ELSEVIER

Journal of Econometrics 63 (1994) 133–151

JOURNAL OF Econometrics

A multivariate approach to modeling univariate seasonal time series

Philip Hans Franses

Econometric Institute, Erasmus University, 3000 DR Rotterdam, The Netherlands

(Received October 1993)

Abstract

A seasonal time series can be represented by a vector autoregressive model for the annual series containing the seasonal observations. This model allows for periodically varying coefficients. When the vector elements are integrated, the maximum likelihood cointegration method can be used to check for the presence of, possibly restricted, cointegration relations between these annual series. In this paper it is shown that this application generalizes a test procedure for seasonal unit roots. Simulations and examples illustrate its empirical performance.

Key words: Seasonality; Vector autoregression; Cointegration
JEL classification: C22

1. Introduction

Common assumptions for models of a seasonally observed economic time series are, e.g., (a) the series is seasonally integrated, (b) seasonal patterns can be represented by deterministic dummies, and (c) a variable is periodically integrated (see, e.g., Osborn, 1988). The Hylleberg et al. (1990) [HEGY] method is designed to discriminate between models implied by assumptions (a) and (b). Thus, the HEGY approach considers only a subset of possible models, and, in

Thanks are due to Robert Engle, Eric Ghysels, David Hendry, Svend Hylleberg, Teun Kloek, Marius Ooms, Denise Osborn, two anonymous referees, and to participants of seminars in Tilburg, Aarhus, the $(EC)^2$ 1991 meeting in Rotterdam, and of the Australasian Meeting of the Econometric Society 1992 in Melbourne for their comments on previous drafts. A special thanks to Peter Boswijk for generating the critical values, and to Richard Paap and Gerbert Romijn for research assistance.

SSDI 0304407693015632

particular, does not allow for periodically varying coefficients. This calls for a model selection strategy in a class of models that considers models of types (a), (b), and (c). In the present paper it is argued that a vector autoregressive [VAR] model with cointegrating restrictions for the annual series of seasonal observations can provide such an extended class. The reason for choosing a VAR specification is that several models for the univariate series imply the presence of, and restrictions on, cointegration relations between these annual series. Each model assumes a different number of cointegrating relations, and therefore an estimate of that number is of importance. Further, since the validity of certain parameter restrictions in the cointegrating vectors are to be tested, the Johansen (1988) test procedure seems to be most suitable for model selection.

Section 2 discusses model representation issues. Particularly, the focus is on writing autoregressive models for a univariate series as a VAR for the vector of stacked observations. The relations between cointegrated annual series and, e.g., (non) seasonal unit roots and periodic integration are also highlighted. For notational convenience and illustrative purposes, I deal with quarterly time series only. Section 3 proposes a model selection strategy which is based on applying the Johansen method to this VAR model. In Section 4 this strategy is compared with the HEGY method via some Monte Carlo simulations. In Section 5 the proposed model selection method is applied to the Japanese consumption and income series studied in Engle et al. (1993). More examples can be found in Franses (1990). The final section concludes with some remarks.

2. Model representation

Consider a univariate quarterly time series x_t, $t = 1, \ldots, n$, when it is generated by an autoregressive process of order p [AR(p)],

$$\phi_p(B)x_t = \delta + \varepsilon_t, \qquad (1)$$

where δ is a constant and $\phi_p(B)$ is a polynomial of order p in the backward shift operator B. This operator is defined by $B^k x_t = x_{t-k}$. The ε_t denotes a standard white noise process, i.e., an uncorrelated zero mean process with constant variance. In case the polynomial $\phi_p(B)$ can be decomposed as $\phi^*_{p-4}(B)(1 - B^4)$, the x_t series is said to be seasonally integrated. Since $(1 - B^4)$ equals $(1 - B)(1 + B)(1 - iB)(1 + iB)$, this means that x_t then contains a nonseasonal unit root 1 and the seasonal unit roots -1, $-i$, and $+i$. Hylleberg et al. (1990) propose a procedure to test for the presence of these roots.

Although the constant δ can be replaced by seasonally varying constants, one of the properties of model (1) is that its dynamic parameters do not vary with the seasons. This variation can be introduced by adding a subscript s to the elements of (1) (see also Osborn, 1991). A simple example is the first-order autoregressive model

$$x_t = \alpha_s x_{t-1} + \varepsilon_{st}, \tag{2}$$

where the values of the autoregressive parameter and the variance of the error process vary with the season, and where $s = 1, 2, 3, 4$. This process is called a periodic process (see Pagano, 1978; Troutman, 1979; *inter alia*). Strictly speaking, periodic processes like (2) are not stationary since they treat each season differently, and variances and correlations therefore vary with the seasons. The latter suggests an alternative representation of a periodic autoregressive process, which is given by a multivariate process for the (4×1) vector X_T containing the annual series X_{sT}, or $X_T = (X_{1T}, X_{2T}, X_{3T}, X_{4T})'$, where X_{sT} is the observation in season s in year T. The annual index T runs from 1 to N, where $N = n/4$. This multivariate process is

$$A_0 X_T = A_1 X_{T-1} + \cdots + A_m X_{T-m} + \mu + \varepsilon_T, \tag{3}$$

where the A_i, $i = 0, 1, \ldots, m$, are (4×4) parameter matrices, and where the μ is a (4×1) vector of constants and ε_T is a (4×1) vector white noise process. The model in (3) can be called a vector of quarters [VQ] representation. Note that this model only allows the parameters to be periodic, and that they do not necessarily have to be seasonally varying, i.e., models like (1) can also be written as (3). The idea of stacking has been introduced in Gladyshev (1961) and is also considered in, e.g., Tiao and Grupe (1980) and Osborn (1991). As an example, the univariate model in (2) can be represented as

$$\begin{bmatrix} 1 & 0 & 0 & -\alpha_1 L \\ -\alpha_2 & 1 & 0 & 0 \\ 0 & -\alpha_3 & 1 & 0 \\ 0 & 0 & -\alpha_4 & 1 \end{bmatrix} \begin{bmatrix} X_{1T} \\ X_{2T} \\ X_{3T} \\ X_{4T} \end{bmatrix} = \begin{bmatrix} \varepsilon_{1T} \\ \varepsilon_{2T} \\ \varepsilon_{3T} \\ \varepsilon_{4T} \end{bmatrix}, \tag{4}$$

where the backward shift operator L is similarly defined as B, i.e., $L^k X_T = X_{T-k}$, and it refers to annual time series.

