The applied perspective for seasonal cointegration testing

André Luís Rossi de Oliveira*
Paulo Picchetti[†]

RESUMO

Enquanto a literatura sobre cointegração lida exclusivamente com o caso de cointegração no longo prazo, ou na freqüência zero, entre séries em um vetor de variáveis econômicas, pode ser que raízes unitárias estejam também presentes nas freqüências sazonais, de forma que o conceito de cointegração pode ser extendido para o caso de cointegração sazonal. Neste artigo, fazemos uma resenha dos procedimentos disponíveis para testar e estimar as relações de cointegração nas freqüências sazonais, bem como na freqüência zero, quando raízes unitárias sazonais estão presentes. Uma motivação importante para este trabalho é a falta de um tratamento sobre cointegração sazonal, mesmo nos livros-texto mais recentes sobre cointegração

Palavras-Chave: Análise de séries temporais, cointegração, sazonalidade.

ABSTRACT

While the literature on cointegration deals exclusively with the case of cointegration at the long-run or zero frequency between series in a vector of economic variables, it may happen that unit-roots are also present at the seasonal frequencies, and hence the concept of cointegration can be extended to the case of seasonal cointegration. In this paper we survey the available procedures for testing and estimating cointegration relationships at the seasonal frequencies, as well as at the zero frequency when seasonal unit-roots are present. A strong motivation for this is the lack of treatment of seasonal cointegration, even in the most recent books on cointegration.

Keywords: Time-series analysis, cointegration, seasonality.

* Universidade de Brasília.

† Universidade de São Paulo.

1 Introduction

The cointegration literature usually assumes that in a vector of economic time series, unit-roots are precisely equal to one (so that the series can have a spectral representation with a peak at the zero frequency.) However, many economic time series, due to a number of natural and economic factors such as, for example, climate changes and time-dependent preferences, exhibit strong seasonality characteristics, so that the spectral representation of these series will also have peaks at the seasonal frequencies. Two important aspects of seasonality are the fact that seasonal variations account for a great part of the variation in many economic series, and the fact that seasonal movements are often varying and changing, and in addition interdependent with the non-seasonal parts. Therefore, it seems desirable to develop tools for detecting the nature of the seasonal component of a series, and to integrate the available information from the data and from economic theory in a multivariate structural modelling of all components of the data. This paper provides the applied techniques implied in an attempt in this direction, represented by the available tests for seasonal cointegration of economic time series, for which a precise definition will be given. Intuitively, a series will be seasonally integrated if it exhibits a varying and changing seasonal pattern, and a group of seasonally integrated series will be seasonally cointegrated if they exhibit a parallel movement in their seasonal component. In other words, in a seasonally cointegrated vector of economic series, an innovation has a permanent effect on the seasonal pattern of each of its elements,whereas it has only a temporary effect on a specific linear combination of these elements.

When a vector of series is seasonally integrated, the Engle-Granger two-step procedure for testing for cointegration turns out to be inappropriate, yielding inconsistent results - see Theorem in Engle, Granger and Hallman(1990). Therefore, testing for cointegration in the usual sense, i.e. testing for the existence of common factors in the long run for nonstationary series, needs a new approach when unit-roots at frequencies other than the long run are believed to be present in the series, and also we need tests for the existence of common factors at seasonal frequencies,which is what the idea of seasonal cointegration conveys.

Hylleberg, Engle, Granger, and Yoo (1990) - from now on refered to as HEGY (1990) -, and Engle, Granger, Hylleberg, and Lee (1993) - from now on refered to as EGHL (1993) -, develop tests for seasonal cointegration which are basically modified versions of the Engle-Granger two-step procedure, trying to avoid the inconsistency problems above mentioned. However, this requires prior information on which seasonal unit-roots are present in order to filter out seasonal unit-roots components and to test for cointegration with the filtered series, which requires pretesting for seasonal unit-roots, the consequences of which HEGY (1990) claim not to have been investigated, merely conjecturing that it may be appropriate. Although we may try to solve this kind of problem by directly using seasonally adjusted series (which should eliminate seasonal unit-

roots), the fact that seasonal adjustment might lead to mistaken inference among time series data, and the fact that it causes loss of information on important seasonal behavior in economic time series when seasonal fluctuations are an important source of variation in the system, are arguments in favor of working with seasonally unadjusted data.

Lee (1992) extends the Johansen procedure - see Johansen (1988) - for testing for cointegration in the zero frequency to the case of testing for cointegration at the zero as well as at the seasonal frequencies. The derived ML estimation of cointegrating vectors and test statistics about cointegrating vectors provide a testing procedure that does not require any prior knowledge about the presence of seasonal unit-roots, avoiding the above mentioned problems of inconsistency and having to work with seasonally adjusted data. In what follows, we give some definitions and review the rationale and practical procedures for the above tests. In the end, some applications are mentioned.

2 Definitions

Definition 1. $x_t \sim I_\theta(d)$ denotes that x_t is seasonally integrated of order d at frequency θ, which means that the spectrum of x_t takes the form $f(\omega) = c(\omega - \theta)^{-2d}$ for ω near θ where c is a constant. We can see that, when $\omega = \theta$, $f(\theta)$ is infinite, what can be graphically depicted by a spike of infinite height at the frequency $\omega = \theta$, where consequently the series has infinite variance, what is in accordance of our understanding of unit-roots.

