On the role of seasonal intercepts in seasonal cointegration

Philip Hans Franses

Econometric Institute, Erasmus University Rotterdam

Robert M. Kunst

Institute for Advanced Studies, Vienna, and Johannes Kepler University, Linz

Econometric Institute Report 9820/A

Abstract

In the paper we consider the role of seasonal intercepts in seasonal cointegration analysis. For the nonseasonal unit root, such intercepts can generate a stochastic trend with a drift common to all observations. For the seasonal unit roots, however, we show that unrestricted seasonal intercepts generate trends that are different across the seasons. Since such seasonal trends may not appear in economic data, we propose a modified empirical method to test for seasonal cointegration. We evaluate our method using Monte Carlo simulations and using a four-dimensional data set of Austrian macroeconomic variables.

Keywords: Deterministic and stochastic seasonality; Seasonal cointegration; Cointegration.

JEL classification: C32.

Corresponding author

Robert M. Kunst

Institute for Advanced Studies, Stumpergasse 56, A-1060 Wien, Austria

Telephone +43-1-59991-255 Fax +43-1-59991-163 e-mail kunst@ihs.ac.at

Acknowledgments

This work was initiated when the second author visited the Tinbergen Institute Rotterdam (February 1995). Its hospitality is gratefully acknowledged. The first author thanks the Royal Netherlands Academy of Arts and Sciences for its financial support. We thank an anonymous referee for some helpful comments.

1 INTRODUCTION

Many quarterly economic time series display trending and seasonal patterns which do not appear to be constant over time. A representation of time series that accounts for time-varying trends and seasonals assumes the presence of stochastic trends at the zero and seasonal frequencies. Formulated otherwise, for several economic time series one can assume the presence of nonseasonal and seasonal unit roots. Hylleberg et al. (1990) propose formal test statistics to investigate these roots in univariate time series. Given a set of economic time series, it is of interest to study whether these have stochastic trends at certain frequencies in common. Hence, a usual next step in analyzing a set of quarterly time series involves testing for cointegration at the nonseasonal and seasonal frequencies. Engle et al. (1993) suggest so-called residual-based tests for seasonal cointegration, while Lee (1992) proposes tests for similar purposes based on a fully specified multivariate time series model.

Inference on cointegration and on common stochastic trends can be shown to depend critically on the presumed empirical model and on the deterministic regressors included in the auxiliary test regressions. See Johansen (1994) for a detailed treatment of the role of the constant and linear terms in analyzing cointegration at the nonseasonal frequency. This role is important since under the null and alternative hypotheses the constant and linear variables may have different implications. For example, an unrestricted constant in a model with imposed cointegration among two variables implies that the driving stochastic trend contains a drift. In the present paper, we consider the role of four seasonal intercepts in the seasonal cointegration model for quarterly data. Although the seasonal cointegration model with seasonal dummies is analyzed in Lee and Siklos (1995), the role of the four intercepts is not discussed. We show that the inclusion of unrestricted seasonal intercept parameters can lead to an undesirable feature of the data, and hence, that one may obtain inappropriate empirical results. This feature is that in case of cointegration at a seasonal frequency, the data are assumed to have deterministic trends that vary with the season. To overcome this, we propose a simple modification of the standard seasonal cointegration analysis.

The outline of this paper is as follows. In Section 2, we start with a discussion of the representation of seasonal cointegration with the inclusion of seasonal intercepts. We examine the role of these intercepts and summarize the main results in a proposition. In the second part of Section 2, we propose an alternative empirical strategy to test for seasonal cointegration. We also provide new tables with critical values for several empirically relevant cases. In a sense, our strategy simply amounts to a partial cointegration analysis per frequency, where the intercept (with a specific form for each frequency) is restricted under the null hypothesis of cointegration. In Section 3, we report some experiments which we conduct in order to assess the potential improvement in test power by using the new procedure. In Section 4, we illustrate our approach for a four-dimensional macroeconomic system of the Austrian economy, and we compare our results with those obtained using the Lee (1992) method. In Section 5, we conclude our paper with some remarks.

2 SEASONAL COINTEGRATION AND SEASONAL INTERCEPTS

In this section we discuss the representation of a seasonal cointegration model. The key results on the impact of unrestricted and restricted seasonal intercept terms are summarized in a proposition. Our results clearly indicate a useful empirical modeling strategy for seasonal cointegration. In Section 2.2, we present this strategy and we provide several tables with critical values.

2.1 Representation

A general representation of an autoregressive process for an $n \times 1$ vector time series X_t $(t = 1, ..., N)$, which allows for cointegration at seasonal and nonseasonal frequencies is:

$$\begin{aligned} \Delta_4 X_t &= \alpha_1 \beta_1 S(B) X_{t-1} + \alpha_2 \beta_2 A(B) X_{t-1} + \alpha_3 \beta_3 \Delta_2 X_{t-2} \\ &\quad + \sum_{i=1}^{4} d_i \delta^i_{t-4[(t-1)/4]} + \varepsilon_t \quad , \end{aligned} \tag{1}$$

where ε_t is an $n \times 1$ vector white noise process. The $\delta^i_{t-4[(t-1)/4]}$ in (1) concern the conventional seasonal dummy variables. The Kronecker symbol expresses the structure of the deterministic seasonal dummies that can be equated to t mod 4. [.] is used to denote the largest integer or entier function. The differencing filter Δ_k, $k = 1, 2, ...$ is defined by $\Delta_k = (1 - B^k)$, where B is the usual backward shift operator defined by $B^k X_t = X_{t-k}$. Model (1) is more special than the most general form that would allow for non-synchronous seasonality at $\pi/2$ and it has the form used by Lee (1992). For ease of exposition, a possible short-run autoregressive influence has been excluded that would allow for VARs of any finite order p. An unrestricted vector autoregression for X_t that corresponds to (1) is of order 4. The matrices α_i and β_i are assumed to have full column ranks r_i with $0 \leq r_i < n$. The operators $S(B)$ and $A(B)$ are defined as $S(B) = 1 + B + B^2 + B^3$ and $A(B) = 1 - B + B^2 - B^3$, where $S(B)$ can be interpreted as the seasonal moving average smoothing operator and $A(B)$ as the alternating-signs summing operator, hence "S" and "A".

The matrix $\alpha_1\beta_1'$ corresponds to nonseasonal cointegration at the zero frequency. The matrix $\alpha_2\beta_2'$ concerns seasonal cointegration at the bi-annual frequency, whereas the matrix $\alpha_3\beta_3'$ relates to seasonal cointegration at the annual frequency. See Lee (1992) for additional discussions of model (1).

Given (1) and fixed starting values for the X_t vector process, X_t has a representation in starting values, innovations ε_s, and deterministic contributions D_s for $s \leq t$. Formally, this representation is achieved by inverting the seasonal operator Δ_4 in (1). For the now classical case of zero-frequency cointegration, the mathematical derivation of such a representation is summarized in the Granger Representation Theorem (cf. Engle and Granger, 1987). In the present case of a seasonal cointegration model as in (1), the influence of the deterministic terms is however more involved, and hence a representation theorem for seasonal cointegration contains complex structures (cf. Johansen and Schaumburg, 1997). To highlight this phenomenon, we decompose the deterministic part D_t of (1) into three components, i.e., $D_t = \mu + a_t + r_t$. Averaging over the seasonal cycle yields the time-constant drift $\mu = (d_1 + d_2 + d_3 + d_4)/4$, with the d_i parameters as defined in (1). The sum of the remaining components is then 0 over the four-quarter cycle. Similarly, a_t is found by applying the alternating operator $A(B)$ to the sequence of seasonal constants. This results in $a_t = a \cos \pi(t-1)$ with $a = (d_1 - d_2 + d_3 - d_4)/4$. The remainder r_t has a distinct pattern of alternating constants of type $(b, c, -b, -c, b, c, ...)$, with $b = (d_1 - d_3)/2$ and $c = (d_2 - d_4)/2$. Then,

$$r_t = b \cos \frac{\pi}{2}(t-1) + c \cos \frac{\pi}{2}(t-2) \quad .$$

From the definition of $D_t = \mu + a_t + r_t$, where μ, a_t, and r_t are defined above, it is clear that formal application of the inverse operator Δ_4^{-1} has different effects on the three deterministic components. For example, $\Delta_4^{-1}\mu$ yields four parallel linear time trends that perpetuate the original seasonal starting pattern. Such a pattern does not seem unusual for seasonal and trending series observed in practice. In striking contrast, a_t and r_t generate divergent linear trends. In the case of a_t and, e.g., $a > 0$, parallel positive trends appear for even t and parallel negative mirror images for odd t. In the case of r_t, the patterns look even more strikingly counterintuitive. However, in the multivariate model such divergent trends do not appear necessarily. Hence, the key problem addressed in this paper is how one should accommodate for and possibly restrict the deterministic part of seasonality in seasonal cointegration analysis such that the aforementioned implausible features are avoided.

For illustrative purposes, we start with a simple VAR(1) model with cointegration at the zero frequency. This model reads as

$$\Delta_1 X_t = \alpha\beta' X_{t-1} + \mu + \epsilon_t \,. \tag{2}$$

From integrating (2), it is clear that X_t depends on a linear trend through $\alpha^{\perp\prime}\mu$ only, where $\alpha^{\perp}$ denotes the orthogonal complement to α. Hence, even if μ is not 0 but $\alpha^{\perp\prime}\mu$ =0, there is no linear trend in the multivariate process X_t. For a similar property in the seasonal cointegration model, let us first re-write (1) with decomposed deterministic terms, i.e., with $D_t = \mu + a_t + r_t$, as:

$$\begin{aligned}\Delta_4 X_t &= \alpha_1\beta_1' S(B) X_{t-1} + \alpha_2\beta_2' A(B) X_{t-1} + \alpha_3\beta_3'\Delta_2 X_{t-2} + \mu \\ &\quad + a\cos\pi(t-1) + (b,c)\left(\cos\frac{\pi}{2}(t-1), \cos\frac{\pi}{2}(t-2)\right)' + \varepsilon_t \end{aligned} \tag{3}$$

where (b, c) is an $n \times 2$ matrix expressing the influence of the two annual dummies, each one of type (1,0,-1,0,1,0,-1,...) with one of them lagged one quarter. Obviously, (3) has exactly the same number of parameters as (1), since (μ, a, b, c) replaces $(d_1, ..., d_4)$. It is now possible to show that the unwanted and implausible divergent seasonal trends appear in the representation of X_t through $\alpha_2^{\perp\prime}a$ and $\alpha_3^{\perp\prime}(b, c)$ only. We state this result in the following proposition.

PROPOSITION: If a vector autoregression is given by (1) or equivalently (3) and X_t has four fixed consecutive starting values, then, in general, its deterministic part consists of an n-dimensional linear time trend proportional to $\alpha_1^{\perp\prime}\mu$, linear time trends diverging over the seasonal cycle proportional to $\alpha_2^{\perp\prime}a$ and $\alpha_3^{\perp\prime}(b,c)$, and a periodic pattern of constants. If $\alpha_2^{\perp\prime}a = 0$ and $\alpha_3^{\perp\prime}(b,c) = 0$, the deterministic part only contains a linear trend and four seasonal constants.