The vector series X_T generated by (3) is stationary when the roots of the characteristic equation,

$$|A_0 - A_1 z - \cdots - A_m z^m| = 0, \tag{5}$$

are outside the unit circle. For the example in (2) and (4) this means that the solution to

$$|A_0 - A_1 z| = (1 - (\alpha_1 \alpha_2 \alpha_3 \alpha_4) z) = 0 \tag{6}$$

should exceed one, i.e., that $\alpha_1 \alpha_2 \alpha_3 \alpha_4 < 1$. When $\alpha_1 \alpha_2 \alpha_3 \alpha_4 = 1$, the system for X_T has one unit root, and the corresponding univariate process x_t is said to be periodically integrated (see Osborn, 1988). Note that at least one of these α_s exceeds one.

Alternative to (5), the presence of unit roots in the system for X_T can be checked by analyzing a rewritten version of the VQ model. For that purpose the VQ model is written in a vector autoregressive form, i.e.,

$$X_T = \Pi_1 X_{T-1} + \cdots + \Pi_m X_{T-m} + v + \omega_T, \tag{7}$$

where $\Pi_i = A_0^{-1} A_i$, $i = 1, \dots, m$, $v = A_0^{-1}\mu$, and $\omega_T = A_0^{-1}\varepsilon_T$. This model can again be written as

$$\Delta X_T = \Gamma_1 \Delta X_{T-1} + \cdots + \Gamma_{m-1} \Delta X_{T-m+1} + \Pi X_{T-m} + v + \omega_T; \tag{8}$$

see Johansen (1988). The $\Delta = 1 - L$ is the first-order differencing filter for annual data and it corresponds to the $\Delta_4 = (1 - B^4)$ filter for quarterly series. The matrices Γ_j, $j = 1, \dots, m-1$, are functions of the Π_i in (7). The Π matrix in (8) conveys information on stationarity and on the cointegration relations between the elements of X_T. When three of the solutions of (5) are outside the unit circle and one of the solutions is $z = 1$, the process (5) has one unit root, and hence the matrix Π has rank $4 - 1 = 3$. This means that there are three cointegration relations between the X_{sT}, $s = 1, \dots, 4$. For example, the model in (4) can be rewritten as $\Delta X_T = \Pi X_{T-1} + \omega_T$, or

$$\begin{bmatrix} \Delta X_{1T} \\ \Delta X_{2T} \\ \Delta X_{3T} \\ \Delta X_{4T} \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & \alpha_1 \\ 0 & -1 & 0 & \alpha_1\alpha_2 \\ 0 & 0 & -1 & \alpha_1\alpha_2\alpha_3 \\ 0 & 0 & 0 & \alpha_1\alpha_2\alpha_3\alpha_4 - 1 \end{bmatrix} \begin{bmatrix} X_{1T-1} \\ X_{2T-1} \\ X_{3T-1} \\ X_{4T-1} \end{bmatrix} + \begin{bmatrix} \omega_{1T} \\ \omega_{2T} \\ \omega_{3T} \\ \omega_{4T} \end{bmatrix}. \tag{9}$$

It is now easily seen that the Π matrix in (9) has rank 3 when $\alpha_1\alpha_2\alpha_3\alpha_4 = 1$. This means that a periodically integrated first-order autoregressive time series assumes three cointegration relations between the X_{sT} variables. From (9) it can also be observed that these relations are simply $X_{1T} - \phi_1 X_{4T}$, $X_{2T} - \phi_2 X_{4T}$, and $X_{3T} - \phi_3 X_{4T}$, where not all ϕ_j, $j = 1,2,3$ are equal to unity. Note that when the α_s are all equal to 1, i.e., when the model $\Delta_1 x_t = \varepsilon_t$ is appropriate, there are also three cointegration relations for which now all the ϕ_j are equal to 1. Of course, when the rank of Π is equal to 0, it can be seen from (8) and (9) that the Δ_4 filter for the x_t series may be adequate.

Altogether, this suggests that a method to select between several models for seasonal time series can be based on an estimate of the rank of Π and on a test for the validity of restrictions on the parameters in the cointegration vectors. The Johansen (1988) maximum likelihood cointegration method may now be useful for this purpose (see also Johansen and Juselius, 1990).

3. Model selection

The model selection method for a univariate seasonal time series x_t is based on the VQ representation (3), its rewritten version (8), and an analysis of the properties of the matrix Π. Consider the first-order VQ model

$$\Delta X_T = \Pi X_{T-1} + v + \omega_T, \tag{10}$$

where the matrix Π will be estimated unrestrictedly, and assume that for the error process applies that $\omega_T \sim \mathrm{N}_4(0, \Lambda)$. Since this model is essentially a vector autoregression of order 1, the sixteen parameters in Π can be estimated using ordinary least squares. Denote r as the rank of Π.

