Small-sample improvements in the statistical analysis of seasonally cointegrated systems

Gianluca Cubadda[a,*], Pieter Omtzigt[b]

[a]*Dipartimento di Scienze Economiche Gestionali e Sociali, Università del Molise, Via De Sanctis, 86100 Campobasso, Italy*
[b]*Department of Quantitative Economics, University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands*

Abstract

New iterative reduced-rank regression procedures for seasonal cointegration analysis were proposed. The suggested methods are motivated by the idea that modelling the cointegration restrictions jointly at different frequencies may increase efficiency in finite samples. Monte Carlo simulations indicate that the new tests and estimators perform well with respect to already existing statistical procedures.
© 2004 Elsevier B.V. All rights reserved.

Keywords: Seasonal cointegration; Reduced rank regression

0. Introduction

It is still common to use seasonally adjusted time series in empirical analyses. However, a body of research has recently shown that seasonal adjustment may alter such time series properties such as invertibility (Maravall, 1995), linearity (Ghysels et al., 1996), cointegration (Granger and Siklos, 1995), and short-run comovements (Cubadda, 1999). Since there is convincing evidence of seasonal unit roots in common macroeconomic time series (Hylleberg et al., 1993), it is important to model them properly. The common practise of adding seasonal dummies to the set of regressors leads to misspecified models when seasonal unit roots are present (Abeysinghe, 1994). The analysis of seasonal cointegration, as

 * Corresponding author. Tel.: +39-0874404467; fax: +39-0874311124.
 E-mail address: gianluca.cubadda@uniroma1.it (G. Cubadda)

first proposed in Hylleberg et al. (1990), has gained recent interest, see *e.g.*, the thorough surveys by Franses and McAleer (1998) and Brendstrup et al. (2004). Indeed, Lof and Franses (2001) shows that seasonal cointegration models tend to yield better forecasts than alternative models of seasonal data.

A set of seasonally cointegrated time series may be represented by a seasonal version of the Error-Correction Model (ECM), see inter alia Ahn and Reinsel (1994). The statistical analysis of the seasonal ECM can be complicated by the existence of cointegration relationships that vary over the frequencies. Moreover, the cointegration vectors at frequencies other than zero and π are generally polynomial. However, Lee (1992) shows that asymptotically optimal inference on seasonal cointegration may be conducted by Reduced-rank Regression (RR) analyses separately for each frequency. Unfortunately, Lee's method applies only to the peculiar case of synchronous cointegration at frequencies different from zero and π. Based on Boswijk (1995), Johansen and Schaumburg [henceforth, JS] (1998) provides a rather involved iterative procedure for detecting and estimating dynamic cointegration relationships at complex root frequencies. Recently, Cubadda (2001) shows that an estimator and a test statistic that are asymptotically equivalent to those proposed by JS can be obtained by RR between complex-valued data.

Although the RR approach considerably simplifies seasonal cointegration analysis, it suffers from two main limitations: it ignores the fact that complex unit roots occur in conjugate pairs in real-valued data, and Maximum Likelihood (ML) analysis of the seasonal ECM requires the cointegration vectors at different frequencies to be jointly estimated (JS, 1998).

The goal of this paper is twofold. First, we propose an iterative RR procedure that allows the cointegration restrictions at the conjugate complex unit root frequencies to be modelled simultaneously. Second, we extend our new procedure to estimate the cointegration vectors jointly at the zero and seasonal frequencies. We investigate the small-sample properties of the proposed methods through simulations and find that they often perform better with respect to separate RR analyses at the different frequencies. As modelling non-stationary seasonality increases substantially the number of parameters of VAR models, we think that new methods are of practical value.

This paper is organized as follows. Section 1 reviews the relevant representations of seasonally cointegrated time series. Section 2 introduces the new tests and estimators. Section 3 compares the performances of our procedures with existing ones by Monte Carlo simulations. Section 4 concludes.

1. Error-correction models for seasonally cointegrated time series

Let X_t be an *n*-vector time series satisfying

$$\Pi(L)X_t = \Phi D_t + \varepsilon_t, \tag{1}$$

where $\Pi(L)$ is a *p*-order polynomial matrix with $\Pi_0 = I_n$, the ε_t are i.i.d. $N_n(0, \Omega)$, the initial values $Y_{-p+1}, \ldots, Y_0$ are fixed, and D_t is a deterministic kernel that may contain a constant, a linear trend, and a set of seasonal dummies.

Suppose for simplicity that the X_t are observed on a quarterly basis. We know (JS, 1998) that if the series are cointegrated of order (1,1) at frequencies 0, π, $\frac{\pi}{2}$, and $\frac{3\pi}{2}$ Eq. (1) may be rewritten in the following ECM:

$$\tilde{\Gamma}(L)\underbrace{(1-L^4)X_t}_{X_{0,t}}=\Phi D_t + \alpha_1\beta_1'\underbrace{(1+L+L^2+L^3)X_{t-1}}_{X_{1,t-1}}$$
$$+\alpha_2\beta_2'\underbrace{(1-L+L^2-L^3)X_{t-1}}_{X_{2,t-1}}$$
$$+\alpha_*\overline{\beta}_*'\underbrace{(-i-L+iL^2+L^3)X_{t-1}}_{X_{*,t-1}}$$
$$+\overline{\alpha}_*\beta_*'\underbrace{(i-L-iL^2+L^3)X_{t-1}}_{\overline{X}_{*,t-1}}+\varepsilon_t, \tag{2}$$

where $\alpha_1\beta_1' = -\frac{1}{4}\Pi(1)$, $\alpha_2\beta_2' = \frac{1}{4}\Pi(-1)$, $\alpha_*\overline{\beta}_*' = -\frac{1}{4}\Pi(i)$, α_j and β_j are $n \times r_j$-matrices with rank equal to r_j for $j = 1, 2$, α_* and β_* are complex $n \times r_3$-matrices with rank equal to r_3, and $\overline{C}$ denotes the complex conjugate C, $\tilde{\Gamma}_0 = I_n$, and $\tilde{\Gamma}_k = -\sum_{l=1}^{[(p-k)/4]}\Pi_{k+4l}$ for $k = 1, 2, \ldots, p-4$.