Proof. We use a result developed by Tsay and Tiao (1990) who build on previous work by Chan and Wei (1988). We re-write the VAR system in state-space form:

$$\begin{aligned} \begin{bmatrix} X_t \\ X_{t-1} \\ X_{t-2} \\ X_{t-3} \end{bmatrix} &= \begin{bmatrix} \mathbf{\Gamma}_1 & \mathbf{\Gamma}_2 & \mathbf{\Gamma}_3 & \mathbf{\Gamma}_4 \\ \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} X_{t-1} \\ X_{t-2} \\ X_{t-3} \\ X_{t-4} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ 0 \\ 0 \\ 0 \end{bmatrix} \\ \text{or} \quad X_t^* &= \mathbf{\Gamma} X_{t-1}^* + \varepsilon_t^* \quad , \end{aligned} \tag{4}$$

where $\mathbf{\Gamma}_j$, $j = 1,2,3,4$ are matrices directly depending on α_i and β_i of (1). With deterministic terms in (1), one must add additional terms to the right of (4), again with block zeros except in the first n-block as in ε_t^* . Tsay and Tiao (1990) use the Jordan canonical form of the state transition matrix $\mathbf{\Gamma} = \mathbf{T}^{-1}\mathbf{D}\mathbf{T}$ to rotate the system into $\mathbf{T}(X_t, X_{t-1}, X_{t-2}, X_{t-3}) = \mathbf{T}X_t^* = Y_t$, where $\mathbf{T}$ denotes a transition matrix. The vector process Y_t is a $4n$ -dimensional process, naturally ordered according to the eigenvalues of the original transition matrix. One could, e.g., consider Y_t with an $(n - r_1)$-dimensional subvector corresponding to the eigenvalue of +1, continue with an $(n-r_2)$-dimensional subvector corresponding to the eigenvalue of -1, which is the unit root that concerns the bi-annual frequency and a pair of $(n - r_3)$-dimensional vectors corresponding to i, which are the complex unit roots that concern the annual frequency. The real and imaginary parts of these latter vectors can also be interpreted in the real numbers as real eigenvectors to the eigenvalue of -1 in the squared transition matrix. The remaining eigenvalues are less than 1 in modulus. The matrix $\mathbf{T}$ contains the (left) eigenvectors, i.e., $\mathbf{T} = (\mathbf{T}_1, ..., \mathbf{T}_4)'$ with the components $\mathbf{T}_i$,$i = 1, ..., 4$ corresponding to +1,-1, i, and the remainder. Hence, the first $(n - r_1)$-component of Y_t is a random walk of the form

$$W_t = W_{t-1} + \mathbf{T}_1' \begin{pmatrix} \epsilon_t \\ 0 \\ 0 \\ 0 \end{pmatrix} + \mathbf{T}_1' \begin{pmatrix} \mu, a_t, r_t \\ 0 \\ 0 \\ 0 \end{pmatrix} . \tag{5}$$

This representation shows that it is only the first n-part of the eigenvectors comprising $\mathbf{T}_1$ that has any further influence on the deterministic components in the rightmost term. We denote this first part as $\mathbf{T}_{11}$. If a starting value for W_1 is given, then (5) can be inverted. The generated random walk is superimposed with a time trend of the form $\mathbf{T}_{11}'\mu t$. The seasonal variables a_t and r_t generate an additional cycle of constants. The whole system can be transformed into the original X_t by the inverse transformation matrix $\mathbf{T}^{-1}$. The contribution of the subsystem (5) is then a stochastic random walk component, a linear trend proportional to $\mathbf{T}_{11}'\mu$, and a basic repetitive pattern of constants. Hence, (5) yields a plausible impact of intercepts.

The second component has dimension $n - r_2$ and looks like

$$\tilde{W}_t = -\tilde{W}_{t-1} + \mathbf{T}_2 \begin{pmatrix} \epsilon_t \\ 0 \\ 0 \\ 0 \end{pmatrix} + \mathbf{T}_2 \begin{pmatrix} \mu, a_t, r_t \\ 0 \\ 0 \\ 0 \end{pmatrix} . \tag{6}$$

This is an $(n - r_2)$-dimensional random jump process. We again denote the first section of $\mathbf{T}_2$ by $\mathbf{T}_{21}$. Inverting the representation (6) using one starting value leads to five parts. Firstly, the purely stochastic alternating sum of white noise. Secondly, an alternating influence of the form $(\mathbf{T}_{21}'\mu, 0, \mathbf{T}_{21}'\mu, 0, \mathbf{T}_{21}'\mu, 0, ...)$. Thirdly, two diverging trends at odd and even indices t corresponding to $\mathbf{T}_{21}'a_t$, i.e. different trends for different seasons. Fourthly, a cyclical pattern of constants corresponding to the r_t influence. Fifthly, an alternating additional term deriving from the starting conditions. The diverging seasonal trends in the third influences deserve further attention, since it is this effect that may not be present in empirical data. That influence, however, is strictly rooted in $\mathbf{T}_{21}'a$. Re-transforming with $\mathbf{T}^{-1}$, it can be shown that there is a stochastic seasonal cycle of periodicity π in the original process depending on $\mathbf{T}_{21}$ and an additional deterministic feature proportional to $\mathbf{T}_{21}'a$.

In order to continue with our proof of the proposition, we need the first sections of the eigenvectors of the transition matrix $\mathbf{\Gamma}$ in (4) with respect to the unit eigenvalues. To this aim, we directly express $\mathbf{\Gamma}$ as

$$\mathbf{\Gamma} = \begin{bmatrix} \alpha_1\beta_1' + \alpha_2\beta_2' & \begin{array}{l}\alpha_1\beta_1' - \alpha_2\beta_2' \\ +\alpha_3\beta_3'\end{array} & \alpha_1\beta_1' + \alpha_2\beta_2' & \begin{array}{l}\alpha_1\beta_1' - \alpha_2\beta_2' \\ -\alpha_3\beta_3' + \mathbf{I}\end{array} \\ \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \end{bmatrix}$$

A (left) eigenvector for the eigenvalue of +1 is then defined by the property

$$\tilde{x}\mathbf{\Gamma} = \tilde{x}$$

Partitioning $\tilde{x} = \left[\begin{array}{cccc} \tilde{x}_1 & \tilde{x}_2 & \tilde{x}_3 & \tilde{x}_4 \end{array}\right]$, we obtain the equation system

$$\begin{aligned} \tilde{x}_1\left(\alpha_1\beta_1' + \alpha_2\beta_2'\right) + \tilde{x}_2 &= \tilde{x}_1 \\ \tilde{x}_1\left(\alpha_1\beta_1' - \alpha_2\beta_2' + \alpha_3\beta_3'\right) + \tilde{x}_3 &= \tilde{x}_2 \\ \tilde{x}_1\left(\alpha_1\beta_1' + \alpha_2\beta_2'\right) + \tilde{x}_4 &= \tilde{x}_3 \\ \tilde{x}_1\left(\alpha_1\beta_1' - \alpha_2\beta_2' - \alpha_3\beta_3' + \mathbf{I}\right) &= \tilde{x}_4 \end{aligned}$$

which, summed up, yields directly

$$4\tilde{x}_1\alpha_1\beta_1 = 0 \quad \Rightarrow \quad \tilde{x}_1 \propto \alpha_1^{\perp}.$$

A similar technique for the seasonal root of -1 confirms the conjecture that the first part of the corresponding eigenvector is proportional to $\alpha_2^{\perp}$. It follows directly that the other part of the $n \times n$-dimensional space, α_2, does not influence the possibly undesirable feature of expanding trends, as it disappears after the transformation into the Jordan coordinates.

For the complex root pair of $\pm i$, we have the basic condition $\tilde{x}\mathbf{\Gamma} = i\tilde{x}$, i.e.,

$$\begin{aligned} \tilde{x}_1\left(\alpha_1\beta_1' + \alpha_2\beta_2'\right) + \tilde{x}_2 &= i\tilde{x}_1 \\ \tilde{x}_1\left(\alpha_1\beta_1' - \alpha_2\beta_2' + \alpha_3\beta_3'\right) + \tilde{x}_3 &= i\tilde{x}_2 \\ \tilde{x}_1\left(\alpha_1\beta_1' + \alpha_2\beta_2'\right) + \tilde{x}_4 &= i\tilde{x}_3 \\ \tilde{x}_1\left(\alpha_1\beta_1' - \alpha_2\beta_2' - \alpha_3\beta_3' + \mathbf{I}\right) &= i\tilde{x}_4 \end{aligned}$$

Subtracting the third from the first equation and the last from the second yields

$$\left.\begin{array}{r}\tilde{x}_2 - \tilde{x}_4 = i\left(\tilde{x}_1 - \tilde{x}_3\right) \\ 2\tilde{x}_1\alpha_3\beta_3' + \tilde{x}_3 - \tilde{x}_1 = i\left(\tilde{x}_2 - \tilde{x}_4\right)\end{array}\right\} \Rightarrow \tilde{x}_1\alpha_3\beta_3' = 0$$

and hence the proposed condition also holds for the complex pair. Note that this proof gets slightly more involved if the general seasonal model instead of specification (1) is considered. Due to arguments entirely analogous to the previous cases, only $\alpha_3^{\perp\prime}r_t$ is able to generate the possibly implausible feature of seasonally expanding trends. Hence, if $\alpha_3^{\perp\prime}(b,c) = 0$, model (3) is free of that undesirable feature. This completes the proof of our proposition. □

Some additional remarks can be made. The first is that one should note that the remaining $3n$ components of the eigenvectors — which are not used in the proof — are not trivial. For example, for the root $+1$, we obtain

$$\tilde{x} = \left[\begin{array}{cccc}\alpha_1^{\perp} & \alpha_1^{\perp}\left(\mathbf{I} - \alpha_2\beta_2'\right) & \alpha_1^{\perp}\left(\mathbf{I} - \alpha_3\beta_3'\right) & \alpha_1^{\perp}\left(\mathbf{I} - \alpha_2\beta_2' - \alpha_3\beta_3'\right)\end{array}\right].$$

This means that, if one wants to extend Granger's definition of a common trend in zero-frequency cointegrated systems to this seasonal case, $\tilde{x}'\left(X_t', X_{t-1}', X_{t-2}', X_{t-3}'\right)'$ may be preferable to simple $\alpha_1^{\perp\prime}X_t$, as the former is a multivariate random walk while the latter is not. The second remark is that our proposition can easily be generalized to higher-order systems of the form

$$\begin{aligned}\Delta_4 X_t \;=\;& \alpha_1\beta_1' S(B)X_{t-1} + \alpha_2\beta_2' A(B)X_{t-1} + \alpha_3\beta_3'\Delta_2 X_{t-2} \\ & + \sum_{i=1}^{p}\varphi_i\Delta_4 X_{t-i} + \sum_{i=1}^{4} d_i\delta^i_{t-4[(t-1)/4]} + \varepsilon_t \quad ,\end{aligned}$$

which extends (1) by the inclusion of p lags of $\Delta_4 X_t$ variables, by a straightforward extension of the proof. For more details on the Representation Theorem for the seasonal case, see Johansen and Schaumburg (1997).