The process X_T is asymptotically stationary in case r equals 4. There is no cointegration between the elements of X_T in case r is equal to 0. In case $0 < r < 4$, one can write $\Pi = \alpha\beta'$, where α and β are $(4 \times r)$ matrices, of which the matrix β contains the cointegration vectors. Johansen (1988) developed test procedures for the value of r and for linear hypotheses in terms of α and β. For (10), the method boils down to the choice of the r linear combinations of elements of X_T which have the largest correlation with ΔX_T after correcting for v. The eigenvectors of the relevant canonical correlation matrix are the columns of β. The corresponding eigenvalues λ_i, where $\lambda_i \geqslant \lambda_{i+1}$, are used to construct statistics like $Q_1(r) = -N\sum_{i=r+1}^{4}\log(1-\lambda_i)$ and $Q_2(r) = -N\log(1-\lambda_r)$. The trace test statistic Q_1 and the maximum eigenvalue test statistic Q_2 can be used to test for the number of cointegration vectors.

Asymptotic fractiles for these statistics are displayed in Johansen and Juselius (1990). However, preliminary Monte Carlo simulations reported in earlier versions of my paper have indicated that for sample sizes as large as $N = 25$, these critical values may not be appropriate. Therefore, small-sample fractiles for the statistics Q_1 and Q_2 have been calculated on the basis of 10000 replications for samples of 25 and 50 observations. Note that these sample sizes for annual data correspond to 100 and 200 quarterly data, respectively. In the Appendix the tables with fractiles are displayed. A comparison of these with the corresponding tables in Johansen and Juselius (1990) indicates that the critical values in small samples differ from the asymptotic ones although a convergence can be observed as sample size grows, and that the differences across the distinct null hypotheses between the values for sample sizes 25 and 50 are not very large.

To test for linear restrictions on the cointegrating vectors β, define a $(4 \times q)$ matrix H, where $r \leqslant q \leqslant 4$, which reduces β to the parameters φ, or $\beta = H\varphi$. For brevity, I shall denote these restrictions by their matrix H. Assuming the validity of the restrictions H, one compares the corresponding eigenvalues ξ_i of the canonical correlation matrix with the λ_i via the test statistic $Q = N\sum_{i=1}^{r}\log\{(1-\xi_i)/(1-\lambda_i)\}$. Under the null hypothesis, the test statistic

Q asymptotically follows a $\chi^2(r(4-q))$ distribution. Whether this distribution is valid in small samples will be investigated below.

An application of the Johansen cointegration method to the model in (10) gives an opportunity to gain insights in the properties of the univariate quarterly x_t series. No differencing filter is needed for x_t in case r is equal to 4. In case there are no cointegration relationships between the elements of X_T, each X_{sT} series is an integrated process (see also Osborn, 1993). Hence, $r = 0$ implies that the filter Δ_4 for x_t may be appropriate. If r is 3 and pairs of successive X_{sT} are cointegrated with parameters $(1, -1)$, a transformation Δ_1 for x_t is required. This Δ_1 filter assumes the cointegration relations $(X_{2T} - X_{1T})$, $(X_{3T} - X_{2T})$, and $(X_{4T} - X_{3T})$. In terms of model (10) this means that the restrictions on the columns of β, given by

$$H_{31} = \begin{pmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}, \tag{11}$$

are not rejected. In terms of (1), this means that x_t has a nonseasonal unit root 1. The $\Delta_1 x_t$ series may now be described by an autoregressive model with seasonally varying parameters. Whether all the parameters are periodic indeed can then be tested along standard lines.

It is also possible to test for the presence of seasonal unit roots (see Hylleberg et al., 1990). When $r = 3$, one can check for the presence of root -1 by testing the restrictions

$$H_{32} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}. \tag{12}$$

If both the hypotheses H_{31} and H_{32} are rejected, one has encountered three general cointegration relationships between the elements of X_{sT}. Specific restrictions on the parameters in Π can now imply the appropriateness of the periodically integrated model like $x_t = \alpha_s x_{t-1} + \varepsilon_{st}$ with $\alpha_1 \alpha_2 \alpha_3 \alpha_4 = 1$ but not all $\alpha_s = 1$ as in (2). When r is equal to 1 or 2, one can proceed along similar lines to test for the presence of nonseasonal and/or specific seasonal unit roots. In Table 1 the relevant restriction matrices H are given.

Summarizing, an application of the Johansen cointegration method to a VAR model for the X_T vector generalizes the HEGY approach since it allows for the presence of periodically varying parameters. When each of the hypothes es H in

Table 1 is rejected, the time series x_t can be said to be periodically integrated. A possibly suitable model for such a series is a univariate periodic error correction model like (10) where Π is replaced by $\alpha\beta'$. The 'error' of over-differencing is then corrected by $\beta' X_{T-1}$, which represents linear relationships between the annual series. The results in Engle and Yoo (1987) suggest that a forecasting gain can be expected when this model is used for forecasting. In Franses and Romijn (1993) it is illustrated that such forecasts compare favourably with those obtained from a model for a Δ_4 transformed x_t series.

Alternatively, one may also want to construct a model containing more than one time series. In case of periodically integrated time series, it may then be worthwhile to consider a periodic cointegration model, which is a model where the cointegration vectors, as well as the adjustment parameters, are allowed to vary over the seasons (see, e.g., Birchenhall et al., 1989; Franses and Kloek, 1991).