Notice that four cointegrating relationships are present in the ECM (2). Indeed, β_1 and β_2 are, respectively, the cointegration matrices at frequencies 0 and π, whereas the conjugate complex cointegration matrices $\overline{\beta}_*$ and β_* are respectively associated with frequencies $\pi/2$ and $3\pi/2$.

Below we refer to (2) when conducting statistical inference on the various cointegration matrices. However, since complex valued coefficients are not amenable to economic interpretation, we observe that (2) can be rewritten more neatly as

$$\Gamma(L)X_{0,t}=\Phi D_t + \alpha_1\beta_1'X_{1,t-1} + \alpha_2\beta_2'X_{2,t-1}$$
$$+ (\alpha_4 - \alpha_3 L)(\beta_3' - \beta_4' L)\underbrace{(1-L^2)X_{t-1}}_{X_{3,t-1}}+\varepsilon_t, \tag{3}$$

where $\Gamma_1 = \tilde{\Gamma}_1 + \alpha_3\beta_4'$, and $\alpha_*\overline{\beta}_*' \equiv \frac{1}{2}(\alpha_3 + \alpha_4 i)(\beta_3' - \beta_4' i)$.

Representation (3) is entirely real-valued and it exhibits a polynomial cointegration matrix, namely $(\beta_3 - \beta_4 L)$, and an intertemporal loading matrix, namely $(\alpha_4 - \alpha_3 L)$, for the annual frequency.

2. Optimal inference on seasonal cointegration

In this section we introduce some new statistical tools for seasonal cointegration analysis. Specifically, we offer various methods for detecting and estimating polynomial cointegration vectors and an iterative procedure for ML estimation of the cointegration vectors at the zero and seasonal frequencies. All the proposed inferential procedures are motivated by the idea that joint modelling of the cointegration restrictions at the different frequencies may increase efficiency in finite samples.

2.1. Statistical analysis of cointegration at the complex root frequencies

Cubadda (2001) observes that asymptotically optimal inference on cointegration at complex root frequencies may be obtained through partial RR applied to model (2).

At frequency $\pi/2$, this RR procedure goes as follows: We regress $X_{0,t}$, $X_{1,t-1}$, $X_{2,t-1}$, and $X_{*,t-1}$ on $(D_t, X_{0,t-1}, X_{0,t-2}, \ldots, X_{0,t-p+4})$ to obtain, respectively, the residuals $R_{0,t}$, $R_{1,t}$, $R_{2,t}$, and $R_{*,t}$. These residuals asymptotically satisfy

$$R_{0,t} = \alpha_1 \beta_1' R_{1,t} + \alpha_2 \beta_2' R_{2,t} + \alpha_* \overline{\beta}_*' R_{*,t} + \overline{\alpha}_* \beta_*' \overline{R}_{*,t} + \varepsilon_t. \tag{4}$$

Since the process $R_{*,t}$ is asymptotically uncorrelated with $R_{1,t}$, $R_{2,t}$, and $\overline{R}_{*,t}$ we can safely ignore reduced rank restrictions at frequencies different from the one of interest. Hence, we solve

$$CanCor\{R_{0,t}, R_{*,t} \mid R_{1,t}, R_{2,t}, \overline{R}_{*,t}\}, \tag{5}$$

where $CanCor(Y, X \mid Z)$ denotes the partial canonical correlations between the elements of Y and X conditional on Z.

A test for the null hypothesis that there exist at most r_3 cointegration vectors at the annual frequency is based on the statistic

$$Q_1(r_3|n) = -2T \sum_{l=r_3+1}^{n} \ln(1 - \hat{\lambda}_l), \quad r_3 = 1, \ldots, n,$$

where $\hat{\lambda}_l$ is the lth largest squared canonical correlation coming from the solution for (5).

The test statistic $Q_1(r_3|n)$ converges weakly in distribution to the same limit as the Likelihood Ratio (LR) test statistic; that is

$$\mathrm{tr}\left\{\int_0^1 \mathrm{d}B_c(u)\overline{F}_c'(u)\left[\int_0^1 F_c(u)\overline{F}_c'(u)\,\mathrm{d}u\right]^{-1}\int_0^1 F_c(u)\,\mathrm{d}\overline{B}_c'(u)\right\}, \tag{6}$$

where $\mathrm{tr}\{\cdot\}$ denotes trace, $B_c(u)$ is the standard complex-valued Brownian motion of dimension $(n - r_3)$, and $F_c(u) = B_c(u)$ if D_t does not include a set of seasonal dummies, and $F_c(u) = B_c(u) - \int_0^1 B_c(v)\,\mathrm{d}v$ otherwise.

Moreover, the eigenvectors associated with the r_3 largest eigenvalues $\hat{\lambda}_1, \ldots, \hat{\lambda}_{r_3}$ are T-consistent estimators for the complex cointegration matrix $\overline{\beta}_*$.

A Monte Carlo study in Cubadda (2001) indicates that the JS and RR procedures have similar performances in small samples. There is evidence of a slight superiority for the JS procedure when testing, but not for estimation. This apparent paradox is explained by the fact that the test based on $Q_1(r_3|n)$ does not use the information that the cointegration restrictions at frequency $\pi/2$ apply also at the aliasing frequency $3\pi/2$. Although there is no asymptotic gain in exploiting this information, it may well matter with finite samples.