Here, we work with the practically important case where $d = 1$, and the integrated seasonal process for quarterly data is $(1-B^4) x_t = \varepsilon_t$.

We can see that the operator $(1-B^4)$ can be decomposed as

$$(1-B^4) = (1-B)(1+B)(1+B^2) = (1-B)(1+B)(1-iB)(1+iB)$$

so that the associated unit-roots with modulus 1 are 1, -1, i and -i, each corresponding to a particular frequency:

1. frequency $\theta = 0 (\omega = 0)$ which can be seen by writing:

$$(1-B^4)x_t = (1-B) S_1(B)x_t = (1-B)y_{1t}$$

where $S_1(B) = (1+B+B^2+B^3)$ is a seasonal filter which applied to x_t produces y_{1t}, a series which has a unit-root only at zero frequency (eliminating the unit-roots at seasonal frequencies $\omega = 1/4$ and $\omega = 1/2$).

2. frequency $\theta = \pi$ $(\omega = 1/2)$, at two cycles per year, which can be seen by writing:

$$(1-B^4)x_t = (1+B) \, S_2(B)x_t = (1+B)y_{2t}$$

where $S_2(B)=(1-B+B^2-B^3)$ is a seasonal filter which applied to x_t produces y_{2t}, a series which has a unit-root only at frequency $\omega = 1/2$ (eliminating the unit-roots at frequencies $\omega = 0$ and $\omega = 1/4$).

3. frequencies $\theta = +/- \dfrac{\pi}{2}$ $(\omega = 1/4)$, at one cycle per year, which can be seen by writing:

$$(1-B^4)x_t = (1+B^2) \, S_3(B)x_t = (1+B^2)y_{3t}$$

where $S_3(B) + (1-B^2)$ is a seasonal filter which applied to x_t produces y_{3t}, a series which has a unit-root only at frequency $\omega = 1/4$ eliminating the unit-roots at frequencies $\omega = 0$ and $\omega = 1/2$).

Definition 2. Let all the components of x_t be $I_\theta(1)$ Then, the components of x_t are said to be seasonally cointegrated at frequency θ, denoted $x_t \sim CI_\theta(1,1)$, if there exists a vector α (different from 0) such that $z_t = \alpha' x_t$ is $I_\theta(0)$.

For a vector of nonstationary series which have unit-roots at some seasonal frequencies as well as in the zero frequency, it is possible for a single cointegrating vector to eliminate all the unit-roots in the series, motivating the following definition :

Definition 3. Let each component of x_t be seasonally integrated of order 1 at some frequencies, including the zero frequency, not necessarily the same frequencies for all components. Then, the components of x_t are said to be fully cointegrated, denoted $x_t \sim CI(1,1)$, if there exists a vector α (different from 0) such that $z_t = \alpha' x_t$ is $I_\theta(0)$, for $\theta = 0$. 1/4, and 1/2.

3 Characterizations of Seasonal Cointegration

As in the case of cointegration at the zero frequency, we can characterize seasonal cointegration possibilities in terms of moving-average, autoregressive, and error-correction representations.

Beginning with the moving-average representation, we need the following proposition, due to Lagrange, which we state without proof:

Proposition. Any (possibly infinite or rational) polynomial j (B), finite valued at the distinct, non zero, possibly complex points q1,   .,q_p , has the following representation:

$$\varphi(B) = \sum_{k=1}^{p} \left[\lambda_k \Delta(B)/ \delta_k(B) \right] + \Delta(B) \; \varphi^{**}(B) \tag{3.1}$$

where the λ_k are a set of constants, $\varphi^{**}(B)$ is a (possibly infinite or rational) polynomial, $\delta_k(B) = 1 - (1/\theta_k)B$, and $\Delta(B) = \prod_{k=1}^{p} \delta_k(B)$

Now, let x_t be an $N \times 1$ vector of zero-mean variables which are all $I_\theta(1)$ at the zero and all seasonal frequencies. Its Wold representation is

$$(1-B^4)x_t = C(B) \, \varepsilon_t \tag{3.2}$$

where ε_t is a $N \times 1$ vector white-noise process with zero mean and positive definite covariance matrix S, and $C(B)$ is an $N \times N$ matrix of lag polynomials. We apply decomposition (3.1) to $C(B)$ to get

$$C(B) = \sum_{k=1}^{p} \left[\Lambda_k \Delta(B)/ \delta k(B) \right] + C^{**}(B) \Delta(B) \tag{3.3}$$

where δ_k and $\Delta(B)$ are as before and Λ_k is the vector counterpart to λ_k

We saw that $(1-B^4)$ has four unit-roots: $1, -1, i$, and $-i$. The root $-i$ is indistinguishable from the one at i when quarterly data is used. Substituting these roots in (3.3) gives us

$$C(B) = \theta-1(1+B+B^2+B^3) + \theta_2(1-B+B^2-B^3) + (\theta_3+\theta_4 B)(1-B^2) + C^{**}(B)(1-B^4) \tag{3.4}$$

where $\theta_1 = C(1)/4$, $\theta_2 = C(-1)/4$, $\theta_3 = \mathrm{Re}[C(i)]/2$, and $\theta_4 = \mathrm{Im}[C(i)]/2$.