4. A Monte Carlo study

The model in (10), which is the simplest VQ model, contains 16 + 14 parameters to be estimated. When some of these parameters are in fact equal to 0, as for example in (9), this may have an impact on the empirical performance of the cointegration method. Further, nonperiodic models like (1) imply parameter restrictions on the elements of Π, and this can also effect size and power of the test strategy. This section reports on the results of a Monte Carlo study in which small periodic and nonperiodic, possibly integrated, time series are the data-generating processes. In the simulations below, only the trace test statistic Q_1 will be used in the VQ method. Further, the VQ model selection approach will be compared with the HEGY method. To save space, it is assumed that the reader is familiar with the details of the latter approach.

Issues of interest are whether the empirical power of the procedure is reasonably high, also in cases where the model in (10) is overparameterized, and whether the asymptotic χ^2 distribution for the tests for restrictions is valid in samples as small as 25 observations. In the simulations the maintained regression model is (10), i.e., the model includes a constant term. Below, this model will be called a VQ model of order 1. The relevant critical values are obtained from Table A.1 in the Appendix. For comparability reasons, I consider an auxiliary regression for the HEGY method when it includes seasonal dummies and a deterministic trend. The significance level for each step of the HEGY method is set equal to 5%. To gain insight in the performance of the VQ method itself, I report on the results obtained at a 5% and at a 10% level.

In case the data-generating process is a periodically integrated process in (2) which is a process not captured by the HEGY method, i.e., (2) with the imposed restriction that $\alpha_1\alpha_2\alpha_3\alpha_4 = 1$, the differences between the two methods are most

Table 1

Testing for (non) seasonal unit roots via an application of the maximum likelihood cointegration method to the model $\Delta X_T = \Pi X_{T-1} + v + \omega_T$, where X_T is the (4 × 1)annual vector containing the quarterly observations

Rank of Π	Restrictions matrix	Cointegration vectors	Differencing filter	(Non) seasonal unit roots
3	$H_{31} = \begin{pmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$	$X_{2T} - X_{1T}$ $X_{3T} - X_{2T}$ $X_{4T} - X_{3T}$	$(1 - B)$	1
3	$H_{32} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$	$X_{2T} + X_{1T}$ $X_{3T} + X_{2T}$ $X_{4T} + X_{3T}$	$(1 + B)$	-1
2	$H_{21} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$	$X_{3T} - X_{1T}$ $X_{4T} - X_{2T}$	$(1 - B^2)$	$1, -1$
2	$H_{22} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$	$X_{3T} + X_{1T}$ $X_{4T} + X_{2T}$	$(1 + B^2)$	$i, -i$

1	$H_{11} = \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix}$	$X_{4T} - X_{3T} + X_{2T} - X_{1T}$	$(1 - B)(1 + B^2)$	$1, i, -i$
1	$H_{12} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$	$X_{4T} + X_{3T} + X_{2T} + X_{1T}$	$(1 + B)(1 + B^2)$	$-1, i, -i$
0			$(1 - B^4)$	$1, -1, i, -i$

Table 2
Monte Carlo comparison based on 5,000 replications of the VQ and HEGY methods when the DGP is a periodically integrated process: $x_t = \alpha_s x_{t-1} + \varepsilon_{st}$ with

I: $\alpha_1 = 1.25$, $\alpha_2 = 0.8$, $\alpha_3 = 0.9$, $\alpha_4 = 1.11$, $\varepsilon_{st} \sim \mathrm{N}(0, \sigma_s)$
II: $\alpha_1 = 2$, $\alpha_2 = 0.5$, $\alpha_3 = 1.5$, $\alpha_4 = 0.67$, $\varepsilon_{st} \sim \mathrm{N}(0, \sigma_s)$

100 quarterly observations

DGP		VQ[a]		HEGY[b]					
		10%	5%	1	−1	$\pm i$	−	Δ_1	Δ_4
I:	$\sigma_s = 1$	0.588	0.438	0.961	0.196	0.196	0.026	0.753	0.159
	$\sigma_s = 1.25, 0.8, 0.5, 2.0$	0.595	0.441	0.960	0.225	0.215	0.023	0.727	0.179
II:	$\sigma_s = 1$	0.619	0.488	0.948	0.859	0.272	0.020	0.117	0.247
	$\sigma_s = 1.25, 0.8, 0.5, 2.0$	0.630	0.491	0.949	0.904	0.302	0.023	0.072	0.280

[a] The values in the cells report the number of times the correct decision is made, i.e., r equals 3 and the filter Δ_1 is not appropriate.
[b] The values in the cells report the number of times the presence of the roots 1, −1, $\pm i$ cannot be rejected. The outcomes are based on the t tests for π_1 and π_2 and the joint F test for π_3 and π_4 in the auxiliary regression (3.8) in HEGY, which here contains a constant, seasonal dummies, and a trend. No filter (−) is chosen when all $\pi_i \neq 0$, Δ_1 when $\pi_1 = 0$ and the other π_i are not, and Δ_4 when all π_i equal 0. The test outcomes for the π_i are based on a 5% significance level. The number of additional lags in the test equation is set equal to that number p ($p = 12, \ldots, 0$) for which there appeared to be no significant residual autocorrelation.

striking. Some simulation results relevant to this case, where only two sets of parameter values have been chosen for which applies that the product equals 1, are reported in Table 2. In about 50% of the cases the VQ method selects the correct model, i.e., r is equal to 3 and the restrictions H_{31} are rejected. One reason for this somewhat low value of the power is that the VQ model is over-parameterized. This seems to be confirmed by the unreported fact that, next to the 50% of the cases that r is found to be equal to 3, generally in about 35% of the cases this r is estimated to be 2. Another cause may be that the asymptotic χ^2 distribution may not apply to samples as small as 25 annual observations. In fact, at a nominal level of 5%, the rejection rate of the Q test statistic is about 20%. It can be expected that the results for higher-order periodic autoregressive models estimated for longer time series will show an improvement of the test performance. The HEGY method indicates that in several cases one is inclined to opt for the Δ_1 filter, although also the Δ_4 filter can often be found to be appropriate. Further, it can be seen that a likely outcome of the HEGY method is that the root −1 or the roots $\pm i$ seem to be present. Of course, given the choice of the parameters, this may not come as a surprise. The performances of the two methods do not seem to be effected by a seasonally heteroscedastic error process.

