Canonical Correlation Statistics for Testing the Cointegration Rank in a Reversed Order

Jörg Breitung
Humboldt University Berlin
Institute of Statistics and Econometrics
Spandauer Strasse 1
D-10178 Berlin, Germany

December 1998

Abstract

In this paper a Canonical Correlation Analysis (CCA) is used to test the hypothesis $r = r_0$ against the alternative $r < r_0$. Such a test flips the null and alternative hypotheses of Johansen's LR test and can be used jointly with the LR test to construct a confidence set for the cointegration rank. As the latter test, our tests are based on the eigenvalues of a CCA between differences and lagged levels of a time series vector. The resulting test statistics can easily be adjusted for nuisance parameters using a nonparametric correction in the spirit of Phillips (1987, 1995). Monte Carlo simulations suggest that variants of the CCA statistic may have better properties than alternative tests and can be used as an alternative to Johansen's LR tests for determining the cointegration rank.

* I like to thank Helmut Lütkepohl and Uwe Hassler for helpful comments and suggestions. The research for this paper was carried out within Sonderforschungsbereich 373 at the Humboldt University Berlin and was printed using funds made available by the Deutsche Forschungsgemeinschaft.

1 Introduction

Kwiatkowski et al. (1992) (*henceforth: KPSS*) suggest a test for the null hypothesis that a time series is (trend) stationary against the alternative that the series is a first order integrated process. Such a test flips the null and alternative hypothesis of the unit root tests suggested by Dickey and Fuller (1979) and can be used to determine the degree of integration in a similar manner as the usual Dickey-Fuller type of tests.

In a multivariate setup, the LR test of Johansen (1988) can be employed to select the cointegration rank r in a vector autoregressive system by testing a sequence of hypotheses on the cointegration rank. There are two different strategies to do so. The "bottom-up" procedure starts with the hypothesis $H_0 : r = 0$ and proceed by increasing the rank until the null hypothesis cannot be rejected anymore. For the "top-down" procedure we start with testing $H_0 : r = n - 1$, where n is the dimension of the time series vector, and reduce the rank by one whenever the null hypothesis cannot be rejected. Both procedures are considered in Section 2. It is shown that by using a test procedure with a reversed set of hypotheses, the bottom-up strategy can be employed to construct a confidence set for the cointegration rank. In this paper such a test based on canonical correlations is suggested. Tests of the null hypothesis $r = 1$ against the alternative $r = 0$ was already suggested by Leybourne and McCabe (1994a), Shin (1994) and Harris and Inder (1994). Harris (1997) and Snell (1998) extend the test procedure to the case $r_0 > 1$ by using a principal components approach.

The principle for constructing these tests follows Stock (1994a) and can be demonstrated most easily in the context of a univariate unit root test. Assume that the univariate time series $\{y_t\}_{t=1}^{T}$ is generated by the AR(1) process:

$$y_t = \phi y_{t-1} + \varepsilon_t \ ,$$

where ε_t is a white noise process uncorrelated with y_{t-1}. Under the null hypothesis y_t is assumed to be stationary, that is, $|\phi| < 1$, whereas under the alternative $\phi = 1$ so that y_t is a random walk. An equivalent formulation of this null hypothesis can be obtained from considering the differenced process

$$\Delta y_t = \phi \Delta y_{t-1} + \varepsilon_t - \psi \varepsilon_{t-1} \ ,$$

where $\Delta = 1 - L$ and L is the backshift operator such that $L^k y_t = y_{t-k}$. If $|\phi| < 1$, then the differenced series has an ARMA(1,1) representation with $\psi = 1$. In other words, under the null hypothesis the moving average polynomial $(1 - \psi L)$ has a unit root. This reasoning suggests to test the null hypothesis that y_t is stationary by testing the MA representation of Δy_t against a unit root. This approach is used by Tanaka (1990), Tsay (1993), Saikkonen and Luukkonen (1993), Leybourne and McCabe (1994b), Choi (1994) and Breitung (1994), among others.