Multiplying (3.2) by a vector α' gives us

$$(1-B^4) \, \alpha' x_t = \alpha' C(B) \, \varepsilon_t \tag{3.5}$$

If there exists $\alpha = \alpha_1$ such that $\alpha_1' \, C(1) = 0$, then x_t is cointegrated at zero frequency with cointegrating vector α_1. To see why this is true, we look at the following argument: when $\alpha_1' \, C(1) = 0$, we get from (3.5) that $\alpha_1' \, \theta_1 = 0$; moreover, $\alpha_1' \, C(1) = 0$ implies that the RHS of (3.5) has a

common factor of (1-B).[1] Using these two peces of information. (3.5) reduces to

$$(1+B+B^2+b^3)\alpha_1' x_t = \alpha_1' [\theta_2(1+B^2) + (\theta_3+\theta_4 B)(1+B) + C^{**}(B)(1+B+B^2+B^3)]\varepsilon_t \quad (3.6)$$

The RHS of (3.6) is stationary, and therefore $\alpha_1' x_t$ has unit-roots at the seasonal frequencies, but not at the zero frequency. It is worth noting that the vector $y_{1t} = S_1(B)x_t$ is I(0), while $\alpha_1' y_{1t}$ is stationary whenever $\alpha_1' C(1) = 0$. Hence, y_{1t} is cointegrated in the original sense, as described in Engle and Granger (1987). This fact will be useful ahead, when testing for cointegration.

Similarly, we can see that $(1+B)y_{2t} = -C(B)\varepsilon_t$, where $y_{2t} = -S_2(B)x_t$. Thus, y_{2t} has a unit-root at -1, and if there exists a vector α_2 such that $\alpha_2' C(-1) = 0$, then $\alpha_2' \theta_2 = 0$, and $\alpha_2' y_{2t}$ will not have a unit-root at -1. This enables us to characterize cointegration at frequency 1/2 as follows: $x_t \sim CI_{1/2}(1,1)$ with cointegrating vector α_2 if $\alpha_2' C(-1) = 0$.

Finally, define $y_{3t} = -S_3(B)xt$. By the same reasoning as above, y_{3t} has unit-roots at frequency 1/4, but this is not the case with $\alpha_3' y_{3t}$ if $\alpha_3' c(i)=0$. This gives us a characterization of cointegration at this frequency. However, this characterization can be made more general, as HEGY (1990) point out, by means of using the polynomial cointegrating vector (PCIV) as introduced by Yoo: $x_t \sim CI_{1/4}(1,1)$ with PCIV $\alpha_3 + \alpha_4 B$ if $(\alpha_3' + \alpha_4' i)(\theta_3 - \theta_4 i) = 0$, which is equivalent to the condition $\alpha(i)'C(i) = 0$.

This analysis can be generalized for the case where the cointegrating rank r > 1. This is easily accomplished by defining A_1, A_2, A_3, and A_4 as $N \times r_1$, $N \times r_2$, $N \times r_3$ and $N \times r_4$ matrices and C(B) as a polynomial matrix. The cointegrating ranks at the frequencies 0,1/2, and 1/4 are, r_1, r_2, and r_3 respectively.

This completes the characterization of seasonal cointegration in terms of the moving-average representation. However, both in the development of some tests, as we shall see, and in applied work, the autoregressive representation, and specially the error-correction representation, are more useful.

---

1 If $\alpha_1' C(1) = 0$, then we can write $\alpha_1' C(B) = (1-B) H(B)$, where $H(1) \neq 0$.

HEGY (1990) provide an error-correction representation by first using the so-called Smith-McMillan-Yoo decomposition [(as presented in Engle (1987)], expressing the moving-average polynomial C(B) as $C(B)=U(B)^{-1}M(B)V(B)^{-1}$ where M(B) is a diagonal matrix whose determinant has roots only on the unit circle, and the roots of the determinants of $U(B)^{-1}$ and $V(B)^{-1}$ lie outside the unit circle. Substituting this decomposition into (3.2) and doing some manipulations they obtained the autoregressive representation

$$\psi(B)x_t = \varepsilon_t \tag{3.7}$$

where $\psi(B) = V(B)\,\overline{M}(B)U(B)$ and $\overline{M}(B)$ is a transformation of M(B). So, $\psi(B)$ contains an integrated part and a stationary part. From this representation, HEGY (1990) obtain the error-correction model, which can be written as

$$\tilde{A}* \Delta_4 x_t = \Gamma_1 A_1 \, y_{1,t-1} + \Gamma A_2 \, y_{2,t-1} - (\Gamma_3 + \Gamma_4 B)(A_3 + A_4 B)y_{3,t-2} + \varepsilon_t \tag{3.8}$$

where $\Gamma_1, \Gamma_2, \Gamma_3$ and Γ_4 are $N \times r_1$, $N \times r_2$, $N \times r_3$ and $N \times r_4$ matrices containing the weights to the cointegrating relations, y_{1t}, y_{2t}, and y_{3t} are the vectors defined above, $\tilde{A}*(B)$ is an $N \times N$ polynomial matrix with all its roots outside the unit circle, $\tilde{A}*(0) = I_n$, and A_1, A_2, A_3, and A_4 are $N \times r_1, N \times r_2, N \times r_3$ and $N \times r_4$ matrices. The columns in A_1 and A_2 are the cointegrating vectors at the frequencies 0 and 1/2, respectively, and r_1 and r_2 are the cointegrating ranks. $A(B)=A_3+A_4 B$ is the cointegrating polynomial matrix, the colums of which are the PCIVs.