Table 3
Monte Carlo evaluation of HEGY and VQ procedure; 100 quarterly observations

		VQ	
Data-generating process	HEGY	10%	5%
(I) $x_t = \alpha_s x_{t-1} + \varepsilon_{st}$, $\varepsilon_{st} \sim \mathrm{N}(0, \sigma_s)$			
$\alpha_s = 0$, $\sigma_s = 1$	0.926	0.916	0.774
$\alpha_s = 0.5$, $\sigma_s = 1$	0.802	0.865	0.697
$\alpha_s = 0.9$, $\sigma_s = 1$	0.121	0.323	0.174
$\alpha_s = 0.5$, $\sigma_s = 1.1, 0.9, 1.5, 0.7$	0.789	0.875	0.699
$\alpha_s = 0.5$, $\sigma_s = 1.25, 0.8, 0.5, 2.0$	0.758	0.853	0.683
$\alpha_s = 0.2, 0.4, 0.6, 0.8$, $\sigma_s = 1.25, 0.8, 0.5, 2.0$	0.812	0.893	0.731
$\alpha_s = 0.6, 0.7, 0.8, 0.9$, $\sigma_s = 1.25, 0.8, 0.5, 2.0$	0.440	0.694	0.486
(II) $\Delta_1 x_t = \alpha_s \Delta_1 x_{t-1} + \varepsilon_{st}$, $\varepsilon_{st} \sim \mathrm{N}(0, \sigma_s)$			
$\alpha_s = 0$, $\sigma_s = 1$	0.890	0.422	0.365
$\alpha_s = 0.5$, $\sigma_s = 1$	0.900	0.482	0.461
$\alpha_s = 0.5$, $\sigma_s = 1.25, 0.8, 0.5, 2.0$	0.890	0.530	0.512
$\alpha_s = 0.2, 0.4, 0.6, 0.8$, $\sigma_s = 1.25, 0.8, 0.5, 2.0$	0.889	0.460	0.427
(III) $\Delta_4 x_t = 0.5 \Delta_4 x_{t-1} + \varepsilon_t$, $\varepsilon_t \sim \mathrm{N}(0, 1)$	0.584	0.774	0.875

The values in the cells report the frequencies that the method selects the correct filter. For case I, this should be no filter or, in terms of VQ, r equals 4. For case II this filter is Δ_1, and for case III it is Δ_4. The evaluation is based on 5,000 replications of series of length 100. All HEGY test outcomes are based on a 5% significance level, while all results for a VQ model of order 1 consider 5% as well as 10% significance levels. The critical values for the VQ method are those displayed in the Appendix.

To further investigate how the VQ method performs in case the generating processes are simple (non) periodic processes, which imply overparameterized VQ models, consider the results in Table 3. In the cases in which the process x_t does not need to be differenced, i.e., the cases in the upper part of the Table 2, the VQ method detects that the Π matrix is of full rank in a large amount of the cases, and it sometimes performs better than the HEGY method does. As expected, the power of the method decreases when the root of the process approaches unity. For example, for an α_s parameter of 0.9 for all s, in only 12.1% of the cases the HEGY method finds that the roots 1, -1, i, and $-i$ are not present jointly, and in only 14.1% of the cases the VQ approach detects that the correct r is equal to 4. Again, allowing for a periodic error process does not dramatically effect the performances. Finally, when the first-order autoregressive parameter can vary with the seasons, this does not effect the outcomes to a great extent either.

When the Δ_1 filter for the x_t series is appropriate, the VQ method does not perform extremely well, as can be seen from part II of Table 3. The empirical powers in these cases are similar to those in Table 2. In about 45% of the cases the filter is found back. This can imply that the 20% significance level for the trace test may be more useful and that, e.g., a 1% level for the Q test for restrictions in the cointegration parameters may be more appropriate.

When the data-generating process is a first-order autoregressive model for a series that needs a Δ_4 filter to reach stationarity, as in panel III of Table 3, the VQ method detects the correct filter in more cases than the HEGY method. Note that the empirical success rate in this case is higher for the 5% significance level. This counterintuitive result is caused by the fact that the figures in the cells correspond to 1 minus the size of the tests. The HEGY approach finds in about 40% of the cases that the Δ_4 is not appropriate. This suggests that the size of the VQ method is not much effected by lagged $\Delta_4 x_t$ terms.

In summary, the VQ and HEGY approaches can yield similar outcomes in case the data-generating processes are close to those assumed for the HEGY method. Hence, even when the multivariate time series model to which the VQ method is applied is highly overparameterized, the VQ method performs reasonably well. A suggestion for practical use of this approach is to consider also a nominal size of 20% for the trace test statistic, and to test restrictions on the cointegration relations using a 1% significance level. When one allows for periodically nonstationary processes, it appears that the VQ approach is a useful generalization of the HEGY method, since the latter method can only suggest the use of inappropriate filters.