Tests for MA unit roots are based on the integrated (or partial sum) process $Y_t = \sum_{i=1}^{t} y_i$. Under the null hypothesis the series Y_t has an ARIMA(1,1,0) representation and under the alternative, Y_t is ARIMA(0,2,0). Therefore, (Dickey-Fuller type) unit root statistics can be applied using critical values from the opposite tail of the null distribution. For example, Tsay (1993) proposes to use the ordinary Dickey-Fuller t-statistic and KPSS (1992) is based on a Sargan and Bhargava (1983) type of unit-root statistic (see Stock (1994b) for an overview).

This test principle can be straightforwardly adopted to test the hypothesis that there exist $r = r_0$ cointegration relationships for the n-dimensional time series vector y_t against the alternative of $r < r_0$ cointegration relationships. The idea for a test of the cointegration rank with a reverse sequence of null hypothesis is to consider the cointegration properties of the n-dimensional partial sum process $Y_t = \sum_{i=1}^{t} y_i$. As in Johansen (1988) we use a test procedure based on a Canonical Correlation Analysis (CCA). However, whereas Johansen's LR test is based on a CCA between Δy_t and y_{t-1}, our test is based on a CCA between $\Delta Y_t = y_t$ and Y_{t-1}.

Alternative approaches suggested by Harris (1997) and Snell (1998) adopt a principal components approach. These tests are based on estimates of the cointegration vectors obtained from the eigenvectors of the matrix $\sum y_t y_t'$. There does not seem to be an ultimate reason for preferring one (the principle components) approach over the other (CCA) so it seems worthwhile to consider Johansen's CCA (or "reduced rank") approach to the partial sum process.

For the special case of testing $r_0 = n$ it is shown in Section 3 that the asymptotic null distributions of the test (corrected for nuisance parameters) is identical to the limiting distributions of Johansen's LR statistic for testing $r = 0$. For hypotheses with $r_0 < n$, the asymptotic null distribution is presented in Section 4. In contrast to Johansen's LR test, the asymptotic distribution depends on r and n. In Section 5 it is argued that the eigenvectors of a CCA between y_t and Y_{t-1} yields T-consistent estimates for the cointegration vectors. However, these estimates can be improved by using additional instruments.

It is well known (e.g. KPSS 1992, Leybourne and McCabe 1994b), that tests of the stationarity hypothesis suffer from the poor properties of the estimated nuisance parameters under the alternative hypothesis. In Section 6 we therefore suggest a modification similar to the one recommended in Breitung (1995) for the case of the KPSS test statistic. Indeed the simulation results reported in Section 7 demonstrate that this small sample modification yields a substantial improvement of the test. Furthermore the simulation results suggest that the augmented CCA statistic proposed in Section 5 is roughly as powerful as the test of Shin, although no prior normalization of the cointegration matrix is required for our test. In fact it is shown that if the normalization used for the latter test is invalid, the test is seriously biased. Finally, a four-variable cointegrated system is considered to assess the ability of the new test to select the cointegration rank. Section 8 considers an empirical example and Section 9 offers some concluding remarks. All proofs can be found in Appendix A.

Finally a word on the notational conventions applied in this paper. The symbol $\Rightarrow$ denotes weak convergence with respect to the associated probability measure and $[x]$ denotes the smallest integer $\leq x$. For notational convenience we write integrals such as $\int_0^1 B(a)da$ simply as $\int B$.

2 A Confidence Set for the Cointegration Rank

There are two mayor principles to select the cointegration rank by using Johansen's LR test procedure. First, we may apply a "general-to-specific" type of test procedure by starting with the hypothesis $H_0 : r = n - 1$ and proceed by *reducing* the rank as long as the LR test renders an *insignificant* test statistic. This procedure will be called "top-down procedure". Second, we may start with the hypothesis $r = 0$ and *increase* the rank as long as the test yields a *significant* test statistic. This procedure is called "bottom-up procedure". The latter procedure is preferred by Johansen (1995, p.167).