2.2 An alternative empirical method

In practice one may want to test for seasonal and nonseasonal cointegration in a model framework which does not allow for diverging seasonal trends in the data, for example, simply because many macroeconomic time series do not display this feature. For that purpose, we use our proposition to re-write (3) in a simple form that permits exclusion of such diverging trends, i.e.,

$$\begin{aligned}\Delta_4 X_t &= \alpha_1\beta_1' S(B) X_{t-1} + \alpha_2(\beta_2' A(B) X_{t-1} + a^* \cos\pi(t-1)) \\ &\quad + \alpha_3\{\beta_3'\Delta_2 X_{t-2} + (b^*, c^*)\left(\cos\frac{\pi}{2}(t-1), \cos\frac{\pi}{2}(t-2)\right)\} + \mu + \varepsilon_t\end{aligned} \tag{7}$$

for $t = 1, 2, ..., N$. Note that in most empirically relevant cases a linear time trend generated by μ is perfectly admissible, as most economic time series are trending. Further note the change in dimensionality between a, b, c in (3) and a^*, b^*, c^* in (7). In (7), the row dimensions of the vectors are only r_2 and r_3 whereas in (3) they have row dimension n. Approximate maximum-likelihood estimation of (7) is straightforward and it amounts to a reduced-rank system regression of $\Delta_4 X_t$, in analogy to the traditional estimation of frequency-zero cointegrated models. Johansen (1994) points out that, in such cases, the rank of e.g. $\alpha_2\beta_2'$ is determined by the canonical correlations between $\Delta_4 X_t$ and $(A(B)X_{t-1}, \cos\pi(t-1))$, conditional on remaining influences. Hence, the right-hand side set of variables is to be extended by the deterministic influences. For the nonseasonal case this deterministic term is 1, in our case of seasonal cointegration at the bi-annual frequency it is $\cos\pi(t-1)$. Lee (1992) outlines that the three terms $S(B)X_t$, $A(B)X_t$, and $\Delta_2 X_t$ are asymptotically independent, hence tests on the various ranks can be conducted by conditioning on the complete set of remaining variables. In the absence of rank restrictions at other frequencies, conditioning can be conducted efficiently by auxiliary least squares regressions preliminary to the canonical analysis. Relative to exact maximum-likelihood estimation, a certain loss in efficiency may occur due to the fact that the rank at different frequencies may not be full but also restricted. Lee (1992) argues that this loss of efficiency is negligibly small.

Because of the independence of the three terms $S(B)X_t$, $A(B)X_t$, and $\Delta_2 X_t$, the asymptotic distribution of the likelihood-ratio test statistic for testing hypotheses on the rank of the matrices does not depend on the remaining frequencies. As a consequence, tests for the rank of $\alpha_2\beta_2'$ can be based on

$$\int_0^1 (dB)\, F \left[\int_0^1 FF\right]^{-1} \int_0^1 F\,(dB)\ ,$$

where F denotes the extended $(n-r_2)$-dimensional limit process of a process of the type

$$X_t = -X_{t-1} + \varepsilon_t \,.$$

The limit process of this X_t is, however, standard Brownian motion again. Replacing every other ε_t by $-\varepsilon_t$, this is obvious from symmetry arguments. Using the same symmetry argument, we can show that extending such a process by the alternating a_t variable is tantamount to extending the usual random walk by 1. Hence, for the frequency π, the standard table 15.2 of Johansen (1995) can be used. For the frequency $\pi/2$, however, we need a slightly different set of critical values.

Based on 10,000 Monte Carlo replications, significance points for the trace test statistic

$$\xi_i = -N \sum_{j=n-i+1}^{n} \log(1-\rho_j)$$

for the i smallest squared canonical correlations at the respective frequencies are given in the upper panels of Table 1a-f for the cases $i = 1, ..., 6$. The distribution depends on i only. If there are only two variables, i.e., $n = 2$, the situation is as follows. The statistic ξ_1 tests for the null hypothesis of integration at the respective frequency against the alternative of no integration at that frequency, while ξ_2 tests the null hypothesis of no cointegration against the alternative of cointegration or no integration. If, in the second case, the maintained hypothesis is integration at that frequency, a variant called the $\lambda_{\max}$ test should be used. It is based on the test statistic

$$\lambda_i = -N \log\left(1 - \rho_{n-i+1}\right)$$

We tabulate significance points for this $\lambda_{\max}$ test statistic in the lower panels of Tables 1b-f. For $i = 1$, the two statistics are equivalent.

The critical values in Tables 1a-f lead to a few remarks. The first is that spuriously augmenting lags (i.e., the cases where $p = 1$) tends to decrease critical values slightly at $\omega = \pi$ and $\omega = \pi/2$ but not at $\omega = 0$. Secondly we note that all cases (drift or no drift, spurious augmenting lags present or absent) produce the same asymptotic distribution for $\omega = \pi$, and a different asymptotic distribution is obtained for $\omega = \pi/2$. This property can be seen from a straightforward extension of the proofs provided by Lee (1992) and Johansen (1995). The theoretical result seems to be confirmed by the simulated fractiles. Thirdly, whereas the asymptotic distribution for $\omega = 0$ depends on the presence or absence of a drift only, lag augmentation does not seem to have much effect for smaller i but appears to dominate the dependence on the drift for larger i. Fourthly, the differences between $N = 100$ and $N = 200$ are not very pronounced. Smaller N are probably uninteresting because of the low power of the tests, larger N are unlikely to occur in quarterly economic time series. Finally, for $\omega = 0$ and $\omega = \pi$, correspondence to existing tables of simulated significance points (see Johansen, 1995, and Osterwald-Lenum, 1992) is close. An exception is the no-drift design for $\omega = 0$, where the statistics are considerably larger even for $N = 200$. Correspondence to published tables is a good indicator of the strength of finite-sample cross effects between frequencies. Lee (1992) demonstrated that these cross effects disappear for $N \to \infty$, and our results show that they also play little role for finite N, confirming the general conjectures expressed by Lee (1992) and Lee and Siklos (1995).

3 TEST POWER

To assess the power of the testing procedure which is based on restricting the seasonal constants (henceforth the RS procedure for restricted seasonals) relative to the unrestricted estimation procedure by Lee and used by Lee and Siklos (henceforth the US procedure for unrestricted seasonals), some Monte Carlo simulations are conducted. We use a bivariate and a trivariate design. Particularly in the former case (Section 3.1), we use the technique of *nuisance randomization.*

To explain nuisance randomization, let us consider a parametric problem with the parameter space $\Theta \subset \Re^{k+l}$. A null hypothesis is defined by $\theta = (\theta_1^*, \theta_2)$ with a given fixed value of the k-dimensional θ_1. We want to investigate the power of a test procedure against the alternative $\theta_1 \neq \theta_1^*$. We suspect that the power properties may depend to a certain but presumably low degree on the l-dimensional θ_2 which is usually regarded as nuisance. Faced with the problem of a possibly high-dimensional θ_2, a power simulation for the test can adopt the following two strategies. Firstly, one may fix the nuisance at a certain plausible value θ_2^* and possibly repeat the experiment for some other θ_2^* . Secondly, one may define a weighting prior distribution on the nuisance and draw from that distribution randomly. To assess the robustness of the weighting prior design, one could then repeat the experiment with different weighting priors. An advantage of nuisance randomization is the potentially exhaustive treatment of the nuisance.

The technique can easily be modified if one is interested in evaluating test power as a function of a certain nuisance parameter. Then, we vary this nuisance parameter of interest over a finite set of specified values and randomize the remaining part of θ_2. It may also be convenient to fix other parameters in θ_2 and the decision on whether to keep a certain parameter at a specified value or to randomize it will be guided by considerations of its presumed influence on the rejection rate, of the specific aim of the experiment, and of simplicity requirements.

An example where we opted for non-randomized nuisance is provided by our trivariate design (Section 3.2). Here we assume the presence of one cointegrating vector at all frequencies and test for the presence of a second one at $\omega = \pi$. It seems that, e.g., the intensity of error correction by the first vector should not be randomized as our interest rather focuses on the power properties conditional on that parameter. We could have used randomization with respect to a rotation of the vector but we felt that it would have made the design too complicated.

3.1 The bivariate simulations

We conduct a variety of simulation experiments to investigate the influence of drifts and seasonals in the DGP on finding cointegration at the three frequencies 0, π, $\pi/2$. We also explore the influence of the sample size on test power by extending the basic sample size of 100 observations to 200. In the interest of brevity we only report three representative experiments in more detail. In experiment I, we generate processes with drift and one cointegration vector at $\omega = \pi$ that operates with the intensity parameter λ explained below. Since Lee and Siklos report their significance points for data-generating processes without drift, we generated new significance points for both procedures from the same experimental design with $\lambda = 0$. We note that the two test procedures are equivalent with respect to finding cointegration at frequency 0.

For experiment I, we use the following data-generating process design:

$$\begin{bmatrix} \Delta_4 X_t \\ \Delta_4 Y_t \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \frac{-\lambda}{1-\xi_1\xi_2} \begin{bmatrix} 1 \\ \xi_2 \end{bmatrix} [-1, \xi_1] \begin{bmatrix} A(B)X_{t-1} \\ A(B)Y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \quad (8)$$

We set $\mu_1 = \mu_2 = 1$. In an unreported experiment we also tried randomization of the drift parameter but this hardly affected the results. The two nuisance parameters ξ_1 and ξ_2 are drawn from a standard Gaussian distribution. These parameters can be interpreted as random rotations of the system coordinates. Also for ε_{1t} and ε_{2t} we used 100 mutually and serially independent draws from a standard Gaussian law. The 2×2–matrix $\alpha_2 \beta_2'$ has two eigenvalues, one at 0 and one at λ, such that the "strength" of attraction to the equilibrium vector does not depend on ξ_1, ξ_2, only on λ. We use 10,000 replications and report the rejection frequency of the tests as a function of λ. For $\lambda = 0.5$ power attains almost 100%, if a 5% significance level is used for the decision. The interesting range for λ is therefore [0,0.4]. We increase λ in steps of 0.1. The resulting rejection frequencies are reported in Table 2.

By construction, for $\lambda = 0$ nominal and actual size coincide and for $\omega = 0$ both RS and US face the same rejection rate. Both procedures have a satisfactory power at the frequency π and both face only slight size distortions at the other frequencies, particularly for larger λ and for $\omega = \pi/2$. The RS procedure is slightly more vulnerable to size distortions than the US procedure but, on the whole, differences remain small. We also find that the power in the direction of cointegration at $\omega = \pi$ is very similar to the power at $\omega = 0$ for the case of no drift. Since there is no seasonal drift, this outcome is plausible.