Hence, we propose the following testing procedure. Solve (5) to obtain the RR estimate $\hat{\overline{\beta}}_*$ of the r_3 complex cointegration vectors. Then regress $R_{0,t}$ on $(R_{*,t}' \hat{\overline{\beta}}_*, \overline{R}_{*,t}' \hat{\beta}_*, R_{1,t}', R_{2,t}')'$

and compute the residual covariance matrix $\Omega(\hat{\beta}_*)$. The proposed test statistic is

$$Q_2(r_3|n) = T \, \log\left(\frac{\left|\Omega(\hat{\beta}_*)\right|}{|\Omega(I_n)|}\right), \quad r_3 = 1, \ldots, n.$$

Since the estimator $\hat{\beta}_*$ is asymptotically equivalent to the ML estimator (Cubadda, 2001), $Q_2(r_3|n)$ has the same limiting distribution (6) as the LR test statistic.

In a similar spirit as JS (1998), we also consider a LR test that is based on an iterative estimation procedure called Alternating Reduced-rank Regression (ARR). The ARR procedure, which increases the likelihood function in each step, goes as follows:

(1) Estimate $\overline{\beta}_*$ by solving (5)
(2) For fixed $\overline{\beta}_* = \hat{\overline{\beta}}_*$, obtain $\hat{\beta}_*$ as the eigenvectors associated with the r_3 largest eigenvalues coming from the solution of

$$CanCor\{R_{0,t}, \overline{R}_{*,t} \mid R_{1,t}, R_{2,t}, \overline{\beta}'_* R_{*,t}\}.$$

(3) For fixed $\beta_* = \hat{\beta}_*$, obtain $\hat{\overline{\beta}}_*$ as the eigenvectors associated with the r_3 largest eigenvalues coming from the solution of

$$CanCor\{R_{0,t}, R_{*,t} \mid R_{1,t}, R_{2,t}, \beta'_* \overline{R}_{*,t}\}.$$

(4) Repeat 2 and 3 until numerical convergence occurs.

The associated test statistic is

$$Q_3(r_3|n) = -2T \sum_{l=r_3+1}^{n} \ln(1 - \hat{\eta}_l), \quad r_3 = 1, \ldots, n, \tag{7}$$

where $\hat{\eta}_l$ indicates the lth largest squared canonical correlation coming from the last iteration of the above switching procedure.

Moreover, the eigenvectors associated with the r_3 largest eigenvalues $\hat{\eta}_1, \ldots, \hat{\eta}_{r_3}$ are the ARR estimator of the complex cointegration matrix $\overline{\beta}_*$.

Notice that both the new tests may be easily adapted to include the seasonal dummies in the seasonal error-correction terms, as suggested in Franses and Kunst (1999). However, in this case their asymptotic distribution will be the one tabulated in Table 2 of JS (1998).

2.2. ML estimation of seasonal cointegration vectors

It is noted in JS (1998) that the ML estimator of the complex ECM requires estimating the cointegration vectors jointly at the zero and seasonal frequencies. However, the switching algorithm proposed there is computationally cumbersome since it typically involves a large number of variables. Hence, a simpler estimation strategy is suggested that focuses on each frequency separately. Although such strategy leads to asymptotically optimal estimation

of the various cointegration relationships, there may be some efficiency loss with finite samples.

We now consider a convenient procedure for the simultaneous ML estimation of the various cointegration vectors that appear in the complex-valued ECM (2). In view of Eq. (4), we propose the following iterative scheme that increases the likelihood in each step:

(1) Fix the various cointegration ranks r_j for $j = 1, 2, 3$ and let $\hat{\beta}_i$ for $i = 1, 2, *$ denote the estimates of the cointegration vectors obtained by RR at the various frequencies.

(2) For fixed $\beta_2 = \hat{\beta}_2$, $\beta_* = \hat{\beta}_*$, and $\overline{\beta}_* = \hat{\overline{\beta}}_*$, obtain $\hat{\beta}_1$ as the eigenvectors associated with the r_1 largest eigenvalues coming from the solution of

$$CanCor\{R_{0,t}, R_{1,t} \mid \beta_2' R_{2,t}, \beta_*' \overline{R}_{*,t}, \overline{\beta}_*' R_{*,t}\}.$$

(3) For fixed $\beta_1 = \hat{\beta}_1$, $\beta_* = \hat{\beta}_*$, and $\overline{\beta}_* = \hat{\overline{\beta}}_*$, obtain $\hat{\beta}_2$ as the eigenvectors associated with the r_2 largest eigenvalues coming from the solution of

$$CanCor\{R_{0,t}, R_{2,t} \mid \beta_1' R_{1,t}, \beta_*' \overline{R}_{*,t}, \overline{\beta}_*' R_{*,t}\}.$$

(4) For fixed $\beta_1 = \hat{\beta}_1$, $\beta_2 = \hat{\beta}_2$ and $\overline{\beta}_* = \hat{\overline{\beta}}_*$, obtain $\hat{\beta}_*$ as the eigenvectors associated with the r_3 largest eigenvalues coming from the solution of

$$CanCor\{R_{0,t}, \overline{R}_{*,t} \mid \beta_1' R_{1,t}, \beta_2' R_{2,t}, \overline{\beta}_*' R_{*,t}\}.$$

(5) For fixed $\beta_1 = \hat{\beta}_1$, $\beta_2 = \hat{\beta}_2$, and $\beta_* = \hat{\beta}_*$, obtain $\hat{\overline{\beta}}_*$ as the eigenvectors associated with the r_3 largest eigenvalues coming from the solution of

$$CanCor\{R_{0,t}, R_{*,t} \mid \beta_1' R_{1,t}, \beta_2' R_{2,t}, \beta_*' \overline{R}_{*,t}\}.$$

(6) Repeat 2–5 until numerical convergence occurs.

We shall refer to this iterative procedure as the Generalized Alternating Reduced-rank Regression (GARR) estimator of the cointegration vectors $(\beta_1, \beta_2, \beta_*)$. Remarkably, the GARR approach can easily be extended to take account of more complex root frequencies, like in the case of monthly data.