5. Applications

To empirically illustrate the VQ approach, I consider the Japanese consumption c_t and income y_t series for 1961.1 to 1987.4 as they are analyzed in Engle et al. (1993). From their graphs it emerges that the series clearly do not show constant patterns. In Franses (1990) it is argued that graphs of the four X_{sT} series can give useful insights in seasonal patterns. The graphs of the X_{sT} series for income in Fig. 1 show patterns of pairs of quarters which seem to evolve similarly over time, and also the distances between the individual lines look rather constant. Furthermore, there is only a brief period where one of the inequalities $X_{4T} > X_{3T} > X_{2T} > X_{1T}$ is violated. From Fig. 2, where the X_{sT} graphs for consumption are displayed, it can be seen that for the c_t series similar patterns emerge, although now none of these inequalities is violated. Hence there seems to be visual evidence for the presence of cointegration relations between the elements of X_T for both series.

The order of a reasonably adequate vector autoregressive model appears to be equal to 1 for both series, i.e., a model as in (10) can be analyzed. This choice is based on the very small number of parameters which are significant in a VQ model of order 2, and on the insignificance of almost all the residual

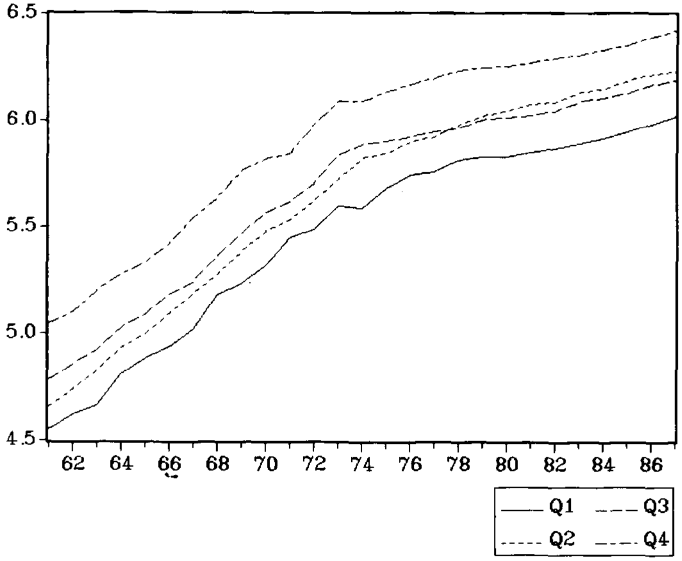

Fig. 1. Real disposable income in Japan per quarter, 1961–1987

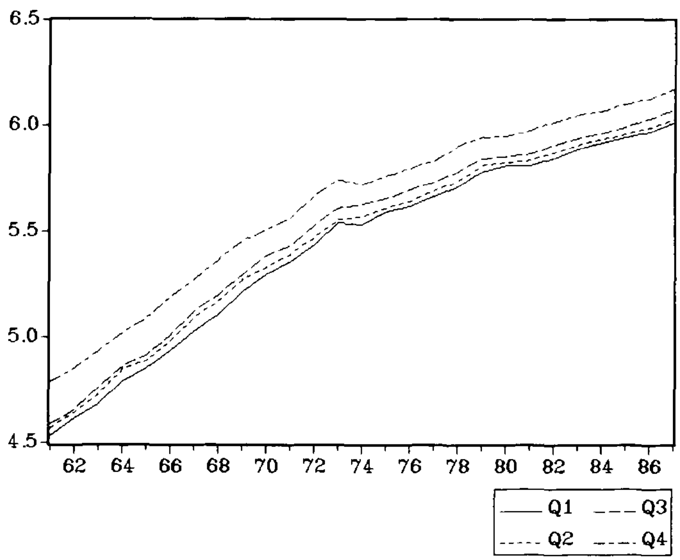

Fig. 2. Total real consumption in Japan per quarter, 1961–1987

Table 4

Estimation and testing results of VQ models of order 1 for the Japanese income and consumption series, $N = 26$

Variable	Π_1				$\Lambda_{ss}^{1/2}$[a]	Correlations		
Income	0.094	0.260	−0.218	0.837	2.648	0.106	0.509	0.568
	0.161	0.610	−0.205	0.427	1.480		0.619	0.522
	0.093	−0.170	0.249	0.806	2.086			0.753
	0.355	−0.246	0.039	0.804	2.955			
Consumption	0.353	−0.877	0.386	1.169	0.755	0.529	0.379	0.386
	−0.079	−0.502	0.270	1.388	1.194		0.642	0.621
	−0.228	−0.683	0.560	1.391	1.397			0.777
	−0.355	−0.303	0.245	1.395	1.778			

Income: autocorrelations of the residuals[b]

Lag 1	0.041	0.016	−0.065	0.021	Lag 2	−0.263	0.082	−0.006	−0.209
	0.184	−0.046	0.373	0.210		0.085	0.107	−0.020	−0.227
	0.250	0.257	0.316	0.194		0.129	0.147	0.228	0.013
	0.187	0.167	0.169	0.076		0.080	−0.031	0.041	−0.131

Consumption: autocorrelations of the residuals[b]

Lag 1	0.000	0.158	0.124	0.227	Lag 2	−0.197	−0.479	−0.202	−0.307
	−0.315	−0.133	−0.068	0.008		0.133	−0.258	−0.197	−0.391
	−0.306	−0.235	−0.286	−0.106		−0.165	−0.344	−0.082	−0.317
	−0.329	−0.194	−0.177	−0.100		−0.087	−0.150	0.004	−0.314

[a] The values of the standard deviations are multiplied by 100.
[b] The standard errors of the autocorrelations are 0.196.

autocorrelations of order 1 and 2; see Table 4. Note that the incorrect assumption of too large a model decreases the empirical power of many cointegration methods (see, e.g., Boswijk and Franses, 1992). The VQ model for consumption contains a dummy for 1974 in all four equations to capture a sequence of outliers. The relevant estimation results of the unrestricted models are displayed in Table 4. These are the estimates of Π_1, see (7), and of the standard deviations of the residual processes. Furthermore, it is clear that periodic models may well describe the time series considered since the elements in several rows of the Π_1 are quite different and the values of $\Lambda_{ss}^{1/2}$ vary across the seasons.