In experiment IIa, we introduce deterministic seasonality in the data-generating process. As we do not believe in the presence of unstable expanding seasonal cycles in the real world — this was the original motivation for introducing our procedure — we exclude their occurrence in the simulation design. In other words, in these experiments we evaluate if the introduction of restricted seasonals leads to a change in power and/or to a change in size. We use

$$\begin{aligned} \begin{bmatrix} \Delta_4 X_t \\ \Delta_4 Y_t \end{bmatrix} &= \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \lambda_1 \begin{bmatrix} 1 \\ \xi_2 \end{bmatrix} \cos(\pi t) \\ &\quad + \frac{-\lambda_2}{1 - \xi_1 \xi_2} \begin{bmatrix} 1 \\ \xi_2 \end{bmatrix} [-1, \xi_1] \begin{bmatrix} A(B) X_{t-1} \\ A(B) Y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \qquad (9) \end{aligned}$$

λ_2 measures the intensity of seasonal error correction at the frequency $\omega = \pi$, as λ does in Experiment I. λ_1 measures the influence of the deterministic seasonal term $\cos(\pi t)$. The case $\lambda_1 = \lambda_2 = 0$ corresponds to the null hypothesis of the test, whereas $\lambda_2 > 0$ describes the alternative irrespective of whether $\lambda_1 = 0$ or $\lambda_1 \neq 0$. The case $\lambda_1 \neq 0, \lambda_2 = 0$ leads to an implausible seasonal expansion of the process trajectories and is not investigated. If $\lambda_2 > 0$, the influence of the seasonal constants is contained in the cointegrating relationship. If $\lambda_2 = 0$ the condition of being contained in the loading space cannot be fulfilled by the deterministic seasonal part.

Although we do not report results on the misspecified case $\lambda_2 = 0, \lambda_1 \neq 0$, this case deserves attention, as it influences test behavior in the adjacent region, in particular where λ_2 is small, λ_1 is large, and N is small. In large samples, the RS test responds to the misspecification by rejection. Note that the misspecified boundary case does not make part of our null hypothesis and hence the rejection rate of the test procedures should not be interpreted as test size. The US test is expected to ignore the peculiarity of the situation and to seemingly accept seasonal non-cointegration in this non-admissible case. In contrast, the RS procedure, conditioning out only part of the seasonal structure, seemingly finds (spurious) seasonal cointegration at frequency π.

An analytic demonstration helps in understanding the behavior of the RS test at $\lambda_2 = 0, \lambda_1 \neq 0$ in this experiment IIa. Suppose we are given the univariate process

$$\Delta_4 X_t = \left[\mu + (b, c)\left(\cos\left(\pi t/2\right), \cos\left(\pi\left(t+1\right)/2\right)\right)\right] + a\cos\left(\pi t\right) + \varepsilon_t$$

with ε_t assumed as i.i.d. N(0,1). The terms in square brackets do not affect the limit distribution and we will omit them for convenience. In the univariate case, the canonical correlation for the test corresponds to the squared correlation between $\Delta_4 X_t$ and $A(B)X_{t-1}$, conditional on $S(B)X_{t-1}$, $\Delta_2 X_{t-1}$, $\Delta_2 X_{t-2}$ and the deterministic terms at frequencies 0 and $\pi/2$. We first note that (defining S_t)

$$A(B)X_t = a(-1)^t t + \sum_{i=0}^{t} (-1)^i \varepsilon_{t-i} = a(-1)^t t + S_t \quad .$$

Then we note that conditioning does not affect $\Delta_4 X_t$ nor $A(B)X_{t-1}$, and hence that all regression coefficients on conditioning variables converge to 0. The only exception is the constant term which corrects $A(B)X_t$ by a periodically changing constant that is proportional to a and by the empirical mean of the stochastic random walk part S_t. The squared sample correlation evolves from dividing the square of the cross-sums

$$\begin{aligned}&\sum \Delta_4 X_t \left(A(B)X_{t-1} - (N-1)^{-1} S_{N-1}\right) \\ &= \sum \left(a\cos(\pi t) + \varepsilon_t\right)\left(a(-1)^{t-1}(t-1) + S^*_{t-1}\right) \quad ,\end{aligned}$$

with * denoting the demeaned sum, by the product of the sums of squares. For the numerator we note

$$\begin{aligned}\sum a\cos(\pi t) a(-1)^{t-1}(t-1) &= -\frac{a^2N^2}{2} + O_p(N) \\ \sum a\cos(\pi t) S^*_{t-1} &= a\sum (N-t)\varepsilon_t = O_p(N)\end{aligned}$$

The first term dominates and we have established that the (squared) numerator diverges at a rate of $N^4(a^4/4)$. As for the denominator sums of squares, the first one is

$$\sum \left(a\cos(\pi t) + \varepsilon_t\right)^2 = N\left(a^2+1\right) + o_p\left(N^{1/2}\right)$$

and the other one is

$$\sum \left(a(-1)^{t-1}(t-1) + S^*_{t-1}\right)^2 = a^2N^3/3 + O_p(N^{5/2})$$

It follows that the ratio converges to

$$\frac{3}{4(1+a^{-2})} \tag{10}$$

The correlation is therefore non-zero and the test, which is consistent under the alternative, rejects asymptotically with probability 1. Unreported Monte Carlo simulations show that the asymptotic value in (10) is fairly accurate for $N = 100$ and $a > 0.5$. In the multivariate cointegration problem, the asymptotic behavior in the presence of an uncontrolled expansion is very similar. Hence, “cointegration” is typically found in the multivariate and “stationarity” in the univariate case.

Figure 1 compares the test power of both procedures in the interesting area $\{(\lambda_1, \lambda_2) \in [0, 0.8] \times [0.025, 0.4]\}$ at a nominal significance level of 5% in the direction of cointegration at frequency $\omega = \pi$. We see that the conjectures are corroborated by the simulation. This feature is well in line with the basic idea of the RS test. As the RS test rejects in the non-admissible limit for $\lambda_2 = 0$, it has higher power for small λ_2 and large λ_1. We used a finer grid than for Experiment I and Table 2. The considerable gains in power by using RS instead of US for small λ_2 come at the cost of a slightly lower power for large deviations from the null. We further note that also for the US test, which formally would treat the inadmissible case $\lambda_1 \neq 0, \lambda_2 = 0$ as belonging to the null hypothesis, the rejection rate does not correspond to the nominal size. For $\lambda_2 < 0.1$, the test power of the US test increases considerably as the deterministic influence λ_1 increases, which for this unrestricted test indicates that the deterministic seasonal structure is eliminated incompletely due to the small sample. Interestingly, in stark contrast to the RS test, we found that the rejection rate of the US test in the unstable case is even slightly smaller than the size. This inadmissible boundary case was excluded from the figures in order not to distort the picture unduly..

Experiment IIa appears to be particularly interesting as it provides clear evidence on differences among the two test procedures. Therefore we repeat it for the larger sample size of 200 observations (experiment IIb) with randomized deterministic seasonal influences λ_1. The parameter λ_1 is drawn from a standard N(0,1) distribution, hence test power depends on $\lambda = \lambda_2$ only. This allows reporting the results in tabular form. To this aim, we simulated new critical values from (7) for a constant drift of 1.0 and $N = 200$. In Table 3 we report the results for a finer λ grid around 0 than in Experiments I and IIa in order to focus on the interesting features.

In the limiting case $\lambda = 0$ we now have a higher rejection rate for the RS test but no important changes for the US test. We note the better test power of the RS procedure for all $\lambda < 0.1$. Its rejection frequency reaches a minimum somewhere in the range [0.025,0.5], depending on the significance level. For $\lambda = 0.1$ and $\lambda = 0.01$, the rejection frequencies are the same at a risk level of 1%. However, for $\lambda = 0.01$ the distribution of the RS statistic is considerably flatter toward its distributional center. On the other hand, for $\lambda = 0.1$ both the RS and US statistic obey the laws dictated by the Ornstein-Uhlenbeck process (see Phillips, 1988, and Johansen, 1995, Ch.14). We also note that the small differences in test size between the two tests have now disappeared but that a slight gain in power is still there for larger λ if the US procedure is used.

In other unreported experiments we focus on the test power in the direction of cointegration at the frequency $\pi/2$. The design parallels the reported experiments I and II but, in accordance with the theoretical part of our paper we only allow for cointegration of the static type at lag 2. Static cointegration at lag 1 is not possible and we did not want to extend our analysis and methods in the direction of dynamic cointegration. It turned out that test power in the direction of cointegration at frequency $\pi/2$ is much lower than at the frequencies 0 or π. Detailed results can be obtained from the authors.

Finally, for experiment III, we augmented the seasonal cointegration model at $\pi/2$ by a non-expanding seasonal pattern:

$$\begin{aligned} \begin{bmatrix} \Delta_4 X_t \\ \Delta_4 Y_t \end{bmatrix} &= \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \xi_3 \begin{bmatrix} 1 \\ \xi_2 \end{bmatrix} \cos(\pi t/2) \\ &\quad + \frac{-\lambda}{1-\xi_1\xi_2} \begin{bmatrix} 1 \\ \xi_2 \end{bmatrix} [-1, \xi_1] \begin{bmatrix} \Delta_2 X_{t-2} \\ \Delta_2 Y_{t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \qquad (11) \end{aligned}$$

We could add another seasonal part based on the shifted cosine function $\cos(\pi(t-1)/2)$ but this is unlikely to change the main results. As in experiment II, the boundary case $\lambda = 0$ is not data-admissible. The rejection frequencies are displayed in Table 4. It appears that the RS test has higher power in the whole reported range for λ.

3.2 The trivariate simulations

In the trivariate experiment we assume the presence of one cointegrating vector at all frequencies $\omega = 0, \pi, \pi/2$ and we test for the presence of a second vector at $\omega = \pi$. In contrast to the bivariate experiment, the restriction imposed by the RS procedure is then non-trivial both under the null hypothesis and under the alternative. Therefore we expect a larger difference between the two testing procedures.

In detail, the following design was used:

$$\begin{aligned} \begin{bmatrix} \Delta_4 X_{1t} \\ \Delta_4 X_{2t} \\ \Delta_4 X_{3t} \end{bmatrix} &= \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} \\ &\quad + \begin{bmatrix} 1 \\ 0.5 \\ 0 \end{bmatrix} \begin{bmatrix} -1 & 1 & 0 \end{bmatrix} (S(B) + A(B) + B\Delta_2) \begin{bmatrix} X_{1,t-1} \\ X_{2,t-1} \\ X_{3,t-1} \end{bmatrix} \end{aligned}$$

$$+\begin{bmatrix} \lambda \\ 0 \\ \lambda \end{bmatrix} \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} A(B) \begin{bmatrix} X_{1,t-1} \\ X_{2,t-1} \\ X_{3,t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \\ \epsilon_{3t} \end{bmatrix} \quad (12)$$

This design results in the rejection frequencies reported in Table 5.

All size biases remain small but the RS procedure is clearly superior. We note that the US procedure has almost zero power for $\lambda = 0.01$ whereas RS would reject in 23% of the cases if a 5% significance point were used. Similar differences can be noted for $\lambda = 0.02$. We also modified design (11) by adding deterministic seasonal structures but this had no impact on the results.

3.3 Empirically determined lag structures

In practice, the lag length p is unknown and has to be determined from the data, unless economic theory suggests such a p, which is rarely the case. Supposing that there is a true finite lag length p^* that has generated the data under investigation, such empirical order selection can make mistakes of two kinds. Firstly, the estimated $\hat{p}$ may be smaller than p^*. This should result in insufficient conditioning on short-run influences $\Delta_4 X_{t-i}$ and may entail a tendency toward blown-up values of canonical correlations. This may result in a positive size bias. Secondly, $\hat{p}$ may be larger than p^*, which, by conditioning on too many terms in the preliminary step of calculating the test statistic, may lead to under-rejection. Thirdly, sampling variation and small-sample distortions in all coefficient estimates may cause size-bias effects in any direction. Also the first two effects are uncertain *a priori*, as cases of underfitting and overfitting may be related to certain combinations of the coefficient parameters or to certain characteristics of trajectories in small samples whose effects are difficult to assess.