3. Monte Carlo experiments

Here we conduct a Monte Carlo study to evaluate the small-sample properties of the different statistical procedures for seasonal cointegration analysis. In particular, we first investigate size and power of the various tests for cointegration at the annual frequencies. Then, we analyze the efficiency of the RR and ARR estimators of the annual cointegration vectors. Finally, we compare the usual RR estimators of the various seasonal cointegration vectors with the ones obtained by GARR.

3.1. Size and power of annual cointegration tests

To evaluate the small-sample performances of the tests statistics $Q_1(r_3|n)$, $Q_2(r_3|n)$ and $Q_3(r_3|n)$, we extend to the seasonal case a Data Generating Process (DGP) which has been used extensively in the zero-frequency cointegration literature (e.g., Gonzalo, 1994; Haug, 1996). The bivariate DGP is

$$X_{0,t} = \begin{bmatrix} -0.2 \\ 0 \end{bmatrix} [\, 1 \quad -1\,] X_{1,t-1} + \begin{bmatrix} 0.2 \\ 0 \end{bmatrix} [\, 1 \quad -1\,] X_{2,t-1}$$
$$+ \begin{bmatrix} \gamma \\ 0 \end{bmatrix} [\, 1 \quad -L\,] X_{3,t-1} + \varepsilon_t, \tag{8}$$

or, equivalently, in a complex-valued format

$$X_{0,t} = \begin{bmatrix} -0.2 \\ 0 \end{bmatrix} [\, 1 \quad -1\,] X_{1,t-1} + \begin{bmatrix} 0.2 \\ 0 \end{bmatrix} [\, 1 \quad -1\,] X_{2,t-1}$$
$$+ \begin{bmatrix} i\gamma/2 \\ 0 \end{bmatrix} [\, 1 \quad -i\,] X_{*,t-1} + \begin{bmatrix} -i\gamma/2 \\ 0 \end{bmatrix} [\, 1 \quad i\,] \overline{X}_{*,t-1} + \varepsilon_t, \tag{9}$$

where $t = 1, 2, \ldots, T$ and

$$E(\varepsilon_t \varepsilon_t') = \begin{bmatrix} 1 & \sigma\rho \\ \sigma\rho & \sigma^2 \end{bmatrix}.$$

The design parameters are

$$\gamma = (0, 0.2), \quad \rho = (-0.5, 0, 0.5), \quad \sigma^2 = (0.5, 1, 2), \quad \text{and} \quad T = (50, 100, 200).$$

Some comments on the choice of the parameter values are in order. The cointegration rank at the annual frequencies is zero when $\gamma = 0$ and is one otherwise. When $\gamma = 0.2$ slow adjustment to equilibrium takes place. This slow adjustment assures that the small-sample properties of the tests statistics differ substantially from the large-sample ones. The value of σ determines the sizes of the non-stationary components in the system. Hence, the small-sample behavior of the tests statistics can be evaluated under rather different signal-to-noise ratios. The various non-stationary components are strictly exogenous when $\rho = 0$ and weakly exogenous when $\rho = \pm 0.5$. However, there is no loss of generality since exogeneity assumptions are not relevant for the comparison of full-information procedures. By letting ρ vary, we also check if the tests statistics are sensitive to the degree and sign of correlation between the innovations.

In all the simulations, 10000 series are generated with initial values set to zero. The first 50 observations are discarded to eliminate dependence from the starting conditions. Based on preliminary experiments, numerical convergence of the ARR procedure is assumed to be reached after six iterations. Notice that a constant and seasonal dummies are included unrestrictedly in the estimated model. Hence, the asymptotic distribution of the three test statistics is the one tabulated in Table 1 of Cubadda (2001).

In Tables 1 and 2 the acceptance frequencies at the 5% level tests based on $Q_1(r_3|n)$, $Q_2(r_3|n)$ and $Q_3(r_3|n)$ are reported, both for $r_3 = 0$ and $r_3 = 1$. All the results are based on the 5% asymptotic critical value. Notice that the last two test statistics assume the same

Table 1
Acceptance Percentages of 5% level tests for the annual cointegration rank r_3 DGP: no cointegration ($\gamma = 0$)

	r_3	$T = 50$			$T = 100$		
		Q_1	Q_2	Q_3	Q_1	Q_2	Q_3
$\sigma^2 = 0.5$	0	91.80	92.17	92.17	94.32	94.38	94.38
$\rho = -0.5$	1	7.69	7.34	7.37	5.32	5.25	5.25
	2	0.51	0.49	0.46	0.37	0.38	0.37
$\sigma^2 = 0.5$	0	91.57	92.03	92.03	94.24	94.26	94.26
$\rho = 0$	1	7.93	7.45	7.49	5.39	5.36	5.36
	2	0.50	0.52	0.48	0.37	0.38	0.38
$\sigma^2 = 0.5$	0	91.30	91.75	91.75	94.15	94.25	94.25
$\rho = 0.5$	1	8.17	7.72	7.77	5.47	5.37	5.37
	2	0.53	0.53	0.48	0.38	0.38	0.38
$\sigma^2 = 1$	0	92.08	92.33	92.33	94.45	94.47	94.47
$\rho = -0.5$	1	7.40	7.17	7.18	5.19	5.15	5.15
	2	0.52	0.50	0.49	0.36	0.38	0.38
$\sigma^2 = 1$	0	91.76	92.08	92.08	94.33	94.23	94.23
$\rho = 0$	1	7.73	7.40	7.44	5.29	5.40	5.40
	2	0.51	0.52	0.48	0.38	0.37	0.37
$\sigma^2 = 1$	0	91.46	91.93	91.93	94.24	94.24	94.24
$\rho = 0.5$	1	8.04	7.53	7.59	5.38	5.36	5.37
	2	0.50	0.49	0.48	0.38	0.40	0.39
$\sigma^2 = 2$	0	92.32	92.45	92.45	94.43	94.52	94.52
$\rho = -0.5$	1	7.19	7.05	7.07	5.22	5.07	5.08
	2	0.49	0.50	0.48	0.35	0.41	0.40
$\sigma^2 = 2$	0	92.07	92.24	92.24	94.37	94.40	94.40
$\rho = 0$	1	7.39	7.26	7.30	5.26	5.21	5.21
	2	0.54	0.50	0.46	0.37	0.39	0.39
$\sigma^2 = 2$	0	91.78	91.95	91.95	94.37	94.39	94.39
$\rho = 0.5$	1	7.71	7.57	7.58	5.26	5.21	5.21
	2	0.51	0.48	0.47	0.37	0.40	0.40