Whenever the sequence of LR tests yields "monotonic" outcome in the sense that there is a rank r_J such that the test accepts the null for all $r \geq r_J$ and rejects for $r < r_J$, then the top-down and the bottom-up procedures yield the same result. However, both procedures differ in the treatment of a "non-monotonic" sequence of test decisions. For illustration assume that the sequence of tests in a five-dimensional system yields the following non-monotonic result:

$$\begin{array}{cccccc} r_0 = & 1 & 2 & 3 & 4 & 5 \\ & * & - & * & - & - \end{array}$$

where "$-$" and "$*$" indicate that the null hypothesis is accepted or rejected, respectively. For such a sequence, the top-down procedure would select the rank 4 and the bottom-up procedure would suggest the rank 2.

To assess the probability for a non-monotonic sequence of test decisions, it is useful to consider the (asymptotic) distribution of the test statistic for the case that the the true rank r^* is lower than the rank under test. Usually, when testing a sequence of nested hypotheses, the test statistics are asymptotically stochastically independent (e.g. Holly 1988, Sec. 4), so that we might expect that for $r_0 > r^*$ the test rejects with a probability equal to the size of test. Intuitively, when a subset of hypotheses is tested then this test does not depend on the validity of another subset of hypotheses. Similarly, we may assume that when testing a subset of eigenvalues against zero, the values of the other eigenvalues does not affect the test decision. However, this is not the case. Since the eigenvalues are ordered by their value, the test will depend on the values of the other eigenvalues, in general.

Let $\text{LR}(r_0)$ denote Johansen's LR trace statistic of the hypothesis $r = r_0$. Then under the assumptions of Johansen (1988) for a n-dimensional VAR model

Table 1: Actual sizes for LR tests with $r_0 \geq r^*$

$q = n - r^*$	$n - r_0$				
	1	2	3	4	5
2	0.94	**4.73**	–	–	–
3	0.28	0.39	**4.87**	–	–
4	0.13	0.07	0.45	**4.79**	–
5	0.17	0.02	0.02	0.51	**4.85**
6	0.13	0.01	0.00	0.05	0.42

Note: Entries report the rejection frequencies in percent for Johansen's trace test with a significance level of 0.05 computed from 10.000 replications of random walk sequences with $T = 500$. The bold numbers are the sizes for using the true cointegration rank in the null hypothesis.

with cointegration rank $r_0 > r^*$ we have as $T \to \infty$:

$$\text{LR}(r_0) \;\Rightarrow \sum_{j=1}^{n-r_0} \lambda_j(q) \ ,$$

where $q = n - r^*$ and $0 < \lambda_1(q) < \cdots < \lambda_q(q)$ are the ordered eigenvalues of the stochastic matrix

$$\int dW_q W_q' \left(\int W_q W_q' \right)^{-1} \int W_q dW_q' \ ,$$

and W_q is a q-dimensional standard Brownian motion.

Since all eigenvalues are positive it follows that $LR(0) < LR(1) < \cdots < LR(r^*)$, and, thus, tests with $r_0 > r^*$ are conservative. To get an impression of the size bias we compute the actual sizes for various combinations of $n - r_0$ and $n - r^*$. The results are presented in Table 1. It turns out that tests with $r_0 > r^*$ are highly conservative. If the r_0 exceeds r^* by more than one, then the actual size is very small (< 0.3 percent). This results demonstrates that the probability of detecting a non-monotonic sequence of test decisions is small and, thus, in practice we usually find that both procedures give the same result.