The experiments reported here are to be interpreted with care only, mainly for two reasons. Firstly, the real-life economic variables are unlikely to have been generated by Gaussian vector autoregressions with any finite p^*. Assuming p^* to grow with N does not capture the full problem either, as empirical series typically suffer from a variety of deviations from the Gaussian VAR framework, such as non-Gaussian errors, local outliers, change in structure, non-linearities etc. The Gaussian VAR framework can be seen as an approximation to real life at best. Secondly, even within the Gaussian VAR framework, rejection frequencies in small samples depend on the complete vector of model parameters. This study can only provide a convenient summary of the effects, based on assumed parameter configurations which need not correspond to good approximations to the observed data. We therefore cannot recommend to make arbitrary adjustments to the significance points as presented in Tables 1a-f based on the simulations reported in the following.

We base the simulations on the empirically determined lag orders on a variant of Experiment IIa. We extend the model design (9) by two short-run lags ($p = 2$):

$$\begin{aligned} \begin{bmatrix} \Delta_4 X_t \\ \Delta_4 Y_t \end{bmatrix} &= \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \lambda_1 \begin{bmatrix} 1 \\ \xi_2 \end{bmatrix} \cos(\pi t) \\ &+ \frac{-\lambda_2}{1 - \xi_1 \xi_2} \begin{bmatrix} 1 \\ \xi_2 \end{bmatrix} [-1, \xi_1] \begin{bmatrix} A(B) X_{t-1} \\ A(B) Y_{t-1} \end{bmatrix} \\ &+ \sum_{i=1}^{2} \begin{bmatrix} \phi_{1i} & 0 \\ 0 & \phi_{2i} \end{bmatrix} \begin{bmatrix} \Delta_4 X_{t-i} \\ \Delta_4 Y_{t-i} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \end{aligned} \quad (13)$$

Note that the coefficient matrices are diagonal. We randomize the nuisance parameters $\phi_{11}, \phi_{12}, \phi_{21}, \phi_{22}$ on the basis of uniform draws from the stability area for univariate second-order regressions

$$S_2 = \{(\phi_{i1}, \phi_{i2} | \phi_{i2} > -1, \phi_{i1} + \phi_{i2} < 1, \phi_{i2} - \phi_{i1} < 1\} \quad (14)$$

where $i = 1, 2$ (cf. Box et al., p.61). Not all combinations of parameters thus generated yield seasonally integrated but otherwise stable models in conjunction with the seasonally cointegrating structures. Therefore, the absence of explosive roots had to be checked separately. Explosive cases were rejected and re-drawing was conducted, hence the effective design does not exactly correspond to uniform drawings from S_2.

Akaike's information criterion AIC is maybe the most frequently used criterion for empirical lag order selection, hence we used it for our first experiment. In detail, lag orders were selected from 'level' vector autoregressions between $\tilde{p} = 4$ and $\tilde{p} = 8$, which appears natural for quarterly data and a sample size of $N = 100$. Note that lag orders less than $\tilde{p} = 4$ are not possible. Added deterministic terms were a constant and the seasonal variables $\cos(\pi t)$ and $\cos(\pi t/2), \cos(\pi(t-1)/2$. All lags and deterministic regressors were added in an unrestricted way, hence the true data generation mechanism is strictly embedded in the search design and the identified lag orders are independent of the testing method used in the sequel.

For the null-hypothesis model with $\lambda_1 = \lambda_2 = 0$, the correct number of $p = 2$ was identified in 69% of the cases. In 1% of the cases, $p = 0$ was found and hence no lagged differences would be added, in 5% of the cases $p = 1$ was found and one difference would be required. Hence, underfitting was relatively infrequent, as was expected from the asymptotically inconsistent AIC criterion. On the other hand, one spurious lag was suggested in 15% of the cases and even two spurious lags in 10%. At first sight, the achieved precision appears to be satisfactory in a sample of $N = 100$. The frequencies of identified lag orders are widely unaffected by changes in λ_1 and λ_2 and are therefore valid for the models under the alternative also.

Unfortunately, the AIC search leads to a substantial size bias if the suggested lag orders are used in seasonal cointegration testing. At the frequency $\omega = 0$, the tests overreject. This rate of overrejection is the same for both the RS and US tests and is also widely unaffected by changes of λ_1 and λ_2 on the alternative. One may conjecture that the 6% of the cases, where underfitting occurs, are mainly responsible for this effect. Indeed, any artificial increase in underfitting, as by restricting the maximum lag order by $\tilde{p} \leq 6$, increases the size bias. At the frequency $\omega = \pi$, the RS test reproduces the correct rejection rates, whereas the US test again overrejects. The additional unrestricted conditioning on seasonal deterministic terms in the US test appears to generate this bias. At the frequency $\omega = \pi/2$, both tests underreject. This negative size bias is worse for the RS test than for the US test. This effect is only partially due to the 25% cases of overfitting. It rather seems to reflect the high probability of complex roots in the simulation design (13). The presence of stable complex roots in the data generation mechanism poses an empirically relevant problem to seasonal unit-root and seasonal cointegration tests. It is worse at frequency $\pi/2$ than at π, as only 25 full 'annual' cycles are observed in a sample of 100 observations. In summary, incorrect specification of lag orders leads to a variety of counteracting effects. We point out once more that these effects are relatively robust with respect to the introduction of cointegration at one of the frequencies.

We now tentatively re-adjust the 5% significance levels on the basis of the AIC-search levels in the experiment with $(\lambda_1, \lambda_2) = (0, 0)$ and report the relative power of the two tests at $\omega = \pi$. Figure 2 can be compared with Figure 1 but note that larger values of the parameter λ_1 are now investigated. Rejection frequencies are hardly affected for $\lambda_1 = 0$ and small λ_2, i.e. near the null hypothesis, whereas there is a considerable drop in power for larger

λ_2, i.e. stronger seasonal error correction. The most striking effect of the AIC lag order determination is the drop in power in the RS test for strong deterministic seasonality with large λ_1. The US test dominates the rival test over a large area of the parameter space $\{(\lambda_1, \lambda_2) \in [0, 1.5] \times (0, 0.4]\}$, excluding only λ_1 values around 1.0 and very small values of λ_2. The effect is partially rooted in the strong drift of 1.0 in the simulations and gets slightly weaker if the drift is reduced or randomized (unreported experiments).

Another commonly used information criterion is the BIC in the version defined by Schwarz (1978). Unlike the AIC, BIC is consistent. In smaller samples, BIC has a strong tendency to identify too parsimonious models and hence AIC is preferred in applications where a slight overparameterization appears to be the lesser risk. In our simulation design (13), BIC achieves a much higher frequency of correct model selection. For $(\lambda_1, \lambda_2) = (0, 0)$, 76% of all 10,000 processes are detected correctly at $p = 2$. In 14% of the cases, $p = 1$ is selected, and in 9% of the cases, $p = 0$ is chosen. The probability of fitting a too large model with $p = 3, 4$ remains extremely low. This overall picture hardly changes when the main parameters λ_1 and λ_2 are varied. In summary, seasonal cointegration poses no threat to lag order determination and BIC has a much bigger chance of finding the correct model. We add that some of the randomized autoregressive coefficients have been close to 0 and that therefore some underfitting of lag orders appears very reasonable.

The effect of the BIC search on test power for the RS and US tests can be seen in Figure 3. For most combinations (λ_1, λ_2), test power increases considerably relative to the AIC search displayed in Figure 2. The gain in power is stronger for the RS than for the US test. The RS test dominates clearly for λ_1 around 1 and small λ_2, i.e., if the deterministic seasonal influence is of a magnitude similar to the drift and innovations variance. The US test dominates for larger λ_1 and λ_2 around 0.2. This dominance for large seasonal error correction and strong deterministic cycles is a paradox and is rooted in the inefficient preliminary estimation of the deterministic seasonality in the US test procedure. The restriction $\alpha_2^{\perp\prime} a = 0$ is ignored and the inefficient primary regression entails a higher rejection rate. Correct identification of the lag length reduces this relative effect but we note that it is still present at a sample size of $N = 200$.

With $N = 200$, BIC finds the correct $p = 2$ in 89% of the cases. Rejection frequencies of both the US and RS tests are very similar for most combinations of (λ_1, λ_2). RS has higher power for λ_1 around 0.5 but US preserves its dominance if both λ_1 and λ_2 are large. The unpleasant lesson from this unreported experiment is that there is a persistent negative size bias at the frequency $\omega = \pi/2$. Increasing λ_1 and λ_2 reduces this size bias effect but the effect is present in all experiments even for larger N. The effect hits the US and RS test similarly and is probably rooted in finite-sample phase shifts of the cointegrating structures at the annual frequency. Note that in our model only correlations between $\Delta_2 X_{t-2}$ and $\Delta_4 X_t$ are investigated whereas the information on correlations between $\Delta_2 X_{t-1}$ and $\Delta_4 X_t$ is ignored. This causes underrejection at fairly large sample sizes, although we are aware of the asymptotic results of Lee (1992) who (implicitly) proves that both the US and RS test have correct size for $N \to \infty$ if p is the true value.

4 AN APPLICATION

As an empirical example, we use a macroeconomic system of four time series from the Austrian national accounts: gross domestic product (Y), private consumer expenditures (C), gross fixed investment (I), and gross wages (W). All variables are in real terms and logarithms of original series have been used. We used quarterly data from 1964.1 to 1994.4. This data extends a subset of the data set used originally by Kunst (1993). The addition of one conditioning lag was suggested by AIC and also by some goodness-of-fit statistics. Among these four variables, economic theory may suggest two or three cointegrating vectors at frequency 0, representing fixed long-run proportions of consumption and investment to total output and also wages growing in proportion to output. However, it has already been found by Kunst and Neusser (1990) that these ratios do not seem to be constant. Also, all four variables show very strong seasonality that appears to be changing over time and which may be captured by seasonal unit roots.

4.1 Testing for cointegration

An application of the standard seasonal cointegration tests in the spirit of Lee (1992) and Lee and Siklos (1995) results in the evidence displayed in Table 6. For the no-dummies case we use the table 15.3 of Johansen (1995) for the frequency 0 and those of Lee and Siklos for the other frequencies, which results in 1 nonseasonal cointegrating vector at 10% risk level, 2 cointegrating vectors at frequency π at 5% risk level, and also 2 cointegrating vectors at frequency $\pi/2$ but only at a 10% risk level. Also for the dummies case in the second part of Table 6 one must use the table 15.3 of Johansen for the frequency 0 and those of Lee and Siklos for the other frequencies. We again find 1 nonseasonal cointegrating vector at 10% risk level, whereas the numbers of cointegrating vectors at $\omega = \pi$ and $\omega = \pi/2$ both increase to 3 at a 5% risk level. Finally, for the restricted dummies case, we use the Tables 1a-d. At $\omega = 0$, no cointegrating vector is found though the test statistic comes close to the 10% fractile. At $\omega = \pi$, the third cointegrating vector is now significant at 10% only, whereas still 3 vectors are found at $\omega = \pi/2$ at the usual risk level of 5%. It is obvious that restricting the seasonal effects plays a role for testing. The cointegrating vector at $\omega = 0$ became insignificant and the third vector at $\omega = \pi$ also lost support. We note that the first effect is due to the modification of significance points and it occurs even though the trace statistic at $\omega = 0$ remains unchanged.