value when $r_3 = 0$. It is apparent that for a given sample size the most important parameter in determining the performance of the test statistics is σ. As expected, all the tests perform better when the variance of the common seasonal component becomes larger. The degree of correlation between the innovations has little effect on the size but is beneficial to the power of the tests. In particular, power is higher when this correlation is negative.

In comparative terms, we consider a difference between the acceptance frequencies of two different tests as significant when it is larger than twice the Monte Carlo standard error at the nominal 5% level, i.e., 0.44%.

Table 2
Acceptance Percentages of 5% level tests for the annual cointegration rank r_3 DGP: one cointegration vector ($\gamma = 0.2$)

	r_3	$T = 50$			$T = 100$		
		Q_1	Q_2	Q_3	Q_1	Q_2	Q_3
$\sigma^2 = 0.5$	0	10.91	11.89	11.89	0.01	0.00	0.00
$\rho = -0.5$	1	83.04	82.28	82.31	94.80	94.65	94.67
	2	6.05	5.83	5.80	5.19	5.35	5.33
$\sigma^2 = 0.5$	0	32.48	32.17	32.17	0.41	0.42	0.42
$\rho = 0$	1	62.90	63.03	63.10	94.29	94.20	94.21
	2	4.62	4.80	4.73	5.30	5.38	5.37
$\sigma^2 = 0.5$	0	24.50	23.50	23.50	0.13	0.11	0.11
$\rho = 0.5$	1	71.30	71.90	72.10	94.35	94.42	94.43
	2	4.20	4.60	4.40	5.52	5.47	5.46
$\sigma^2 = 1$	0	6.63	6.01	6.01	0.00	0.00	0.00
$\rho = -0.5$	1	87.12	87.80	87.83	94.47	94.42	94.42
	2	6.25	6.19	6.16	5.53	5.58	5.58
$\sigma^2 = 1$	0	26.09	22.18	22.18	0.08	0.05	0.05
$\rho = 0$	1	68.75	72.52	72.59	94.51	94.51	94.51
	2	5.16	5.30	5.23	5.41	5.44	5.44
$\sigma^2 = 1$	0	19.97	15.64	15.64	0.02	0.02	0.02
$\rho = 0.5$	1	75.09	79.32	79.44	94.26	94.22	94.23
	2	4.94	5.04	4.92	5.72	5.76	5.75
$\sigma^2 = 2$	0	2.62	1.71	1.71	0.00	0.00	0.00
$\rho = -0.5$	1	90.85	91.71	91.76	94.36	94.31	94.31
	2	6.53	6.58	6.53	5.64	5.69	5.69
$\sigma^2 = 2$	0	15.90	10.57	10.57	0.01	0.01	0.01
$\rho = 0$	1	78.40	83.72	83.81	94.26	94.35	94.38
	2	5.70	5.71	5.62	5.73	5.64	5.61
$\sigma^2 = 2$	0	13.49	7.53	7.53	0.00	0.00	0.00
$\rho = 0.5$	1	80.96	87.01	87.13	94.28	94.36	94.36
	2	5.55	5.46	5.34	5.72	5.64	5.64

When $r_3 = 0$, the tests statistic $Q_2(0|n) = Q_3(0|n)$ is better sized than $Q_1(0|n)$ for all the 9 experiments with $T = 50$. However, only two differences between the acceptance rates of the two tests are significant according to the adopted rule. With $T = 100$ the test statistics $Q_2(0|n)$ and $Q_3(0|n)$ are better sized than $Q_1(0|n)$ in seven experiments but no difference between the acceptance rates is significant.

When $r_3 = 1$, the experimental evidence is more clear-cut. Indeed, with $T = 50$ the test statistics $Q_2(1|n)$ and $Q_3(1|n)$ are significantly more powerful than $Q_1(1|n)$ in seven

experiments, whereas the opposite is true for one experiment only. Moreover, the test statistic $Q_3(1|n)$ dominates both the alternatives in eight experiments. The power of the three tests become similar with $T = 100$.

To save space, we do not report the tables for $T = 200$. Indeed, the performances of the three testing procedures are almost identical both when $r_3 = 0$ and $r_3 = 1$.

Overall, the new testing procedures appear to be superior to $Q_1(r_3|n)$ when a limited sample size is available. In particular, the test statistic $Q_3(r_3|n)$ leads more often to the right decision than both the alternatives. However, if one wishes to avoid the use of an iterative procedure for computational reasons, the test statistic $Q_2(r_3|n)$ generally performs better than $Q_1(r_3|n)$.

3.2. Efficiency of complex cointegration vectors estimators

It is not obvious how to analyze the finite sample properties of the RR and ARR estimators of the annual cointegration vectors in Eq. (9). Indeed, the asymptotic distribution of such estimators is the complex-valued analog to the distribution of the usual Johansen (1996) estimator, see JS (1998) and Cubadda (2001). Although Abadir and Paruolo (1997) shows that the normalized Johansen estimator has asymptotically finite second moments, the use of the minimum standard error criterion remains problematic due to the Cauchy-like tails of the exact distribution of such estimator, see Phillips (1994). Hence, we compare the RR and ARR estimators on the basis of three criteria, namely the standard error, the mean bias module of the normalized estimators, and the distance between the actual and nominal size of the associated LR tests for the null hypothesis that the annual cointegration vector is equal to the "true" one, i.e., $\overline{\beta}'_* = [1, -i]$. The last criterion is used as a dispersion measure that is robust to the possible presence of extreme outliers in the simulated distributions of the two estimators.