Nevertheless, in situations where the test has a poor power (e.g. in small samples), the procedures may select different ranks more frequently. Therefore, it is interesting to compare the properties of both procedures. It is well known that in a sequence of tests the *overall size* is different from the size of the individual tests. In the case that the tests statistics are uncorrelated it is easy to calculate the overall significance level (see, e.g., Lütkepohl 1991, p. 126). However, in our case the test statistics are correlated and we only can give a quite conservative upper bound for the top-down procedure.

In contrast, for bottom-up procedure the overall type I error is bounded by the size of the individual tests. To see this, assume that the tests are performed for the whole sequence of n hypotheses rather then stopping if the null hypothesis is accepted. Then, for $H_0 : r_0 = r^*$ (the true rank) we will find that the test accepts in $(1 - \alpha^*)100\%$ of the cases, where α^* denotes the size of the individual tests. By construction, for these cases the bottom-up procedure selects a rank $r_J^b \geq r^*$ and, thus, we get

$$P(r_J^b < r^*) \leq \alpha^* \ . \tag{1}$$

Thus, the advantage of the button-up strategy is that we can easily control the overall size of the procedure. A similar result is obtained by Dickey and Pantula (1987) for the determination of the degree of integration of a univariate time series.

Next we show that by using two different bottom-up procedures it is possible to construct a confidence set for the unknown cointegration rank. From (1) it is seen that by using Johansen's LR procedure it is possible to control the probability that the bottom-up procedure selects a lower rank. Assume that we have a different type of test procedure that allows to test the hypotheses

$$H_0 : \quad r = r_0 \quad \text{versus} \quad H_1 : \quad r < r_0 \ .$$

Such a test procedure flips the null and alternative hypotheses of Johansen's LR test. We then can construct a bottom-up procedure by starting with a test of the hypothesis $r = n$. If the hypothesis is rejected, we test the hypothesis $r = n - 1$ and will proceed so until the test accept the hypothesis. We denote the selected rank of such a procedure as r_R^b, where the index R indicates that the test uses a reversed sequence of hypotheses. Although the rank is tested in a descending order, it is essentially a bottom-up strategy because we proceed with testing as long as the test *rejects* the null hypothesis.

As for r_J^b, it is possible to control the overall size such a test sequence so that

$$P(r_R^b > r^*) \leq \alpha^* \ , \tag{2}$$

where again α^* denotes the size of the individual tests. Using (1) and (2) it is possible to construct a $1 - 2\alpha^*$ confidence set for the rank r^*:

$$P(r_J^b < r^* < r_R^b) \leq 1 - 2\alpha^* \ . \tag{3}$$

It should be noticed that this confidence set may be conservative. If the power of the test is unity and the test statistics are perfectly correlated such that both tests always reject $H_0 : r = r^*$ together, then the probability in (3) is $1 - \alpha$. If $r_R = 0$ and $r_J = n$, then the confidence set is uninformative.

The rest of the paper deals with a test based on a CCA between y_t and Y_{t-1} which can be used to obtain r_R^b. Of course, the tests of Harris (1997) and Snell (1998) can be used as well.

3 Testing for Stationarity

Assume that the $n \times 1$ vector y_t is generated by a linear process given by

$$\Delta y_t = \Pi y_{t-1} + u_t \ , \tag{4}$$

where $\{u_t\}$ obeys the following assumption:

Assumption 3.1: *Let* $u_t = A(L)\varepsilon_t = \sum_{j=0}^{\infty} A_j \varepsilon_{t-j}$ *with* $\sum_{j=0}^{\infty} j^2 ||A_j||^2 < \infty$ *and* ε_t *is i.i.d. with* $E(\varepsilon_t) = 0$ *and positive definite covariance matrix* $E(\varepsilon_t \varepsilon_t') = \Sigma_\varepsilon$.

A similar assumption is used in Bewley and Yang (1995) and Quintos (1998). Although it is possible to relax this assumption to allow for some kinds of heteroscedasticity, this assumption is used to facilitate the exposition.