In Engle et al. (1993) it is found via the HEGY method that for y_t the roots ± 1, and $\pm i$ and for c_t the roots ± 1 cannot be rejected. The results of the VQ procedure are displayed in Table 5. The number of observations equals 26, and hence the critical values in Table A.1 are used. The eigenvalues indicate that for y_t the hypothesis $r = 3$ cannot be rejected. Moreover, the null hypothesis of cointegration of the sequential quarters with parameters $(1, -1)$ can neither be rejected. This implies that a Δ_1 filter for the y_t seems appropriate. For c_t , the hypothesis that $r = 2$ cannot be rejected at a 10% level. The restrictions H_{21} and H_{22} are rejected, though. The univariate consumption series may therefore be described by a univariate periodically integrated model. Given this last result, one can easily recognize that the consumption and income series for Japan may not fit into a seasonal cointegration model as in Engle et al. (1993), but may possibly be more appropriately modeled with a periodic cointegration model. An estimation method for the latter model is proposed in Franses and Kloek (1991).

Table 5
VQ results for the Japanese income and consumption series, $N = 26$

Income	$\lambda_1 = 0.854^a$	$Q_1(3) = 3.297$	$r = 3$	$Q(H_{31}) = 0.753$
	$\lambda_2 = 0.609^b$	$Q_1(2) = 18.587^c$		$Q(H_{32}) = 23.167^a$
	$\lambda_3 = 0.445^b$	$Q_1(1) = 42.989^b$		
	$\lambda_4 = 0.119$	$Q_1(0) = 93.052^a$		
Consumption	$\lambda_1 = 0.951^a$	$Q_1(3) = 6.694^d$	$r = 2$	$Q(H_{21}) = 18.035^a$
	$\lambda_2 = 0.527^d$	$Q_1(2) = 16.372^d$		$Q(H_{22}) = 49.135^a$
	$\lambda_3 = 0.311$	$Q_1(1) = 35.854^c$		
	$\lambda_4 = 0.227^d$	$Q_1(0) = 114.322^b$		

The expressions for $Q_1(r)$ and that related to the λ_i can be found in the text. Critical values are displayed in Table A.1 of the Appendix. The Q statistics for the hypotheses H_{3i} asymptotically follow $\chi^2(3)$ distributions, and the Q statistics for the hypotheses H_{2i} asymptotically follow $\chi^2(4)$ distributions, where $i = 1, 2$.

[a] Significant at a 1% level.
[b] Significant at a 5% level.
[c] Significant at a 10% level.
[d] Significant at a 20% level.

6. Concluding remarks

The multivariate approach to modeling univariate seasonal time series proposed in this paper amounts to considering an autoregressive model for the vector containing the annual observations per season. In case the elements of this vector are integrated, an application of the Johansen cointegration method yields insights in whether a time series contains a nonseasonal unit root and/or seasonal unit roots, or whether it is periodically integrated. Hence, this application extends the HEGY method by allowing for periodically varying coefficients. Since the critical values of the Johansen method are of an asymptotic nature, and our application deals with small samples, new critical values are tabulated. From Monte Carlo simulations it appears that the success rate of our method is satisfactory, even in cases where the multivariate time series model is highly overparameterized. An application to the Japanese data in Engle et al. (1993) yields new insights in the univariate properties of these series, i.e., the consumption and income series may not be seasonally integrated, and hence a seasonal cointegration model may not be adequately representing the bivariate series.

An often applied transformation for nonstationary seasonal time series is the double filter, i.e., a seasonal and a first-order differencing filter. Such a filter is appropriate in case the annual time series are integrated of order 2 and certain restrictions on the cointegration relations between the first-order transformed annual series are valid. Further, an extension to, e.g., monthly time series is in principle relatively straightforward. The expressions in Beaulieu and Miron (1993) and Franses (1991), where the HEGY method is applied to the monthly case, can then be used. Similarly, an extension to, e.g., the bivariate case is easily made. However, as with all methods for testing for cointegration, the inclusion of more variables has a deteriorating effect on the empirical performance. Therefore, the VQ approach in the present paper may be most suitable as a tool for univariate data analysis, and serve as a starting-point for building periodic cointegration models.

Appendix

Critical values of the Johansen cointegration tests

This appendix contains the critical values of the Johansen cointegration tests for sample sizes 25 and 50. These quantiles are based on 10,000 replications, and the test statistics are computed from the original formulas in Johansen and Juselius (1990). For each sample size, the tables correspond to the tables numbered as A.2 and A.3 in Johansen and Juselius (1990). See also that paper for more details.

Table A.1
Sample size is 25; the data-generating process contains no trend; and the constant term μ is unrestricted

Dim	50%	80%	90%	95%	97.5%	99%	Mean	Var
(A) *Maximal eigenvalue*								
1	2.43	4.93	6.70	8.29	9.91	12.09	3.06	7.36
2	7.86	11.38	13.70	15.75	17.88	20.51	8.54	14.76
3	13.80	18.16	20.90	23.26	25.66	28.57	14.46	22.80
4	20.36	25.56	28.56	31.66	34.47	37.61	21.07	32.97
(B) *Trace*								
1	2.43	4.93	6.70	8.29	9.91	12.09	3.06	7.36
2	9.78	13.99	16.56	18.90	21.26	23.70	10.45	20.64
3	21.79	27.69	31.22	34.37	37.44	40.98	22.54	42.57
4	39.32	47.10	51.59	55.92	59.60	64.33	40.09	77.08