We rely on the previous Monte Carlo design with γ fixed to 0.2. From Table 3 we see that with $T = 50$ the RR estimator has the smallest standard error in seven experiments whereas the ARR is less biased in all the experiments. According to the criterion of the acceptance rate of the LR test for $\overline{\beta}'_* = [1, -i]$ at the 5% level, the ARR test is always better sized even if no difference between the rejection rates is significant. The results in Table 4 indicate that the performances of the two estimators become similar when $T = 100$.

Though outside the scope of this paper, we observe that the actual rejection probabilities of both of the LR tests are far away from the nominal size. This means that some kind of small-sample correction, such as a Bartlett correction or bootstrap, is called for, see e.g., Johansen (2000) and Omtzigt and Fachin (2002) for the zero-frequency case.

Overall, the efficiency gains of the ARR over the RR estimator appear quite limited. Interestingly enough, a similar conclusion is found in a previous Monte Carlo study in Cubadda (2001) where the RR estimator is compared with the JS switching procedure. Notice that although we are not able to prove that the JS and the ARR algorithms numerically converge to the same limit, their simulated values are indistinguishable in our experiments. This implies that all the results that we find for the ARR can be practically referred to the JS switching procedure as well. However, we emphasize that ARR numerically converges much faster than JS.

Table 3
Standard errors (SE) and bias modules (BM) of the RR and ARR estimators of $\bar{\beta}_*$, rejection percentages (RP) of the 5% level LR-tests for the null hypothesis: $\bar{\beta}'_* = [1, -i]$

$T = 50$	RR estimator			ARR estimator		
	SE	BM	RP	SE	BM	RP
$\sigma^2 = 0.5$						
$\rho = -0.5$	0.548	0.0111	16.79	0.571	0.0095	16.50
$\rho = 0$	0.662	0.0231	19.58	0.651	0.0218	19.23
$\rho = 0.5$	0.509	0.0258	18.12	0.552	0.0230	17.97
$\sigma^2 = 1$						
$\rho = -0.5$	0.397	0.0054	15.79	0.413	0.0052	15.52
$\rho = 0$	0.436	0.0165	18.07	0.437	0.0150	17.66
$\rho = 0.5$	0.359	0.0181	16.98	0.362	0.0167	16.80
$\sigma^2 = 2$						
$\rho = -0.5$	0.246	0.0023	14.61	0.251	0.0020	14.30
$\rho = 0$	0.298	0.0082	16.21	0.296	0.0075	16.14
$\rho = 0.5$	0.251	0.0100	15.36	0.252	0.0095	15.32

Table 4
Standard errors (SE) and bias modules (BM) of the RR and ARR estimators of $\bar{\beta}_*$, rejection percentages (RP) of the 5% level LR-tests for the null hypothesis: $\bar{\beta}'_* = [1, -i]$

$T = 100$	RR estimator			ARR estimator		
	SE	BM	RP	SE	BM	RP
$\sigma^2 = 0.5$						
$\rho = -0.5$	0.202	0.0016	9.89	0.203	0.0018	9.63
$\rho = 0$	0.246	0.0070	10.89	0.240	0.0073	10.76
$\rho = 0.5$	0.206	0.0084	10.25	0.207	0.0087	10.27
$\sigma^2 = 1$						
$\rho = -0.5$	0.143	0.0012	9.22	0.143	0.0012	9.08
$\rho = 0$	0.168	0.0045	9.96	0.169	0.0049	10.07
$\rho = 0.5$	0.146	0.0058	9.53	0.147	0.0060	9.43
$\sigma^2 = 2$						
$\rho = -0.5$	0.101	0.0009	8.75	0.101	0.0009	8.64
$\rho = 0$	0.119	0.0025	8.94	0.119	0.0027	9.01
$\rho = 0.5$	0.104	0.0031	8.54	0.105	0.0032	8.63

3.3. Efficiency of seasonal cointegration vectors estimators

In order to examine if the GARR estimator provides any efficiency gain in small samples over separated RR analyses at the different frequencies, we make use of the previous Monte Carlo design with γ fixed to 0.2. Hence, we fix a cointegration vector proportional to $[1, -1]$

Table 5
Standard errors (SE) and biases of the Johansen and GARR estimators of β_1, rejection percentages (RP) of the 5% level LR-tests for the null hypothesis: $\beta'_1 = [1, -1]$

$T = 50$	Johansen estimator			GARR estimator		
	SE	Bias	RP	SE	Bias	RP
$\sigma^2 = 0.5$						
$\rho = -0.5$	2.903	−0.0374	12.90	5.878	−0.1087	12.49
$\rho = 0$	62.91	−0.6844	16.15	23.41	−0.2749	15.52
$\rho = 0.5$	29.36	0.2439	16.88	12.60	0.1691	16.26
$\sigma^2 = 1$						
$\rho = -0.5$	21.52	−0.3319	12.02	10.31	0.0707	11.74
$\rho = 0$	8.247	0.0776	14.98	7.568	0.0113	14.34
$\rho = 0.5$	11.76	0.2070	16.52	5.148	0.0539	16.15
$\sigma^2 = 2$						
$\rho = -0.5$	19.96	−0.2001	11.18	0.489	−0.0153	10.88
$\rho = 0$	2.713	−0.0078	13.53	1.693	−0.0091	13.29
$\rho = 0.5$	2.150	−0.0052	15.65	5.667	0.0534	15.46

and a slow adjustment to equilibrium at both frequencies zero and π. We assume that the various cointegration ranks are known. The comparison of the separated RR estimators with GARR is again evaluated according to the three criteria used above.