If the rank of the matrix Π is $0 < r < n$, then the factorization $\Pi = \alpha\beta'$ applies, where α and β are $n \times r$ matrices. Furthermore, it is assumed that Δy_t has a Wold representation of the form:

$$\Delta y_t = C\varepsilon_t + C^*(L)\Delta\varepsilon_t \ , \tag{5}$$

where $\beta' C = 0$ and $C^*(L) = C_0^* + C_1^* L + C_2^* L^2 + \cdots$ is a matrix polynomial with all roots outside the unit circle and C is an $n \times n$ matrix with $rk(C) = n - k$. This assumption ensures that Δy_t is stationary.

If u_t is white noise, then Johansen's LR test for the cointegration rank is based on a CCA between Δy_t and y_{t-1} leading to the problem:

$$|\lambda S_{11} - S_{10} S_{00}^{-1} S_{10}'| = 0 \ , \tag{6}$$

where

$$S_{11} = \sum_{t=2}^{T} y_{t-1} y_{t-1}', \quad S_{00} = \sum_{t=2}^{T} \Delta y_t \Delta y_t', \quad S_{10} = \sum_{t=2}^{T} y_{t-1} \Delta y_t' \ .$$

The eigenvalues are equivalent to the eigenvalues of the matrix products $\widehat{\Pi}\widehat{\Pi}^*$ (or $\widehat{\Pi}^*\widehat{\Pi}$), where $\widehat{\Pi}$ is the least-squares estimate from a regression of Δy_t on y_{t-1} and $\widehat{\Pi}^*$ denotes the estimate from a (reverse) regression of y_{t-1} on Δy_t. The LR test statistic is (approximately) the sum of the r smallest eigenvalues. If y_t is cointegrated with rank r, then $n - r$ eigenvalues for (6) tend to zero with the rate T^{-1}.

The hypothesis on the cointegration rank is tested by analyzing the cointegration properties of the partial sum process $Y_t = \sum_{i=1}^{t} y_i$. Under the null hypothesis we assume that the cointegration rank is r, that is, there exists an $n \times r$ matrix β such that $\beta' y_t \sim I(0)$. The eigenvalues from a CCA between y_t and Y_{t-1} result from the problem

$$|\lambda S_{22} - S_{21} S_{11}^{-1} S_{21}'| = 0 \tag{7}$$

$$or \qquad |\lambda S_{11} - S_{21}' S_{22}^{-1} S_{21}| = 0 \ , \tag{8}$$

where

$$S_{11} = \sum_{t=2}^{T} y_t y_t', \quad S_{22} = \sum_{t=2}^{T} Y_{t-1} Y_{t-1}', \quad S_{21} = \sum_{t=2}^{T} Y_{t-1} y_t' \ .$$

As in Johansen (1995, p. 151f) we first consider the limiting distribution of a special case. To test the null hypothesis $r_0 = n$ (y_t is stationary) against the alternative $r_0 < n$ we use the normalized sum of the eigenvalues of problem (8) as the test statistic. The following theorem gives the asymptotic null distribution for this test statistic.

Theorem 3.1: *Let y_t be a vector of stationary time series with positive definite covariance matrix $E(y_t y_t') = \Gamma_0$. The test statistic for testing $H_0 : r = n$ is $\varphi_n = T \sum_{j=1}^{n} \lambda_j$, where λ_j, $j = 1, \ldots, n$ denote the eigenvalues of the problem (8). For $T \to \infty$ the asymptotic null distribution is given by*

$$\varphi_n \Rightarrow tr\left[\left(\int W_n W_n'\right)^{-1} \left(\int W_n dW_n' + \Upsilon\right) \Omega_y^{1/2} \Gamma_0^{-1} \Omega_y^{1/2} \left(\int dW_n W_n' + \Upsilon'\right)\right],$$

where

$$\begin{aligned} \Upsilon &= \Omega_y^{1/2} \Psi \Omega_y^{1/2} \\ \Psi &= \sum_{i=1}^{\infty} \Gamma_i \\ \Gamma_i &= E(y_t y_{t+i}') \end{aligned}$$

and it is assumed that $\Omega^{1/2}$ is a symmetric matrix such that $\Omega^{1/2}\Omega^{1/2} = \Omega$.