Table A.2
Sample size is 25; the data-generating process contains no trend; and the constant term μ is restricted by $\mu = \alpha\beta_0$

Dim	50%	80%	90%	95%	97.5%	99%	Mean	Var
(A) *Maximal eigenvalue*								
1	3.55	6.01	7.72	9.35	10.97	12.09	4.14	7.08
2	8.82	12.15	14.40	16.51	18.36	20.56	9.36	14.33
3	14.56	18.89	21.56	23.90	26.21	29.44	15.24	22.84
4	21.01	26.15	29.26	32.18	34.74	38.12	21.69	32.79
(B) *Trace*								
1	3.55	6.01	7.72	9.35	10.97	12.90	4.14	7.08
2	11.95	16.09	18.63	20.96	22.78	25.71	12.55	20.81
3	25.01	31.01	34.44	37.85	40.56	44.60	25.74	44.16
4	43.40	51.38	55.78	59.98	63.51	67.74	44.20	78.19

Table A.3
Sample size is 50; the data-generating process contains no trend; and the constant term μ is unrestricted

Dim	50%	80%	90%	95%	97.5%	99%	Mean	Var
(A) *Maximal eigenvalue*								
1	2.44	4.89	6.40	8.09	9.54	11.39	3.02	6.73
2	7.71	11.09	13.15	15.18	16.98	19.18	8.30	13.42
3	13.34	17.44	19.94	22.29	24.31	26.98	13.92	20.61
4	19.00	23.78	26.63	29.15	31.93	35.20	19.64	28.05
(B) *Trace*								
1	2.44	4.89	6.40	8.09	9.54	11.39	3.02	6.73
2	9.51	13.57	16.06	18.25	20.13	22.81	10.18	19.02
3	21.07	26.78	30.07	32.94	35.59	39.10	21.75	38.87
4	36.86	44.10	48.25	51.98	55.88	59.94	37.59	65.89

Table A.4
Sample size is 50; the data-generating process contains no trend; and the constant term μ is restricted by $\mu = \alpha\beta_0$

Dim	50%	80%	90%	95%	97.5%	99%	Mean	Var
(A) *Maximal eigenvalue*								
1	3.49	5.91	7.59	9.22	10.93	13.06	4.10	7.05
2	8.58	12.05	14.05	15.99	17.92	20.60	9.19	13.94
3	13.99	18.01	20.57	23.01	25.24	27.95	14.59	20.88
4	19.63	24.44	27.23	29.79	32.47	35.33	20.32	27.78
(B) *Trace*								
1	3.49	5.91	7.59	9.22	10.93	13.06	4.10	7.05
2	11.74	15.79	18.25	20.61	22.85	25.49	12.32	20.40
3	24.16	29.91	33.08	36.33	39.28	45.58	24.79	40.81
4	40.88	48.44	52.71	56.62	60.00	64.29	41.73	69.24

References

Beaulieu, J.J. and J.A. Miron, 1993, Seasonal unit roots in aggregate U.S. data, Journal of Econometrics 55, 305–328.

Birchenhall, C.R., R.C. Bladen-Hovell, A.P.L. Chui, D.R. Osborn, and J.P. Smith, 1989, A seasonal model of consumption, Economic Journal 99, 837–843.

Boswijk, H.P. and P.H. Franses, 1992, Dynamic specification and cointegration, Oxford Bulletin of Economics and Statistics 54, 369–381.

Engle, R.F., C.W.J. Granger, S. Hylleberg, and H.S. Lee, 1993, Seasonal cointegration: The Japanese consumption function, Journal of Econometrics 55, 275–298.

Engle, R.F. and B.S. Yoo, 1987, Forecasting and testing in cointegrated systems, Journal of Econometrics 35, 143–159.

Franses, P.H., 1990, A multivariate approach to modeling univariate seasonal time series, Econometric Institute report no. 9101 (Erasmus University, Rotterdam).

Franses, P.H., 1991, Seasonality, nonstationarity, and the forecasting of monthly time series, International Journal of Forecasting 7, 199–208.

Franses, P.H. and T. Kloek, 1991, A periodic cointegration model of quarterly consumption, Econometric Institute report no. 9172 (Erasmus University, Rotterdam).

Franses, P.H. and G. Romijn, 1993, Periodic integration of quarterly UK macroeconomic variables, International Journal of Forecasting 9, 467–476.

Gladyshev, E.G., 1961, Periodically correlated random sequences, Soviet Mathematics 2, 385–388.

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

Johansen, S., 1988, Statistical analysis of cointegration vectors, Journal of Economic Dynamics and Control 12, 231–254.

Johansen, S. and K. Juselius, 1990, Maximum likelihood estimation and inference on cointegration – With applications to the demand for money, Oxford Bulletin of Economics and Statistics 52, 169–210.

Osborn, D.R., 1988, Seasonality and habit persistence in a life-cycle model of consumption, Journal of Applied Econometrics 3, 255–266.

Osborn, D.R., 1991, The implications of periodically varying coefficients for seasonal time-series processes, Journal of Econometrics 28, 373–384.

Osborn, D.R., 1993, Discussion of Engle et al. (1993), Journal of Econometrics 55, 299–303.

Osborn, D.R. and J.P. Smith, 1989, The performance of periodic autoregressive models in forecasting seasonal U.K. consumption, Journal of Business and Economic Statistics 7, 117–128.

Pagano, M., 1978, On periodic and multiple autoregressions, Annals of Statistics 6, 1310–1317.

Tiao, G.C. and M.R. Grupe, 1980, Hidden periodic autoregressive-moving average models in time series data, Biometrika 67, 365–373.

Troutman, B.M., 1979, Some results in periodic autoregression, Biometrika 66, 365–373.