4.2 Estimating the cointegrating structures

Restricting the influence of seasonal constants does not only change the identified cointegrating ranks at the different frequencies, it also changes the estimated cointegrating vectors. In Table 7 we compare the canonical vectors found by the US and by the RS procedures. We note that the RS vectors are extended by the deterministic seasonal variables. With respect to the identified significance of the structures as cointegrating vectors, we refer to Table 6 and to the last subsection. This change in the identified seasonal structures can be quite important for forecasting projections.

As in other empirical examples, the canonical vectors are difficult to interpret economically. However, we see that the first vectors at both frequencies $\omega = \pi$ and $\pi/2$ appear to link seasonality in wages to seasonality in output (Y). The other two significant vectors then relate all four variables in such a way that there seems to exist just a single source of stochastic seasonality in the system. With regard to the RS method, it is also interesting to see the comparatively strong influence of the deterministic part in the third vectors. This can be interpreted as follows. For example at $\omega = \pi$, the first two cointegrating relationships succeed in establishing variables with both stochastic seasonality at $\omega = \pi$ and their deterministic cycle $\cos(\pi t)$ completely absent. On the other hand, the third cointegrating relationship defines a variable that has a very strong deterministic cycle of the form $\cos(\pi t)$. Roughly the same argument holds also for $\omega = \pi/2$.

5 CONCLUDING REMARKS

In this paper we have shown that unrestricted seasonal intercepts in a seasonal cointegration model can lead to diverging seasonal trends. Since such seasonal trends may be implausible in certain practical occasions, we proposed an alternative empirical method to investigate seasonal cointegration where we impose restrictions on the seasonal intercept parameters. We tabulated critical values for the various test statistics in the case of one to six variables. A comparison of the standard method and our proposed restricted method to four quarterly macroeconomic aggregates for Austria was used to show that our method can lead to different estimates of the rank of long-run matrices and to different estimates of the cointegrating vectors.

The analysis in the present paper can be extended in at least two ways. Firstly, it seems interesting to re-evaluate earlier empirical studies of seasonal cointegration to investigate the robustness of the findings to the restriction of seasonal dummy parameters. Indeed, the finding of more or less similar results across the two methods may yield additional confidence in the reported outcomes. Secondly, since estimated ranks and cointegration vectors can differ across the methods applied, it seems relevant to study the forecasting performance of both approaches in practice. A first attempt of such a study appears in Kunst and Franses (1997).

REFERENCES

Box, G.E.P., Jenkins, G.M., and Reinsel, G.C. (1994). *Time Series Analysis: Forecasting and Control*. Prentice-Hall.

Chan, N.H. and Wei, C.Z. (1988). 'Limiting distributions of least squares estimates of unstable autoregressive processes', *Annals of Statistics* Vol. 16, pp. 367-401.

Engle, R.F. and Granger, C.W.J. (1987). 'Co-integration and error correction: Representation, estimation and testing', *Econometrica* Vol. 55, pp. 251-276.

Engle, R.F., Granger, C.W.J., Hylleberg, S. and Lee, H.S. (1993). 'Seasonal cointegration: The Japanese consumption function', *Journal of Econometrics* Vol. 55, pp. 275-298.

Hylleberg, S., Engle, R.F., Granger, C.W.J. and Yoo, B.S. (1990). 'Seasonal integration and cointegration', *Journal of Econometrics* Vol. 44, pp. 215-238.

Johansen, S. (1994). 'The role of the constant and linear terms in cointegration analysis of nonstationary variables', *Econometric Reviews* Vol. 13, pp. 205-229.

Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press.

Johansen, S., and Schaumburg, E. (1997). 'Likelihood Analysis of Seasonal Cointegration', EUI Working Paper ECO No. 97/16, European University Institute, Florence.

Kunst, R.M. (1993). 'Seasonal Cointegration in Macroeconomic Systems: Case Studies for Small and Large European Countries', *Review of Economics and Statistics* LXXV, 325-330.

Kunst, R.M., and Franses, P.H.F. (1997). 'The impact of seasonal constants on forecasting seasonally cointegrated time series', *Journal of Forecasting* 17, 109–124.

Kunst, R.M., and Neusser, K. (1990). 'Cointegration in a Macroeconomic System', *Journal of Applied Econometrics* Vol.5, 351-365.

Lee, H.S. (1992). 'Maximum likelihood inference on cointegration and seasonal cointegration', *Journal of Econometrics* Vol. 54, pp. 1-47.

Lee, H.S., and Siklos, P. (1995). 'A note on the critical values for the maximum likelihood (seasonal) cointegration tests', *Economics Letters* 49, 137-145.

Lütkepohl, H. (1991). *Introduction to multiple time series analysis.* Springer Verlag, Berlin.

Osterwald-Lenum, M. (1992). 'A note with quantiles of the asymptotic distribution of the maximum likelihood cointegration rank statistics: Four cases', *Oxford Bulletin of Economics and Statistics* Vol. 54, pp. 461-472.

Phillips, P.C.B. (1988). 'Regression Theory for Near-Integrated Time Series,' *Econometrica* 56, 1021-1044.

Schwarz, G. (1978). 'Estimating the dimension of a model,' *Annals of Statistics* 6, 461–464.

Tsay, R.S. and Tiao, G.C. (1990). 'Asymptotic properties of multivariate nonstationary processes with applications to autoregressions', *Annals of Statistics* 18, 220-250.

TABLES

Table 1a
Simulated significance points (based on 10,000 Monte Carlo replications) for trace test statistics ξ_1 under the restriction of no diverging seasonal trends, i.e., model (7) in text. The data generating process is $\Delta_4 X_t = \mu + \varepsilon_t$, where ε_t is univariate Gaussian with $\sigma^2 = 1$.

			$\omega = 0$			$\omega = \pi$			$\omega = \pi/2$		
N	p	μ	.90	.95	.99	.90	.95	.99	.90	.95	.99
100	0	0	6.5	8.1	11.3	7.4	8.9	12.4	10.8	12.8	16.5
100	1	0	6.6	8.3	11.9	7.2	8.6	11.9	11.1	12.9	17.5
100	0	1	2.8	3.8	7.1	7.3	8.8	12.1	10.8	12.6	16.4
100	1	1	2.9	4.1	7.3	7.2	8.7	12.1	11.0	12.9	17.3
200	0	0	6.7	8.2	11.5	7.4	9.0	12.4	11.0	12.9	17.0
200	1	0	6.7	8.4	11.8	7.3	8.8	12.3	11.2	13.0	17.0
200	0	1	2.7	3.9	6.4	7.4	9.0	12.4	11.0	12.8	17.0
200	1	1	2.8	4.0	6.7	7.4	8.9	12.4	11.1	12.9	17.1

Notes: The ξ_1 test statistic considers the hypothesis of integration at frequency ω against the alternative of no integration at that frequency. N is the number of observations, p denotes whether the auxiliary regression (7) includes an additional lag of $\Delta_4 X_t$ (1) or not (0), and μ is the constant term in the DGP.

Table 1b

Simulated significance points (based on 10,000 Monte Carlo replications) for trace test statistics ξ_2 and eigenvalue statistic λ_2 under the restriction of no diverging seasonal trends, i.e., model (7) in text. The data generating process is $\Delta_4 X_t = \mu(1,1)' + \varepsilon_t$, where ε_t is bivariate Gaussian with $\Sigma = \mathbf{I}$.

				$\omega = 0$			$\omega = \pi$			$\omega = \pi/2$	
N	p	μ	.90	.95	.99	.90	.95	.99	.90	.95	.99
trace statistic ξ_2											
100	0	0	16.0	18.2	23.1	17.9	20.1	24.8	23.9	26.5	31.9
100	1	0	16.4	18.8	23.8	17.0	19.2	23.4	23.3	25.8	31.0
100	0	1	13.6	15.5	20.1	17.8	20.1	24.8	23.9	26.4	31.7
100	1	1	13.9	16.3	21.1	17.0	19.2	23.7	23.3	25.8	30.8
200	0	0	16.0	18.3	22.8	17.9	20.1	24.9	24.0	26.3	31.8
200	1	0	16.1	18.4	23.2	17.2	19.4	23.6	23.7	25.9	30.9
200	0	1	13.3	15.5	19.8	17.9	20.1	24.8	23.9	26.3	31.8
200	1	1	13.7	15.7	20.3	17.2	19.3	23.8	23.6	25.9	31.0
eigenvalue statistic λ_2											
100	0	0	13.3	15.1	18.8	13.7	15.7	19.5	17.3	19.4	23.8
100	1	0	13.4	15.5	20.9	13.1	15.1	18.8	16.9	18.9	23.4
100	0	1	12.4	14.3	18.6	13.8	15.7	19.6	17.3	19.3	23.8
100	1	1	12.6	14.8	19.3	13.2	15.1	18.7	16.8	18.9	23.4
200	0	0	13.1	15.0	19.5	13.8	15.8	20.0	17.4	19.5	23.9
200	1	0	13.2	15.2	19.7	13.3	15.2	18.9	17.1	19.1	23.3
200	0	1	12.1	14.1	18.3	13.7	15.8	20.0	17.4	19.5	23.9
200	1	1	12.3	14.3	18.9	13.3	15.2	18.9	17.1	19.2	23.2

Notes: The ξ_2 test statistic considers the hypothesis of $r \leq n-2$ at frequency ω against the alternative of $r > n-2$ at that frequency. The λ_2 test statistic considers the alternative $r = n-1$ at frequency ω. N is the number of observations, p denotes whether the auxiliary regression (2) includes an additional lag of $\Delta_4 X_t$ (1) or not (0), and μ is the constant term in the DGP.

Table 1c
Simulated significance points (based on 10,000 Monte Carlo replications) for trace test statistics ξ_3 and eigenvalue test statistic λ_3 under the restriction of no diverging seasonal trends, i.e., model (7) in text. The data generating process is $\Delta_4 X_t = \mu(1,1,1)' + \varepsilon_t$, where ε_t is 3–variate Gaussian with $\Sigma = \mathbf{I}$.