This result suggests to correct the test statistic for the nuisance parameters by using the expressions

$$\begin{aligned} \widetilde{S}_{21} &= \left[\sum_{t=2}^{T} Y_{t-1} y_t' - T\widehat{\Psi}'\right] \widehat{\Omega}^{-1/2} \\ \widetilde{S}_{11} &= \sum_{t=2}^{T} \widehat{\Gamma}_0^{-1/2} y_t y_t' \widehat{\Gamma}_0^{-1/2} \end{aligned}$$

instead of S_{21} and S_{11} in (7), where $\widehat{\Psi}, \widehat{\Omega}$ and $\widehat{\Gamma}_0$ are consistent estimates of Ψ, Ω and Γ_0. Following Phillips (1995) the following estimators are used:

$$\widehat{\Psi}(k) = \sum_{i=1}^{k} w(i) \widehat{\Gamma}_i \tag{9}$$

$$\widehat{\Omega}(k) = \widehat{\Gamma}_0 + \widehat{\Psi}(k) + \widehat{\Psi}(k)' \tag{10}$$

$$\widehat{\Gamma}_i = T^{-1} \sum_{t=1}^{T-i} y_t y_{t+i} \ , \tag{11}$$

where $w(i)$ is an appropriate weight function and k denotes the truncation lag, which increases with the sample size such that $k \to \infty$ as $T \to \infty$ but $k/T \to 0$. Further kernel conditions and bandwidth expansion rates are given in Phillips (1995).

A natural estimator for Γ_0 is $\widehat{\Gamma}_0 = T^{-1}\sum y_t y_t'$ so that the term $\widetilde{S}_{11}$ reduces to T yielding a standard eigenvalue problem:

$$|\widetilde{\lambda} I_n - \widehat{S}_{21}(k)' S_{22}^{-1} \widehat{S}_{21}(k)| = 0 \ , \tag{12}$$

where the factor T is absorbed in $\widetilde{\lambda}$. The resulting test statistic is

$$\widetilde{\varphi}_n(k) = \sum_{j=1}^{n} \widetilde{\lambda}_j = tr[\widehat{S}_{21}(k)' S_{22}^{-1} \widehat{S}_{21}(k)]. \tag{13}$$

Using Theorem 3.1 it is easy to verify that this statistic has the same distribution as Johansen's LR trace statistic.

4 The Asymptotic Null Distribution for $r_0 < n$

In this section we consider a test of the null hypothesis $H_0 : r = r_0 < n$ against the alternative $r < r_0$. The special case H_0: $r = 1$ is the situation considered in Leybourne and McCabe (1994a), Shin (1994) and Harris and Inder (1994). Without loss of generality we will consider the transformed system $x_t = Qy_t$, where Q is an invertible $n \times n$ matrix. This transformation is used to separate r stationary linear combinations from the remaining $n - r$ nonstationary components. A further feature of this transformation is that the resulting components are asymptotically independently distributed with unit covariance matrix. Note that such a "rotation" of the system does not affect the eigenvalues for our test procedure. It is merely introduced to facilitate the asymptotic analysis of the system.

Lemma 4.1: *There exists an invertible matrix $Q = [\beta^*, \gamma^*]'$, where β^* is an $n \times r$ cointegration matrix and γ^* is an $n \times (n - r)$ matrix linearly independent of β^* such that*

$$\begin{aligned} x_t = \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} &= Qy_t = \begin{bmatrix} \beta^{*\prime} y_t \\ \gamma^{*\prime} y_t \end{bmatrix} \\ T^{-1/2} \sum_{i=1}^{[aT]} x_{1i} &\Rightarrow W_r(a) \\ T^{-1/2} x_{2,[aT]} &\Rightarrow W_{n-r}(a), \end{aligned}$$

where W_r and W_{n-r} are uncorrelated r and $(n-r)$ dimensional Brownian motions with unit covariance matrix.