Lee (1992) departs from an autoregressive representation and derives a slightly different form for the error-correction model, which will provide the basis for his test for seasonal cointegration. Consider a nonstationary VAR process which has unit-roots at seasonal frequencies as well as at the zero frequency. The observed data is a $N \times 1$ sequence of vectors x_t, and the process can be written as

$$x_t = \psi_1 x_{t-1} + \psi_2 x_{t-2} + \quad + \psi_p x_{t-p} + \varepsilon_t, \ t = 1, \quad ,T \tag{3.9}$$

where ε_t are n.i.d.(0,S). This is a VAR model of order p, where the $N \times N$ matrices $(\psi_1, \ldots, \psi_p, S)$ are parameters to be estimated. Let $\Delta s \equiv 1-B^s$, where s is the number of observations taken per year, and $\psi(z)$ be the matrix polynomial $\psi(z) = I - \psi_1 z - \psi_2 z^2 - \quad - \psi_p z^p$.

In the case where $\Delta_s x_t$ is stationary, the determinant $|\psi(z)|$ has unit-roots at the zero frequency $(\omega = 0)$ and all the seasonal frequencies $(\omega = j/s; j = 1, \quad , s/2)$. The general structure for the error correction representation of this model is:

$$\Delta_s x_t = \beta_1 \Delta_s x_{t-1} + \quad + \beta_{p-s} \Delta_s x_{t-p+s} + \Pi_1 S_1(B) x_{t-1} + \quad + \Pi_s S_s(B) x_{t-1} + \varepsilon_t$$

where $\beta_i = -I + \psi_1 + \quad + \psi_i; i = 1, \quad , p\text{-s}.$

As in Lee (1992), we simplify the exposition by working with the practically interesting particular case of quarterly data, taking s = 4, so that the error correction representation, which can be directly derived by using decomposition (3.1), becomes

$$\Delta_4 x_t = \beta_1 \Delta_4 x_{t-1} + \quad + \beta_{p-4} \Delta_4 x_{t-p+4} + \Pi_1 y_{1,t-1} + \Pi_2 y_{2,t-1} + \Pi_3 y_{3,t-1} + \Pi_4 y_{3,t} + \varepsilon_t \qquad (3.10)$$

where $y_{j,t-1} \equiv Sj(B) x_{t-1}; j = 1,2,$ and $y_{3,t-1} \equiv BS_3(B) x_{t-1}$. The parameters $\beta_1, \quad , \beta_{p-4}$ are unrestricted, whereas the parameters $\Pi_1, \quad , \Pi_4$ are restricted to singularity, because of cointegration and seasonal cointegration. As all the components of x_t are $I_\theta(1)$, $\Delta_4 x_t, \beta_1 \Delta_4 x_{t-1}, \quad , \beta_{p-4} \Delta_4 x_{t-p+4}$ are $I_\theta(0)$ Now, the terms $\Pi_j y_{j,t-1}$ are linear combinations of $I_\theta(1)$ variables, and they should be $I_\theta(0)$, given that $\Delta_4 x_t$ is $I_\theta(0)$ Therefore, each one of the matrices Π_j contains a number of cointegrating vectors (each matrix corresponding to a particular frequency, as we shall see), or they are null matrices.

Note. $\Pi_1 = -\psi(1)/4$, and the filtered series $y_{1,t-1}$ has a unit-root only at the frequency $\omega = 0$. $\Pi_2 = \psi(-1)/4$, and the filtered series $y_{2,t-1}$ has a unit-root only at the frequency $\omega = 1/2$. $\Pi_3 = \text{Re}\,[\psi(i)]/2$, and $\Pi_3 = -\text{Im}\,[\psi(i)]/2$ and the filtered series $y_{3t}, y_{3,t-1}$ have unit-roots only at frequency $\omega = 1/4$.

4 Tests for Seasonal Cointegration

4.1 The HEGY and EGHL Procedures for Testing for Seasonal Cointegration

In order to keep matters as simple as possible, here we deal with the bivariate case, letting $x_t = [c_t, z_t]'$, where c_t is a consumption series and z_t is an income series. The first test we have to provide is for cointegration at the zero frequency. The natural idea is to follow Engle and Granger (1987) and run the static OLS regression

$$c_t = a z_t + u_t \qquad (4.1)$$

and test for the existence of a unit-root at the zero frequency in the residuals. However, if c_t and z_t are cointegrated at both the zero and the seasonal frequencies, with cointegrating vectors, say, α_1 and $\alpha_2, \alpha_1 \neq \alpha_2$, then we don't know what value of a will be chosen by the static regression.

Nevertheless, EGHL (1993) provide us a test which avoids this sort of problem. The test can be better understood if we write (3.7) for the bivariate case as

$$\Delta_4 c_1 = \sum_{j=1}^{q} \delta_j \Delta_4 z_{t-j} + \sum_{j=1}^{q} \beta_i \Delta_4 c_{t-i} + \gamma_{11}(c_{1,t-1} \quad \alpha_{12} z_{1,t-1}) + \gamma_{12}(c_{2,t-1} - \alpha_{22} z_{2,t-1}) - $$

$$-(\gamma_{13} + \gamma_{14} B)(c_{3,t-2} - a_{32} z_{3,t-2} \quad a_{41} c_{3,t-3} - a_{42} z_{3,t-3}) + \varepsilon_t \qquad (4.2)$$

and a similar form for z_t. Note that we are defining $y_{i,t} = [c_{i,t.} \ z_{i,t}]'$; $i = 1, 2, 3$, where $y_{i,t}$ is defined as in section 3.