			$\omega = 0$			$\omega = \pi$			$\omega = \pi/2$		
N	p	μ	.90	.95	.99	.90	.95	.99	.90	.95	.99
trace statistic ξ_3											
100	0	0	29.7	32.8	39.1	32.1	34.9	41.0	40.2	43.3	49.7
100	1	0	30.9	34.0	40.8	30.9	33.7	39.8	38.3	41.3	47.8
100	0	1	27.7	30.4	36.5	32.0	34.9	40.7	40.2	43.2	49.8
100	1	1	29.0	31.9	37.6	30.9	33.6	39.8	38.3	41.2	47.8
200	0	0	29.3	32.5	38.3	32.2	35.3	40.9	40.7	44.0	50.7
200	1	0	29.9	32.8	38.6	31.0	33.8	39.8	39.4	42.7	49.0
200	0	1	27.3	30.0	35.5	32.2	35.3	40.8	40.7	44.1	50.4
200	1	1	27.8	30.6	35.8	31.1	33.8	40.0	39.4	42.6	48.9
eigenvalue statistic λ_3											
100	0	0	19.9	22.2	27.4	19.9	22.3	27.5	23.4	25.7	30.5
100	1	0	20.7	23.1	28.7	19.3	21.6	27.3	22.2	24.4	29.4
100	0	1	19.4	21.8	26.1	19.9	22.2	27.4	23.3	25.6	30.5
100	1	1	20.2	22.6	28.0	19.4	21.6	27.2	22.2	24.5	29.4
200	0	0	19.3	21.5	26.3	20.1	22.3	26.9	23.7	26.0	31.0
200	1	0	19.7	22.0	26.8	19.3	21.4	26.3	22.9	25.0	29.8
200	0	1	19.0	21.1	25.9	20.0	22.2	26.8	23.7	26.0	30.8
200	1	1	19.3	21.7	26.1	19.3	21.5	26.6	22.8	25.0	29.8

Notes: The ξ_3 test statistic considers the hypothesis of $r \leq n-3$ at frequency ω against the alternative of $r > n-3$ at that frequency. The λ_3 test statistic considers the alternative $r = n-2$ at frequency ω. N is the number of observations, p denotes whether the auxiliary regression (2) includes an additional lag of $\Delta_4 X_t$ (1) or not (0), and μ is the constant term in the DGP.

Table 1d

Simulated significance points (based on 10,000 Monte Carlo replications) for trace test statistics ξ_4 and for eigenvalue test statistic λ_4 under the restriction of no diverging seasonal trends, i.e., model (7) in text. The data generating process is $\Delta_4 X_t = \mu(1,1,1,1)' + \varepsilon_t$, where ε_t is 4–variate Gaussian with $\Sigma = \mathbf{I}$.

			$\omega = 0$			$\omega = \pi$			$\omega = \pi/2$		
N	p	μ	.90	.95	.99	.90	.95	.99	.90	.95	.99
trace statistic ξ_4											
100	0	0	48.3	51.9	59.9	51.0	54.6	61.6	61.1	64.8	71.7
100	1	0	50.4	54.2	61.6	49.5	53.1	60.2	57.6	61.5	69.1
100	0	1	46.5	50.2	57.2	51.0	54.4	61.5	61.0	64.9	72.0
100	1	1	48.4	52.5	60.0	49.5	53.0	59.7	57.6	61.4	69.5
200	0	0	47.0	50.4	57.2	50.8	54.3	61.7	60.9	64.8	72.8
200	1	0	48.0	51.7	59.2	48.9	52.0	58.6	58.9	62.3	69.3
200	0	1	45.1	48.4	54.9	50.8	54.3	61.7	60.9	64.8	72.6
200	1	1	46.0	49.3	56.3	48.8	51.9	58.4	58.8	62.4	69.2
eigenvalue statistic λ_4											
100	0	0	26.9	29.7	35.4	26.4	29.0	34.9	29.8	32.2	36.8
100	1	0	28.0	31.0	36.8	26.2	28.9	34.3	28.3	31.0	35.8
100	0	1	26.6	29.1	35.0	26.4	29.0	34.7	29.6	32.1	37.0
100	1	1	27.6	30.6	36.7	26.1	28.9	34.3	28.2	30.8	35.9
200	0	0	25.8	28.3	33.5	26.2	28.6	33.9	29.8	32.3	38.2
200	1	0	26.5	29.1	34.6	25.3	27.6	33.3	28.6	31.0	36.1
200	0	1	25.5	28.0	33.1	26.2	28.6	33.7	29.8	32.4	37.8
200	1	1	26.0	28.5	33.8	25.3	27.6	33.1	28.7	31.0	36.1

Notes: The ξ_4 test statistic considers the hypothesis of $r \leq n-4$ at frequency ω against the alternative of $r > n-4$ at that frequency. The λ_4 test statistic considers the alternative hypothesis $r = n-3$ at frequency ω. N is the number of observations, p denotes whether the auxiliary regression (7) includes an additional lag of $\Delta_4 X_t$ (1) or not (0), and μ is the constant term in the DGP.

Table 1e

Simulated significance points (based on 10,000 Monte Carlo replications) for trace test statistics ξ_5 and for eigenvalue test statistic λ_5 under the restriction of no diverging seasonal trends, i.e., model (2) in text. The data generating process is $\Delta_4 X_t = \mu(1,...,1)' + \varepsilon_t$, where ε_t is 5–variate Gaussian with $\Sigma = \mathbf{I}$.

			$\omega = 0$			$\omega = \pi$			$\omega = \pi/2$		
N	p	μ	.90	.95	.99	.90	.95	.99	.90	.95	.99
trace statistic ξ_5											
100	0	0	72.0	76.5	86.1	73.9	77.8	86.8	85.6	90.2	98.6
100	1	0	76.6	81.3	90.3	72.7	77.0	86.2	81.0	85.5	94.7
100	0	1	70.5	74.9	83.2	73.9	78.1	86.2	85.5	90.0	99.1
100	1	1	75.1	79.8	90.3	72.6	76.8	86.0	80.7	85.8	94.1
200	0	0	69.0	73.2	82.3	73.0	77.0	86.0	85.6	90.1	99.3
200	1	0	70.7	75.3	84.4	70.8	75.1	83.1	81.7	86.0	95.0
200	0	1	67.6	71.9	80.4	73.1	76.9	86.1	85.5	90.2	99.3
200	1	1	69.3	73.7	82.5	70.8	74.9	83.1	81.6	85.9	94.8
eigenvalue statistic λ_5											
100	0	0	34.2	37.0	43.6	32.9	36.0	42.1	35.8	38.4	44.9
100	1	0	36.3	39.3	46.0	33.4	36.4	42.9	34.5	37.4	43.8
100	0	1	34.2	37.3	43.5	33.1	36.1	42.3	35.9	38.5	43.9
100	1	1	36.3	39.5	47.0	33.3	36.4	43.2	34.6	37.4	44.1
200	0	0	32.2	34.7	40.8	32.2	34.9	39.9	35.8	38.6	43.8
200	1	0	33.2	35.9	42.0	31.8	34.2	39.7	34.1	36.8	42.2
200	0	1	32.1	34.8	40.7	32.2	35.0	40.1	35.8	38.6	43.9
200	1	1	33.3	36.1	42.0	31.7	34.2	39.4	34.3	36.7	42.2

Notes: The ξ_5 test statistic considers the hypothesis of $r \leq n-5$ at frequency ω against the alternative of $r > n-5$ at that frequency. The λ_5 test statistic considers the hypothesis of $r = n-4$ at frequency ω. N is the number of observations, p denotes whether the auxiliary regression (7) includes an additional lag of $\Delta_4 X_t$ (1) or not (0), and μ is the constant term in the DGP.

Table 1f

Simulated significance points (based on 10,000 Monte Carlo replications) for trace test statistics ξ_6 and for eigenvalue test statistic λ_6 under the restriction of no diverging seasonal trends, i.e., model (2) in text. The data generating process is $\Delta_4 X_t = \mu(1, ..., 1)' + \varepsilon_t$, where ε_t is 6–variate Gaussian with $\Sigma = \mathbf{I}$.

		$\omega = 0$			$\omega = \pi$			$\omega = \pi/2$		
p	μ	.90	.95	.99	.90	.95	.99	.90	.95	.99
trace statistic ξ_6, $N = 100$										
0	0	101.4	106.9	118.1	102.1	107.0	116.9	114.0	119.2	129.0
1	0	109.3	115.1	127.8	103.3	109.0	120.4	108.7	113.9	125.3
0	1	100.1	105.6	116.4	102.0	106.8	116.7	113.7	119.1	129.1
1	1	107.9	113.8	126.6	103.1	109.0	120.3	108.4	114.0	125.2
trace statistic ξ_6, $N = 200$										
0	0	95.1	99.8	108.8	100.0	104.9	114.0	113.5	119.0	129.0
1	0	98.3	102.8	113.1	97.4	102.1	111.3	108.0	113.0	122.2
0	1	93.9	98.6	108.7	99.9	105.0	113.9	113.5	118.7	128.8
1	1	96.8	102.0	112.1	97.4	102.2	111.1	107.7	113.0	122.4
eigenvalue statistic λ_6, $N = 100$										
0	0	42.0	45.4	52.5	40.5	43.7	50.5	42.1	45.5	51.9
1	0	45.5	49.1	56.5	42.3	46.0	53.4	41.8	45.0	51.9
0	1	41.7	45.0	51.8	40.4	43.3	50.5	42.1	45.2	51.8
1	1	45.1	48.8	56.8	42.4	45.9	53.3	41.8	44.8	51.8
eigenvalue statistic λ_6, $N = 200$										
0	0	38.6	41.6	47.5	38.4	41.4	47.3	41.8	44.8	50.2
1	0	40.0	42.9	48.8	38.0	40.8	47.0	40.0	42.8	48.4
0	1	38.6	41.6	47.7	38.5	41.4	47.4	41.6	44.7	50.2
1	1	40.0	43.1	49.5	37.9	40.8	46.9	40.0	42.8	48.7

Notes: The ξ_6 test statistic considers the hypothesis of $r \leq n - 6$ at frequency ω against the alternative of $r > n - 6$ at that frequency. The λ_6 test statistic considers the alternative hypothesis $r = n - 5$ at frequency ω. N is the number of observations, p denotes whether the auxiliary regression (7) includes an additional lag of $\Delta_4 X_t$ (1) or not (0), and μ is the constant term in the DGP.

Table 2

Rejection of the US and RS tests in the case of cointegration at $\omega = \pi$. For $\lambda = 0$ or at $\omega = 0, \pi/2$ the rejection frequencies can be interpreted as size, for $\lambda > 0$ and $\omega = \pi$ the frequencies can be interpreted as power of the test. Sample size is 100.

$\lambda =$	0		0.1		0.2		0.3		0.4	
size	US	RS	US	RS	US	RS	US	RS	US	RS
$\omega = 0$										
1%	0.01	0.01	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010
5%	0.05	0.05	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050
10%	0.1	0.1	0.010	0.101	0.101	0.101	0.101	0.101	0.101	0.101
$\omega = \pi$										
1%	0.01	0.01	0.233	0.232	0.502	0.479	0.767	0.734	0.935	0.916
5%	0.05	0.05	0.378	0.365	0.711	0.679	0.921	0.904	0.991	0.986
10%	0.1	0.1	0.492	0.466	0.818	0.783	0.968	0.956	0.997	0.995
$\omega = \pi/2$										
1%	0.01	0.01	0.010	0.009	0.010	0.009	0.009	0.009	0.009	0.008
5%	0.05	0.05	0.048	0.048	0.047	0.045	0.045	0.041	0.042	0.039
10%	0.1	0.1	0.097	0.097	0.096	0.092	0.092	0.089	0.089	0.082

Table 3
Rejection of the US and RS tests in the case of cointegration at $\omega = \pi$ with superimposed seasonal cycle. At $\omega = 0, \pi/2$ the rejection frequencies can be interpreted as size. For $\lambda = 0$ the model is not data-admissible. For $\lambda > 0$ and $\omega = \pi$ the frequencies can be interpreted as power of the test. Sample size is 200.