Furthermore, to abstract from nuisance parameters we will make the following assumption, which will be relaxed below.

Assumption 4.1: *x_{1t} and Δx_{2t} are white noise with $E(x_{1t}|x_{t-1}, x_{t-2}, \ldots) = 0$ and $E(\Delta x_{2t}|x_{t-1}, x_{t-2}, \ldots) = 0$ for all t.*

For notational convenience we define the matrices $X = [x_2, x_3, \ldots, x_T]'$ and $Z = [X_1, X_2, \ldots, X_{T-1}]'$, where $X_t = \sum_{i=1}^{t} x_i$. Similar as in the case $r_0 = n$ the eigenvalue problem is of the form

$$|\lambda I_n - X'Z(Z'Z)^{-1}Z'X| = 0. \tag{14}$$

Originally, a CCA between x_t and Z_t would require to set $X'X$ instead of I_n. However, as argued in Section 3, the term $X'X$ drops out when the test statistic is corrected for nuisance parameters.

Let b_j denote the eigenvector corresponding to λ_j. If b_j falls inside the cointegration subspace, then λ_j is $O_p(1)$. That is, there exist r eigenvalues with a nondegenerate limiting distribution. On the other hand, if b_j falls outside the cointegration subspace, then the corresponding eigenvalues diverge at the rate[1] T^2.

It is interesting to compare this asymptotic behavior with the properties of the eigenvalues from the ML estimation in a VAR system. In the latter case Johansen (1988) shows that r eigenvalues are $O_p(1)$ and $n-r$ eigenvalues are $O_p(T^{-1})$. Whereas Johansen's test is based on the (normalized) $n-r$ eigenvalues, our test is based on the smallest r eigenvalues. Accordingly, the test flips the null and alternative hypotheses.

In the following theorem the asymptotic null distribution of the test ist given.

[1]Note that by replacing $X'X$ by I_n in (14) the eigenvalues need no longer be smaller than one, as it is the case for the original CCA problem.

Theorem 4.1: *Let y_t be generated as in (4), with cointegration rank $1 \leq r \leq n-1$. Furthermore $\varphi_r = \sum_{j=1}^{r} \lambda_j$, where $\lambda_1 \leq \cdots \leq \lambda_n$ are the eigenvalues of the problem (14). Then, under Assumption 3.1 and $T \to \infty$ we have*

$$\varphi_r \;\Rightarrow\; tr\left[\left(\int dV_r \xi_n'\right)\left(\int \xi_n \xi_n'\right)^{-1}\left(\int \xi_n dV_r'\right)\right] ,$$

where

$$\begin{aligned} dV_r = dW_r \quad &- \quad \left\{\left[\int dW_r \xi_n' \left(\int \xi_n \xi_n'\right)^{-1} \int \xi_n W_{n-r}'\right]\right. \\ &\times \quad \left.\left[\int W_{n-r} \xi_n' \left(\int \xi_n \xi_n'\right)^{-1} \int \xi_n W_{n-r}'\right]^{-1}\right\} W_{n-r}, \end{aligned}$$

$\xi_n = [W_r', \int W_{n-r}']'$, W_r and W_{n-r} are r and $(n-r)$ dimensional standard Brownian motions.

This limiting distribution is more complicated as for the case $r_0 = n$ and depends on the dimensions r and $n-r$. Critical values obtained from this limiting distribution are presented in Appendix B (Table B.1).