In Table 5 we report the results of the simulations of the usual Johansen (1996) estimator and GARR for the zero-frequency case with $T = 50$. Visual inspection of the biases and standard errors of both the estimators reveals a high incidence of abnormal values, which is likely due to the Cauchy-like tails of such estimators in finite samples. However, the GARR estimator has the smallest standard errors in seven experiments and the smallest bias in six experiments. Interestingly, the GARR test for $\beta'_1 = [1, -1]$ at the 5% level is always better sized than the RR one and three differences between the rejection rates of the two LR tests are indeed significant. From Table 6 we notice that when $T = 100$ the simulated distributions of the two estimators are much less affected by the presence of large outliers. In comparative terms, the GARR estimator is less dispersed and biased in six experiments and the GARR test is better sized in eight, although no difference between the rejection rates is significant.

From Table 7 we see the results of the comparison of the Lee (1992) estimator with GARR for the case of frequency π with $T = 50$. We again notice that the presence of large outliers in the simulated distributions of both the estimators inflates their biases and standard errors. Remarkably, although the Lee estimator exhibits the smallest standard error in five experiments and the GARR estimator is less biased in five experiments, efficiency gains and bias reductions are more relevant when the GARR estimator is superior. For instance, when the Lee estimator is less dispersed than GARR the average standard error ratio of these estimators is 0.805, whereas when the reverse is true the average standard error ratio of the GARR and Lee estimators is 0.434. Moreover, the GARR test for $\beta'_2 = [1, -1]$ is closer to the nominal size than RR in eight experiments even if just one difference between

Table 6
Standard errors (SE) and biases of the Johansen and GARR estimators of β_1, rejection percentages (RP) of the 5% level LR-tests for the null hypothesis: $\beta_1' = [1, -1]$

$T = 100$	Johansen estimator			GARR estimator		
	SE	Bias	RP	SE	Bias	RP
$\sigma^2 = 0.5$						
$\rho = -0.5$	0.252	−0.0038	8.17	0.298	−0.0000	8.21
$\rho = 0$	0.375	−0.0022	9.31	0.398	0.0002	9.20
$\rho = 0.5$	0.387	0.0168	9.98	0.340	0.0154	9.77
$\sigma^2 = 1$						
$\rho = -0.5$	0.174	−0.0029	8.01	0.172	−0.0008	7.99
$\rho = 0$	0.234	−0.0017	9.09	0.280	−0.0020	9.06
$\rho = 0.5$	0.490	−0.0001	9.74	0.264	0.0055	9.54
$\sigma^2 = 2$						
$\rho = -0.5$	0.122	−0.0023	7.76	0.121	−0.0006	7.70
$\rho = 0$	0.226	−0.0013	8.54	0.154	−0.0022	8.41
$\rho = 0.5$	1.176	−0.0130	9.03	0.300	−0.0030	8.99

Table 7
Standard errors (SE) and biases of the Lee and GARR estimators of β_2, rejection percentages (RP) of the 5% level LR-tests for the null hypothesis: $\beta_2' = [1, -1]$

$T = 50$	Lee estimator			GARR estimator		
	SE	Bias	RP	SE	Bias	RP
$\sigma^2 = 0.5$						
$\rho = -0.5$	1.389	−0.0371	12.95	1.231	−0.0466	12.80
$\rho = 0$	136.6	1.3292	15.49	6.303	−0.0042	15.52
$\rho = 0.5$	42.77	−0.4288	17.43	5.976	0.0086	16.97
$\sigma^2 = 1$						
$\rho = -0.5$	1.459	−0.0919	12.04	0.912	−0.0344	11.91
$\rho = 0$	2.507	−0.0501	14.16	2.822	−0.0243	13.90
$\rho = 0.5$	3.715	−0.0177	15.87	4.971	−0.0123	15.60
$\sigma^2 = 2$						
$\rho = -0.5$	0.358	−0.0149	11.64	0.331	−0.0174	11.55
$\rho = 0$	0.776	−0.0250	12.79	1.031	−0.0341	12.74
$\rho = 0.5$	1.599	−0.0673	13.72	3.161	−0.0801	13.83

the rejection rates is really significant. The results in Table 8 indicate that when $T = 100$ the simulated moments of both the estimators appear much less influenced by the Cauchy-like tails. Moreover, the two estimators perform very similarly in terms of standard error and LR test size, whereas the GARR estimator exhibits a smaller bias in seven experiments.

Table 8
Standard errors (SE) and biases of the Lee and GARR estimators of β_2, rejection percentages (RP) of the 5% level LR-tests for the null hypothesis: $\beta_2' = [1, -1]$

$T = 100$	Lee estimator			GARR estimator		
	SE	Bias	RP	SE	Bias	RP
$\sigma^2 = 0.5$						
$\rho = -0.5$	0.250	−0.0041	8.52	0.249	−0.0038	8.43
$\rho = 0$	0.305	−0.0075	9.68	0.304	−0.0034	9.71
$\rho = 0.5$	3.056	0.0247	9.89	0.380	0.0072	9.93
$\sigma^2 = 1$						
$\rho = -0.5$	0.173	−0.0016	7.98	0.173	−0.0020	8.02
$\rho = 0$	0.207	−0.0044	9.15	0.207	−0.0023	9.16
$\rho = 0.5$	0.190	−0.0049	9.48	0.190	−0.0007	9.47
$\sigma^2 = 2$						
$\rho = -0.5$	0.122	−0.0004	7.93	0.122	−0.0012	7.89
$\rho = 0$	0.143	−0.0021	8.67	0.143	−0.0017	8.66
$\rho = 0.5$	0.125	−0.0029	8.84	0.125	−0.0011	8.89

Table 9
Standard errors (SE) and bias modules (BM) of the GARR estimator of $\overline{\beta}$, rejection percentages (RP) of the 5% level LR-tests for the null hypothesis: $\overline{\beta}_*' = [1, -i]$