		λ									
		0		0.001		0.025		0.05		0.1	
ω	size	US	RS	US	RS	US	RS	US	RS	US	RS
0	1%	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010
	5%	0.050	0.050	0.050	0.050	0.051	0.051	0.051	0.051	0.050	0.050
	10%	0.099	0.099	0.099	0.099	0.100	0.100	0.098	0.098	0.097	0.097
π	1%	0.004	0.824	0.606	0.685	0.499	0.583	0.540	0.556	0.689	0.671
	5%	0.027	0.858	0.542	0.750	0.623	0.681	0.685	0.686	0.852	0.826
	10%	0.053	0.878	0.437	0.791	0.693	0.737	0.764	0.756	0.911	0.893
$\pi/2$	1%	0.010	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.010	0.009
	5%	0.050	0.050	0.051	0.049	0.049	0.049	0.049	0.049	0.049	0.047
	10%	0.095	0.097	0.090	0.097	0.095	0.097	0.095	0.097	0.094	0.095

Table 4

Rejection of the US and RS tests in the case of cointegration at $\omega = \pi/2$ with superimposed seasonal cycle. For $\omega = 0$ and π, the rejection frequencies can be interpreted as size, for $\omega = \pi/2$ the frequencies can be interpreted as power of the test. For $\lambda = 0$ the generated process is not data-admissible. Number of simulated observations is 100.

	$\lambda =$	0.0		0.025		0.05		0.1	
ω	size	US	RS	US	RS	US	RS	US	RS
0	1%	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012
	5%	0.048	0.048	0.048	0.048	0.049	0.049	0.049	0.049
	10%	0.101	0.101	0.101	0.101	0.101	0.101	0.101	0.101
π	1%	0.012	0.008	0.010	0.008	0.011	0.008	0.011	0.008
	5%	0.052	0.053	0.053	0.053	0.054	0.053	0.053	0.051
	10%	0.105	0.103	0.104	0.103	0.104	0.104	0.106	0.102
$\pi/2$	1%	0.003	0.605	0.109	0.450	0.206	0.374	0.306	0.358
	5%	0.019	0.677	0.200	0.555	0.333	0.497	0.471	0.509
	10%	0.040	0.712	0.274	0.615	0.427	0.574	0.583	0.604

Table 5
Trivariate design (12) with one cointegrating vector at all frequencies. Rejection of the US and RS tests in the case of a second cointegration vector at $\omega = \pi/2$. For $\lambda = 0$ and for $\omega = 0$ and π, the rejection frequencies can be interpreted as size, for $\omega = \pi/2$ the frequencies can be interpreted as power of the test. Number of simulated observations is 100.

	$\lambda =$	0.0		0.01		0.02		0.05		0.1	
ω	size	US	RS	US	RS	US	RS	US	RS	US	RS
0	1%	0.012	0.012	0.012	0.012	0.013	0.013	0.012	0.012	0.012	0.012
	5%	0.055	0.055	0.053	0.053	0.053	0.053	0.054	0.054	0.054	0.054
	10%	0.116	0.116	0.113	0.113	0.112	0.112	0.111	0.111	0.110	0.110
π	1%	0.011	0.010	0.014	0.128	0.292	0.564	0.941	0.962	0.998	0.999
	5%	0.054	0.053	0.059	0.232	0.406	0.652	0.956	0.973	1.000	1.000
	10%	0.104	0.104	0.113	0.318	0.478	0.704	0.964	0.980	1.000	1.000
$\pi/2$	1%	0.009	0.009	0.009	0.010	0.010	0.010	0.009	0.010	0.009	0.009
	5%	0.051	0.052	0.050	0.053	0.050	0.052	0.048	0.053	0.045	0.051
	10%	0.103	0.104	0.102	0.103	0.099	0.102	0.098	0.100	0.093	0.098

Table 6
Results of seasonal cointegration tests for four-dimensional Austrian macroeconomic data set. Trace test statistics.

	no dummies			unrestricted dummies			restricted dummies		
$\omega =$	0	π	$\pi/2$	0	π	$\pi/2$	0	π	$\pi/2$
$n = 1$	1.68	0.22	1.27	1.44	6.08	6.74	1.74	6.34	8.51
1%	6.6	7.3	7.2	6.6	12.4	16.2	7.1	12.1	16.4
5%	3.8	4.3	4.4	3.8	8.6	11.9	3.8	8.8	12.6
10%	2.7	3.1	3.1	2.7	6.9	9.9	2.8	7.3	10.8
$n = 2$	8.71	6.90	4.78	8.69	24.28	25.37	8.69	18.08	27.91
1%	19.7	17.3	16.7	19.7	24.5	30.2	20.1	24.8	31.7
5%	15.3	13.0	12.3	15.3	19.3	24.3	15.5	20.1	26.4
10%	13.3	11.1	10.3	13.3	17.0	21.6	13.6	17.8	23.9
$n = 3$	21.63	26.36	23.65	20.55	50.38	58.09	20.55	37.18	48.16
1%	34.9	31.7	29.8	34.9	41.0	47.7	36.5	40.7	49.8
5%	29.4	26.2	24.0	29.4	34.4	40.6	30.4	34.9	43.2
10%	26.7	23.4	21.2	26.7	31.3	37.1	27.7	32.0	40.2
$n = 4$	47.08	57.86	53.90	45.18	85.42	99.04	45.18	72.20	85.07
1%	53.9	na	na	53.9	na	na	57.2	61.5	72.0
5%	47.2	na	na	47.2	na	na	50.2	54.4	64.9
10%	43.8	na	na	43.8	na	na	46.5	51.0	61.0

Notes: The auxiliary regression in (1) includes 1 lag of $\Delta_4 X_t$. The solid numbers are test statistics calculated from the data, the numbers in italics denote significance points. For columns 1 and 4, these were taken from Table 15.3 in Johansen (1995); for columns 2, 3, 5, 6, we consulted Lee and Siklos (1995); columns 7 to 9 correspond to our Tables 1a-1d.

Table 7
Canonical vectors in the Austrian data set as identified by the methods of unrestricted and restricted seasonal dummy variables.

$\omega = 0$ (both methods)

variable	Y	C	I	W
vector 1	0.70	0.19	-0.68	-0.09
vector 2	-1.15	0.35	0.14	0.61
vector 3	-12.60	11.65	0.37	0.12
vector 4	-0.40	0.75	0.32	-0.52

$\omega = \pi$

variable	Y	C	I	W	$\cos(\pi t)$
		unrestricted dummies			
vector1	0.11	-0.23	-0.33	0.91	
vector2	1.05	-0.19	0.09	0.97	
vector3	-1.47	0.46	0.32	0.10	
vector4	-0.42	-0.81	0.40	0.05	
		restricted dummies			
vector1	-0.08	0.24	0.35	-0.90	0.00
vector2	-1.26	0.36	0.19	-0.19	0.03
vector3	0.09	-0.26	-0.75	-2.43	0.76
vector4	-1.59	-3.30	1.59	0.17	0.07

$\omega = \pi/2$

variable	Y	C	I	W	$\cos(\frac{\pi t}{2})$	$\cos(\frac{\pi(t-1)}{2})$
		unrestricted dummies				
vector1	-0.93	-0.10	-0.15	0.80		
vector2	-0.34	-0.52	0.66	1.38		
vector3	-0.89	0.56	0.10	0.00		
vector4	-0.71	-0.56	0.17	-0.39		
		restricted dummies				
vector1	0.59	0.17	-0.09	-0.78	-0.02	-0.02
vector2	-0.89	0.58	-0.06	-0.11	0.07	0.01
vector3	1.72	-0.87	-0.87	-0.37	0.32	0.09
vector4	0.29	0.86	-0.07	0.38	-0.08	-0.15

FIGURES

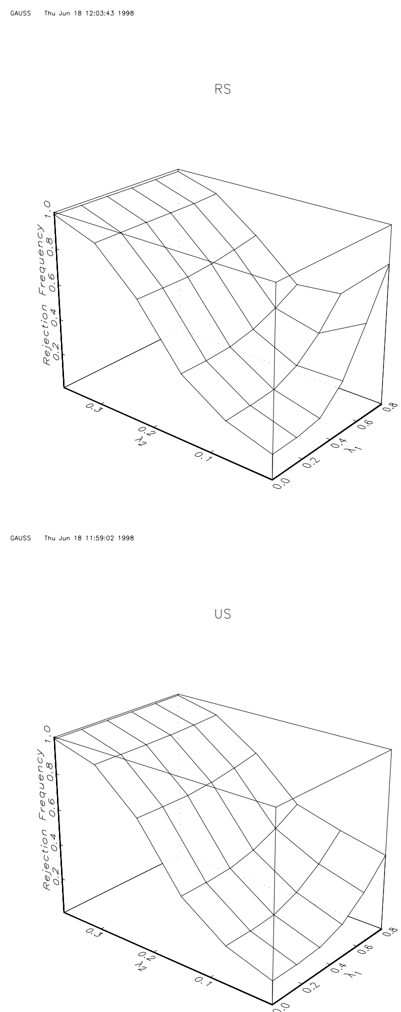

Figure 1: Simulated rejection frequencies of RS and US testing procedures as a function of λ_2 (seasonal cointegration) and λ_1 (deterministic seasonality). Sample size is 100.

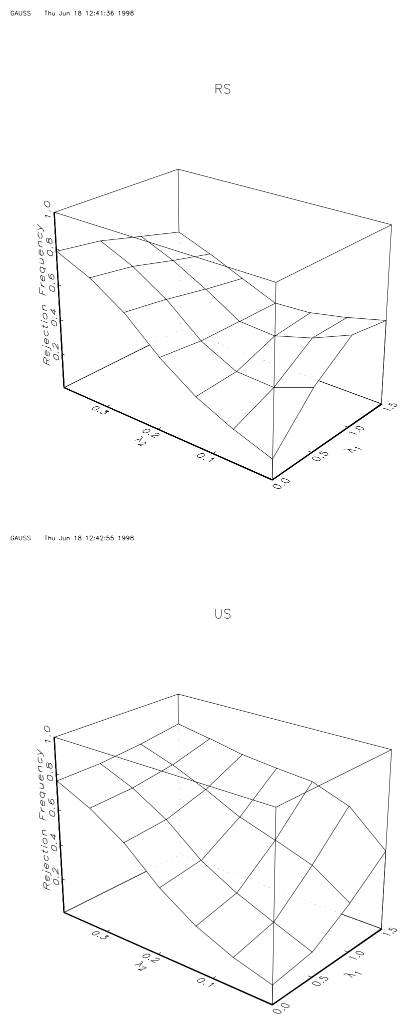

Figure 2: Simulated rejection frequencies of RS and US testing procedures as a function of λ_2 (seasonal cointegration) and λ_1 (deterministic seasonality) if lag length is determined by AIC. Sample size is 100.

Figure 3: Simulated rejection frequencies of RS and US testing procedures as a function of λ_2 (seasonal cointegration) and λ_1 (deterministic seasonality) if lag length is determined by BIC. Sample size is 100.