The terms in (4.2) must be stationary, since we have an I(0) process in the LHS. So now, although $y_{i,t}$ and $c_{i,t}$; $i = 1, 2, 3$, have asymptotically infinite variance, the following particular linear combination will be I(0) at all frequencies:

$$s_{1t} = c_{1t} - \alpha_{12} z_{1t} \qquad (4.3)$$

$$s_{2t} = c_{2t} - \alpha_{22} z_{2t} \qquad (4.4)$$

$$s_{3t} = c_{3t} - \alpha_{32} z_{3t} - \alpha_{41} c_{3,t-1} - \alpha_{42} z_{3,t-1} \qquad (4.5)$$

The testing procedure for (4.3) and (4.4) is clear. A least squares regression will give a superconsistent estimate of the cointegrating parameters as in the Engle-Granger two-step method. Notice that now we don't have the same problem as before when we were running the static OLS regression (4.1), because c_{it} and $z_{i,t}$, $i = 1,2$ will be I(1) only at the frequencies zero and 1/2, respectively. Moreover, the residuals from (4.3) and (4.4) can be used to test for any remaining unit-roots at the particular frequencies zero and 1/2 using the tests for seasonal unit-roots derived in HEGY (1990).

In the case of z_{3t}, the fact that it is I(0) is not sufficient to identify the parameters to be estimated. This is so because z_{3t} is a dynamic relation, and we can find another linear combination of its elements that is also I(0). The solution, then, is to eliminate one of the variables, say $c_{3,t-1}$, which is the procedure suggested by EGHL (1993).

Now we can describe the testing procedure. Let u_t, v_t, and w_t be the residuals from regressing c_{1t} on z_{1t}, c_{2t} on z_{2t} and c_{3t} on z_{3t}, respectively. To test for non-cointegration at the zero frequency we run a regression of Δu_t on u_{t-1}, and test for unit-roots. This regression can include deterministic parts and/or be augmented by the necessary lagged values of Δu_t to whiten the errors. The

distribution of the t-ratio for the coefficient of u_{t-1} is not the usual Dickey-Fuller distribution because the cointegrating vector is estimated. The fatter-tailed distribution obtained by Engle and Granger (1987) should be used.

To test for noncointegration at the frequency 1/2 we run a regression of Δv_t on $-(v_{t-1})$, where the minus sign is needed in order to use the Engle and Granger distribution mentioned above.

Finally, the test for non-cointegration at the frequency 1/4 corresponds to testing if both π_3 and π_4 are zero in the auxiliary regression of the form:

$$(w_t + w_{t-2}) = \pi_3(-w_{t-2}) + \pi_4(-w_{t-1}) + \text{error.}$$

The asymptotic distribution of the t-ratios and the F statistics are given in Theorems 1 and 2 in EGHL (1993), where we can also find the critical values calculated via a series of Monte Carlo experiments.

One final remark is that the test described can be performed by first filtering the data to get rid of "excessive" variance at those frequencies not being tested, what is accomplished by use of the filters y_{it}, $i = 1, 2, 3$, and then performing the appropriate regressions.

4.1.1 An Application: The Japanese Consumption Function

We now review the application of the testing and estimation procedures of section 4.1 to data on japanese consumption and income, which is in logarithmic form, denoted as c_t and z_t, as presented in EGHL (1993). Starting with a cointegrating vector $\alpha = [1,-1]$, which seems to be a sensible candidate, based on the theory of permanent income, and using the distribution given in HEGY (1990), they couldn't reject the hypothesis of a unit-root at any of the frequencies 0, 1/4, 1/2, and 3/4.

The next step is to run a regression of c_{1t} on z_{1t}. Several different regression were used: (a) with an intercept and a trend, (b) with an intercept only; and (c) without any deterministic part. Then, the residuals of these regressions were tested for unit-roots at the zero frequency based on the ordinary augmented Dickey-Fuller regression. We show the results for regression (a) only:

coefficient (coint. vector)	R^2	Test for unit-root in residuals	
		$t_{\hat{n}_1}$	critical value
0.772 (0.015)	0.998	-0.59	3.17

We can see that the hypothesis of a unit-root cannot be rejected, implying non-cointegration at the long-run frequency. Tests based on (b) and (c) had the same results. A test based on the residuals from the cointegrating regresssion of c_{2t} on z_{2t} also could not reject the hypothesis of non-cointegration at the frequency 1/2, as can be verified below, where we present the results for the cointegrating regression that includes an intercept and seasonal dummies.

coefficient (coint. vector)	R^2	Test for unit-root in residuals	
		$t_{\hat{\pi}_2}$	critical value
0.236 (0.015)	0.939	-1.81	3.17

For the frequency 1/4 (and 3/4), the results are given below for the regression including seasonal dummies and an intercept, and where the cointegrating regression of c_{3t} on z_{3t} and $z_{3,t-1}$ gives us the polynomial cointegrating vector between c_{3t} and z_{3t}.

coefficient (coint. vector)		R^2	Test for unit-root in residuals		
			$t_{\hat{\pi}_3}$	$t_{\hat{\pi}_4}$	$F_{\hat{\pi}_3 \cap \hat{\pi}_4}$
0.264 (0.067)	0.109 (0.034)	-0.974	3.60	-2.17	8.90

The critical values for the t and F statistics on π_3 and π_4 are -4.40, -2.14, and 10.12 at a 5 level of significance, and for the F statistic at a 10 level, 8.66. The critical values are given in Table A.1 in the appendix of HEGY (1990).