	$T = 50$			$T = 100$		
	SE	BM	RP	SE	BM	RP
$\sigma^2 = 0.5$						
$\rho = -0.5$	0.565	0.0129	16.45	0.202	0.0027	9.61
$\rho = 0$	0.625	0.0337	19.23	0.239	0.0109	10.80
$\rho = 0.5$	0.553	0.0373	17.81	0.206	0.0138	10.31
$\sigma^2 = 1$						
$\rho = -0.5$	0.390	0.0061	15.43	0.143	0.0205	9.08
$\rho = 0$	0.432	0.0213	17.87	0.168	0.0070	10.02
$\rho = 0.5$	0.365	0.0267	16.79	0.146	0.0097	9.58
$\sigma^2 = 2$						
$\rho = -0.5$	0.296	0.0025	14.50	0.101	0.0019	8.65
$\rho = 0$	0.299	0.0091	16.26	0.119	0.0033	9.12
$\rho = 0.5$	0.254	0.0144	15.42	0.104	0.0048	8.76

In Table 9 we report the results relative to the GARR estimator for the case of frequency $\pi/2$. These results must be compared with those corresponding to the RR and ARR estimators of $\overline{\beta}_*$ that are reported in Tables 3 and 4. Interestingly, we notice that even when $T = 50$ the three estimators do not exhibit anomalous standard errors and bias modules in our simulations. An intuitive explanation of this different behavior of RR-type estimators in

the complex-root case is that the occurrence of large outliers in the complex plane is more unlikely than on the real axis only. In comparative terms, with $T = 50$ the GARR estimator has the smallest standard error in three experiments and the largest bias module in all the experiments, and the GARR test for $\overline{\beta}'_* = [1, -i]$ has the best size in five experiments, but no difference between the rejection rates is significant. When $T = 100$ the performances of the three methods become similar even if the GARR estimator remains slightly more biased than both RR and ARR in all the experiments.

4. Conclusions

In this paper we have evaluated new statistical procedures for seasonally cointegrated systems. A Monte Carlo study has revealed that our new tests for the cointegration rank at the annual frequency outperform the trace test proposed in Cubadda (2001) for small sample sizes.

Moreover we have presented two novel iterative RR estimation procedures; the first allows for estimating jointly the conjugate complex cointegration vectors, the second is designed for the simultaneous ML estimation of all the cointegration vectors at the zero and seasonal frequencies. Our simulations suggest that the joint ML estimator is a clear improvement over the individual RR estimators of cointegration vectors at frequencies zero and π, whereas the efficiency gains of the new estimators appear more limited in the complex-root frequency case.

Acknowledgements

References

Abadir, K.M., Paruolo, P., 1997. Two mixed normal densities from cointegration analysis. Econometrica 65, 671–680.

Abeysinghe, T., 1994. Deterministic seasonal models and spurious regressions. J. Econometrics 61, 259–272.

Ahn, S.K., Reinsel, G.C., 1994. Estimation of partially non-stationary vector autoregressive models with seasonal behavior. J. Econometrics 62, 317–350.

Boswijk, H.P., 1995. Identifiability of cointegrated systems. Tinbergen Institute, pp. 78.

Brendstrup, B., Hylleberg, S., Nielsen, M., Skipper, L., Stentoft, L., 2004. Seasonality in economic models. Macroeconomic Dynamics 8, 362–394.

Cubadda, G., 1999. Common cycles in seasonal non-stationary time series. J. Appl. Econometrics 14, 273–291.

Cubadda, G., 2001. Complex reduced rank models for seasonally cointegrated time series. Oxford Bull. Econom. Statist. 63, 497–511.

Franses, P.H., Kunst, R.M., 1999. On the role of seasonal intercepts in seasonal cointegration. Oxford Bull. Econom. Statist. 61, 409–433.

Franses, P.H., McAleer, M., 1998. Cointegration analysis of seasonal time series. J. Econom. Surveys 12, 651–678.

Ghysels, E., Granger, C.W.J., Siklos, P.L., 1996. Is seasonal adjustment a linear or nonlinear data-filtering process? (with comments). J. Business Econom. Statist. 14, 374–397.

Gonzalo, J., 1994. Five alternative methods of estimating long-run equilibrium relationships. J. Econometrics 60, 203–223.

Granger, C.W.J., Siklos, P.L., 1995. Systematic sampling, temporal aggregation, seasonal adjustment, and cointegration: theory and evidence. J. Econometrics 66, 357–369.

Haug, A.A., 1996. Tests for cointegration. A Monte Carlo comparison. J. Econometrics 71, 89–115.

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

Hylleberg, S., Jorgensen, C., Sorensen, N.K., 1993. Seasonality in macroeconomic time series. Empirical Econom. 18, 321–335.

Johansen, S., 1996. Likelihood-Based Inference in Cointegrated Vector Autoregressive Models. Oxford University Press, Oxford.

Johansen, S., 2000. A Bartlett correction factor for tests on the cointegrating relations. Econometric Theory 16, 740–778.

Johansen, S., Schaumburg, E., 1998. Likelihood analysis of seasonal cointegration. J. Econometrics 88, 301–339.

Lee, H.S., 1992. Maximum likelihood inference on cointegration and seasonal cointegration. J. Econometrics 54, 1–47.

Lof, M., Franses, P.H., 2001. On forecasting cointegrated seasonal time series. Internat. J. Forecasting 17, 607–621.

Maravall, A., 1995. Unobserved components in economic time series. in: Pesaran, M.H., Wickens, M.R. (Eds.), Handbook of Applied Econometrics, Vol. 1. Macroeconomics. Blackwell, Oxford, pp. 12–72.

Omtzigt, P., Fachin, S., 2002. Bootstrapping and Bartlett corrections in the cointegrated VAR model. Technical report, University of Amsterdam.

Phillips, P.C.B., 1994. Some exact distribution theory for maximum likelihood estimators of cointegrating coefficients in error correction models. Econometrica 62, 73–93.