Hylleberg, Svend, Robert F. Engle, Clive W. J. Granger and Byung Sam Yoo, "Seasonal Integration and Cointegration," *Journal of Econometrics* 44 (1990), 215–238.

Johansen, Søren, "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," *Econometrica* 59 (1991), 1551–1580.

_____, "Cointegration in Partial Systems and the Efficiency of Single-Equation Analysis," *Journal of Econometrics* 52 (1992), 389–402.

Magnus, Jan R., and Heinz Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics* (Chichester: John Wiley, 1988).

Ooms, Marius, *Empirical Vector Autoregressive Modelling*, Lecture Notes in Economics and Mathematical Systems No. 407 (Berlin: Springer-Verlag, 1994).

Osborn, Denise R., "Seasonality and Habit Persistence in a Life-Cycle Model of Consumption," *Journal of Applied Econometrics* 3 (1988), 255–266.

_____, "A Survey of Seasonality in UK Macroeconomic Variables," *International Journal of Forecasting* 6 (1990), 327–336.

_____, "The Implications of Periodically Varying Coefficients for Seasonal Time-Series Processes," *Journal of Econometrics* 48 (1991), 373–384.

_____, "Discussion of 'Seasonal Cointegration'," *Journal of Econometrics* 55 (1993), 299–303.

Osborn, Denise R., Alice P. L. Chui, Jeremy P. Smith and Chris R. Birchenhall, "Seasonality and the Order of Integration for Consumption," *Oxford Bulletin of Economics and Statistics* 50 (1988), 361–377.

Park, Joon Y., and Peter C. B. Phillips, "Statistical Inference in Regressions with Integrated Processes: Part I," *Econometric Theory* 4 (1988), 468–497.

_____, "Statistical Inference in Regressions with Integrated Processes: Part II," *Econometric Theory* 5 (1989), 95–131.

Phillips, Peter C. B., and Bruce E. Hansen, "Statistical Inference in Instrumental Variable Regression with $I(1)$ Processes," *Review of Economic Studies* 57 (1990), 99–125.

Sims, Christopher A., James H. Stock, and Mark W. Watson, "Inference in Linear Time Series Models with Some Unit Roots," *Econometrica* 58 (1990), 113–144.

Stock, James H., "Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors," *Econometrica* 55 (1987), 1035–1056.

Stock, James H., and Mark W. Watson, "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems," *Econometrica* 61 (1993), 783–820.

Tiao, George, and M. R. Grupe, "Hidden Periodic Autoregressive-Moving Average Models in Time Series Data," *Biometrika* 67 (1980), 365–373.