Hence, there seems to be a "weak" evidence of cointegration between c_t and y_t at the seasonal frequencies 1/4 and 3/4, where this "weakness" becomes more clear when we know that the results for the cointegration regression with no seasonal dummies indicate non-cointegration.

Summarizing, the tests applied to total consumption and disposable income in Japan from 1961.1 to 1987 4 indicate that the logs of the income and the consumption series are integrated of order 1 at both the long-run frequency and the seasonal frequencies. However, the results of the cointegration tests indicate that the seasonally adjusted series are neither cointegrated at the long-run frequency nor at the bi-annual frequency ($\omega = 1/2$), while there are some signs that they may be cointegrated at the annual frequency ($\omega = 1/4, 3/4$).

The interpretation for this cointegration at the annual frequency is that we are dealing with slightly impatient, borrowing-constrained, utility-mazimizing consumers that use their end-of-the-year bonus payments to replace their worn out clothes, furniture, etc. They spend the bonuses as

soon as the payments occur. This bonus system is characteristic of the japanese economy.

4.2 The Lee Procedure for Testing for Seasonal Cointegration

The approach followed by Lee (1992) is an extension of the Johansen procedure (Johansen 1988) to incorporate the possibility of seasonal cointegration in this system. Recall the error-correction representation (3.10):

$$\Delta_4 x_t = \beta 1 \, \Delta_4 x_{t-1} + \quad + \beta_{p-4} \Delta_4 x_{t-p+4} + \Pi_1 y_{1,t-1} + \Pi_2 y_{2,t-1} +$$

$$+ \Pi_3 y_{3,t-1} + \Pi_4 y_{3,t} + \varepsilon_t \qquad\qquad\qquad (3.10)$$

We need to investigate the properties of the matrices Π_j, $j = 1,$. 4 in order to determine whether the components of x_t are seasonally cointegrated in the presence of unit-roots at other frequencies. If the components of x_t are $I_\theta(1)$, but are not $CI_\theta(1,1)$ then the term $\Pi_j y_{j,t-i}$ would be nonstationary and the true Π_j must be a null matrix, whereas if the elements of xt are $CI_\theta(1,1)$, the the term $\Pi_j y_{j,t-i}$ will be stationary, which means that we can express Π_j as the product of two n × r matrices $\Gamma_j A_j'$, where A_j is a matrix of cointegrating vectors at frequency θ, where j =1 for θ = 0 (ω = 0), j = 2 for θ = π (ω = 1/2), and j = 3, 4 for θ = π/2 (ω = 1/4).

4.2.1 Procedure for testing for cointegration at frequency $\omega = 0$

Following the same reasoning as in Johansen (1988), we consider the maximum likelihood estimation of the parameters in the unrestricted model (3.10) We need to maximize the likelihood function with respect to the parameters $(\beta_1, \quad, \beta_4, S)$ and $(\Pi_1, \quad \Pi_4)$. This can be practically done by doing the OLS regression of $\Delta_4 x_t$ on $\Delta_4 x_{t-1}$, $\quad \Delta_4 x_{t-p+4}$ and computing the residuals R_{0t}, and then doing the OLS regression of $y_{k,t-1}$ on $\Delta_4 x_t$ on $\Delta_4 x_{t-1}$, $\quad , \Delta_4 x_{t-p+4}$ and computing the residuals R_{kt}; k = 1,2,3,4. The maximum likelihood estimates of the Π_k's are equivalent to the parameters of the OLS regression of R_{0t} on R_{1t}, $\quad , R_{4t}$. From then, we also obtain the estimates of the other parameters in the likelihood function, S and β_1, $\quad, \beta_4$ by appropriate substitution of the estimated Π_k's and R_{kt}'s in the formulas given in Lee (1992). Turning to the model under $H_1: \Pi_1 = \Gamma_1 A_1$;, a practical procedure for obtaining the constrained maximum likelihood function is to do an OLS regression of R_{0t} on R_{2t}, R_{3t}, R_{4t} and compute the residuals Q_{0t}, then do an OLS regression of R_{1t} on R_{2t}, R_{3t}, R_{4t} and compute the residuals Q_{1t}, which allows the computation of $D_{i,j} = T^{-1} \sum_{t=1}^{T} Q_{it} Q_{jt}'$; i, j = 0,1. By the same reasoning as in Johansen (1988) the likelihood function is maximized when the quotient $| A'D_{11} A1 - A_1' D_{10} D_{00}^{-1}$

$D_{01} A_1 |/| A_1 D_{11} A_1 |$ is minimized with respect to all matrices A_1. We can get $\hat{A}_1$ by choosing the first r eigenvectors of $D_{10} D_{00}^{-1} D_{01}$ w.r.t D_{11} ($\hat{\lambda}_{1,r+1}$, , $\hat{\lambda}_{1,n}$)

This allows the implementation of the following theorem in Lee (1992):

Theorem 3.1. The Likelihood Ratio test statistic for the hypothesis $H_1: \Pi_1 = \Gamma_1 A_1$ (i.e., that there are at most r cointgrationg vectors at zero frequency) is $-2 \ln (Q) = -T \sum_{i=r+1}^{n} \ln(1- \hat{\lambda}_{1,i})$, where Q is the quotient of the restricted and unrestricted maximized likelihoods, and ($\hat{\lambda}_{1,r+1}$, , $\hat{\lambda}_{1,n}$) are the N - r smallest partial canonical correlations of R_{1t} with respect to R_{0t}, given R_{2t}, R_{3t}, R_{4t}. The LR test statistic has an asymptotic distribution which is a function of an (N - r) dimensional Brownian Motion, for which a table with critical values can be found in Lee (1992).

4.2.2 Procedure for testing for seasonal cointegration at frequency $\omega = 1/2$

When we want to test if the components of x_t are seasonally cointegrated at the seasonal frequency $\omega = 1/2$ in the presence of unit-roots at the zero and other seasonal frequencies, the testing procedure is very similar to the above case where $\omega = 0$, except for the fact that the roles of R_{1t} and R_{2t} are reversed, when we can implement the following theorem in Lee(1992):

Theorem 3.2. The LR test statistic for the hypothesis that there are at most r seasonal cointegrating vectors at frequency $\omega = 1/2$ is $-2 \ln (Q) = -T \sum_{i=r+1}^{n} \ln(1- \hat{\lambda}_{2,i})$, where ($\hat{\lambda}_{2,r+1}$, , $\hat{\lambda}_{2,n}$) are the (N - r) smallest partial canonical correlations of R_{2t} with respect to R_{0t}, given R_{2t}, R_{3t}, R_{4t}. This statistic has the same asymptotic distribution as the one above (for $\omega = 0$).

4.2.3 Procedure for testing for seasonal cointegration at frequency $\omega = 1/4$

In this case, we would like to look simultaneously at the two parameter matrices Π_3 and Π_4 since the cointegrating vectors and the coefficients of the error correction term may be different at different lags. Assuming for simplicity, as in Lee(1992) that cointegration, if any, is contemporaneous, the hypothesis of interest can be formulated as a joint test such that $H_q . \Pi_3 = \Gamma_3 A_3 \cap \Pi_4 = \Gamma_4 A_4$ (i.e., that there exist at most r seasonal cointegrating vectors at frequency), for which the following theorem provides a test:

Theorem 3.3. The LR test statistic for the hypothesis that there are at most r seasonal cointegrating vectors at frequency $\omega = 1/4$ is $-2 \ln (Q) = -T \sum_{i=r+1}^{T} \ln (1- \hat{\lambda}_{3,i} - \hat{\lambda}_{4,i})$ where ($\hat{\lambda}_{k,r+1}$, $\hat{\lambda}_{k,n}$)

are the $(N - r)$ smallest partial canonical correlations of R_{kt}; $(k = 3,4)$ with respect to R_{0t}. Again, the asymptotic distribution and critical values are in Lee (1992).

4.2.3 Procedure for testing for full cointegration

As mentioned above, it is possible that one cointegrating vector, say A_f, could eliminate unit-roots of the series at all frequencies. This will happen in the case where the cointegrating vectors A_1, A_2 and A_q coincide, which, along with the assumption of contemporaneous cointegration, allows the error correction model to be written as

$$\Delta_4 x_t = \Pi_f x_{t-4} + \beta_1 \Delta_4 x_{t-1} + \quad + \beta_{p-4} \Delta_4 x_{t-p+4} + \varepsilon_t$$

Then the testing procedure for the hypothesis H_f. $\Pi_f = \Gamma_f A_f'$ is given in the following theorem:

Theorem 3.5. The LR test statistic for the hypothesis that there are at most r full cointegrating vectors is $-2 \ln (Q) = -T \sum_{i=R+1}^{T} \ln (1 - \hat{\lambda}_{f,i})$, where $(\hat{\lambda}_{f,r+1}, \quad , \hat{\lambda}_{f,n})$ are the $(N - r)$ smallest partial canonical correlations of R_{ft} with respect to R_{0t}, where R_{ft} is the residual from regressing x_{t-4} on $\Delta_4 x_{t-j}, j = 1, \quad . p-4$. The asymptotic distribution for this test statistic is in Lee (1992), along with critical values.

4.2.5 An Application: Canadian immigration and unemployment

The procedures above are applied in Lee (1992) to test for the existence of cointegration and/or seasonal cointegration between series of Canadian unemployment and immigration data. From visual inspection of the plot of these series, we can expect to find long-run and seasonal unit-roots in both series, which is confirmed by the tests reported in Table 5 in Lee (1992). So, there exists the possibility for testing for cointegration at the zero and seasonal frequencies between the series. The LR test statistics, also reported in Table 5 in Lee (1992) do not allow the rejection of the hypothesis of no cointegration at frequencies 0 (i.e., H_0: rank $(\Pi_1) = 0$) and 1/4 (i.e., H_0: rank(Π_3) $= $ rank$(\Pi_4) = 0$). However, the LR test statistic for testing H_1. rank$(\Pi_1) = 0$ is 13.567. which exceeds the critical value at the 5 significance value in Table 4.c of Lee (1992), leading us to reject the hypothesis of no cointegration at frequency 1/2. Therefore, we conclude that the data exhibit one seasonal cointegration relationship at $\omega = 1/2$ only.

The equilibrium error process, given by the seasonal cointegration relation between the two series, indicates a negative relationship between unemployment and immigration rates:

$$z_t = (1 - B + B^2 - B^3)\,(UN_t + 11.49IMMt)$$

where the estimated seasonal cointegrating vector has been normalized. When the system is put in error-correction form, the equation for unemployment shows that the coefficients of the lagged values of immigration an the coefficient for the error-correction term are all significant, implying a highly significant relationship for immigration causing unemployment, which is not true when we look at the other way, i.e., the model does not imply that past values of unemployment affect immigration. The existence of a seasonal cointegration relationship between immigration and unemployment is therefore interpreted as an important information on the effectiveness of immigration policy in Canada, in the sense that it appears to be able to offset the impact of the varying pattern of seasonality in immigration on the seasonal pattern of unemployment.

5 Concluding remarks

The procedures for estimation and testing of seasonal cointegrating relationships reviewed here provide a way for modelling the long-run relationship between seasonal patterns of economic variables in a way that can be linked to economic theory. We also saw that these tests allow us to circumvent the potential problems involved in the estimation and testing of cointegration relationships at the zero frequency when unit-roots are present at the seasonal frequencies as well. These potential problems could imply that if the seasonal aspects of a vector of economic series are overlooked, we could mistakenly infer that long-run relationships between these variables do not exist, when in fact they may be present.

The procedure suggested by Lee (1992) seems to be superior in relation to the ones suggested by HEGY (1990) and EGHL (1993) which involve normally undesirable filtering of the series. There is one point regarding these tests which should be mentioned. In general, the inclusion of seasonal dummies and/or an intercept term in the model has the effect of slightly modifying the distributions of the tests statistics.

A useful application of seasonal cointegration is described in Engle, Granger and Hallman (1989), where it is argued that if we want to forecast the behavior of series that are affected by short-run and long-run factors, a superior forecasting strategy is to construct an error-correction model that incorporates both these aspects. As the construction of the error-correction model is based on the consistent estimation of the cointegrating relationship between the series, it is fundamental for the forecasting performance of the model to account for the presence of seasonal unit-roots in the series, specially if it is believed that the short-run aspects of the model are subject to important seasonal variations. To estimate these relationships accordingly, the procedures reviewed here are available.

Yet another useful application of the seasonal cointegration model is to provide superior representations of economic series subject to seasonal fluctuations, in the sense that modelling the seasonality of these series stochastically allows for the information contained in the relationships between them not to be wasted. Ermini and Chang (1996), for example, conduct alternative tests for the macro rational expectations hypothesis of rationality and money neutrality using Korean data on series of money supply, price level, output, and interest rates. When the test is first applied to the series desasonalized individually by the standard X-11 method, they reject the rational expectations hypothesis. However, when they applied the same test to the series with the seasonal factor modelled stochastically, as proposed here, they do not reject the rational expectations hypothesis. The seasonal cointegration relation between these series apparently provides important information which is lost when the traditional deseasonalizing procedure transforms filters the seasonal factors individually.

Another alternative extension of the usual cointegrating concept, taking into account periodic variations in the time patterns of economic series,is provided by Franses (1993). The model for periodic cointegration proposed in this article assumes that the parameters in the cointegrating vectors, as well as the ajustment parameters in the error-correction context, can vary over the seasons. The two approaches, the seasonal cointegration reviewed here, and the periodic cointegration proposed by Franses (1993) produce non-nested models, and Franses (1993) also proposes a test procedure for selecting between these models.

References

ENGLE, R.F., "On the theory of cointegrated economic time series" *U.S.C.D. discussion paper* no. 87-26, presented to the European meeting of the Econometric Society, Copenhagen, 1987.

ENGLE, R.F. & Granger, C.W.J. "Cointegration and error-correction: Representation, estimation, and testing" *Econometrica*, bf 55, 251-276, 1987.

ENGLE, R.F., C.W.J. Granger, & J. Hallman, "Merging short and long run forecasts: An application of seasonal cointegration to monthly electricity sales forecasting". *Journal of Econometrics*, bf 40, 45-62, 1989.

ENGLE, R.F., GRANGER, C.W.J. HYLLEBERG, S. & LEE, H.S. "Seasonal cointegration: The Japanese consumption function" *Journal of Econometrics*, bf 55, 275-298, 1993.

ERMINI, L., & CHANG, D., "Testing the joint hypothesis of rationality and neutrality under seasonal cointegration" *Journal of Econometrics*, bf 74, 363-386, 1996.

FRANCES, P.H., "A method to select between periodic cointegration and seasonal cointegration. *Economic Letters*, bf 41, 7-10, 1993.

HYLLEBERG, S., ENGLE, R.F., GRANGER, C.W.J. & YOO, B.S. "Seasonal integration and cointegration" *Journal of Econometrics*, bf 44, 215-238, 1990,.

JOHANSEN, S., "Statistical analysis of cointegrating vectors" *Journal of Economic Dynamics and Control*, bf 12, 231-254, 1988.

LEE, H.S., "Maximum likelihood inference in cointegration and seasonal cointegration". *Journal of Econometrics*, bf 54, 1-47, 1992.