In order to allow for a constant or a trend the test can be performed using the mean-adjusted series $\tilde{y}_t = y_t - T^{-1} \sum y_t$ or the trend adjusted series $\hat{y}_t$ that results as the residuals from a regression of y_t on t and a constant. The partial sums are then constructed by using $\tilde{y}_t$ or $\hat{y}_t$. As usual the limiting distribution of the resulting test statistics is different from the case without any deterministics. Although the general form of the asymptotic distribution is the same, the Brownian motions are replaced by multivariate Brownian bridges in case of mean adjusted series and by second order Brownian bridges (cf KPSS 1992) in the case of a trend adjustment. Corresponding critical values for these cases can be found in the Appendix (Table B.2 and Table B.3).

To accommodate more general processes we allow x_{1t} and Δx_{2t} to be serially correlated. As a consequence, the limiting distribution of the test statistic depends on nuisance parameters. Therefore, to adjust the test statistic for nuisance parameters we use the same estimators (9) – (11) as for the case $r_0 = n$ and replace $Z'X$ in (14) by

$$\widehat{S}_{21}^{x}(k) = [Z'X - T\widehat{\Psi}^{x}(k)']\widehat{\Omega}^{x}(k)^{-1/2}$$

where $\widehat{\Psi}^x(k)$ and $\widehat{\Omega}^x(k)$ are computed as in (9) – (11) but with x_t instead of y_t. This may appear inappropriate since for $r < n$ the covariances $\widehat{\Gamma}_j$ are $O_p(T)$ and, thus, the nuisance parameters tend to infinity as $T \to \infty$. Nevertheless, under appropriate assumptions on the asymptotic behavior of the nuisance parameters it is shown that the asymptotic null distribution is not affected by using estimates for the nuisance parameters.

Assumption 4.2: *Let Ψ^x and Ω^x be partitioned according to $x_t = [x_{1t}', x_{2t}']'$ such that*

$$\Psi^x = \begin{bmatrix} \Psi_{11}^x & \Psi_{21}^{x\,\prime} \\ \Psi_{21}^x & \Psi_{22}^x \end{bmatrix} \qquad \text{and } \Omega^x = \begin{bmatrix} \Omega_{11}^x & \Omega_{21}^{x\,\prime} \\ \Omega_{21}^x & \Omega_{22}^x \end{bmatrix} .$$

It is assumed that the estimates of the submatrices of Ψ^x and Ω^x obey the following assumptions:

$$\begin{aligned} \widehat{\Psi}_{11}^x(k) &= \Psi_{11}^x + o_p(1) \\ \widehat{\Omega}_{11}^x(k) &= \Omega_{11}^x + o_p(1) = I_r + o_p(1) \\ \widehat{\Psi}_{22}^x(k) &= O_p(kT) \\ \widehat{\Omega}_{22}^x(k) &= O_p(kT) \\ \widehat{\Psi}_{21}^x(k) &= O_p(k) \\ \widehat{\Omega}_{21}^x(k) &= O_p(k) \end{aligned}$$

The usual kernel estimates such as the ones considered in Phillips (1995) satisfy this assumption.

Theorem 4.2: *Let y_t be generated as in (4), with cointegration rank $1 \le r \le n-1$. Furthermore $\widetilde{\varphi}_r(k) = \sum_{j=1}^{r} \widetilde{\lambda}_j$, where $\widetilde{\lambda}_j$ $(j = 1, \ldots, n)$ denote the eigenvalues of the problem*

$$|\widetilde{\lambda} I_n - \widehat{S}_{21}^x(k)'(Z'Z)^{-1}\widehat{S}_{21}^x(k)| = 0. \tag{15}$$

For $k/T \to \infty$, and under Assumption 3.2, a test based on $\widetilde{\varphi}_r(k)$ has the same limiting distribution as φ_r.

For a similar set of conditions an analogous result is obtained for the KPSS statistic. There are two reasons for this result to hold. First, by rotating the system as in Lemma 4.1, we obtain two sets of nuisance parameters. The estimates of the nuisance parameters involved by the smallest r eigenvalues converge to the true values as $T \to \infty$, whereas the estimates of the nuisance parameters corresponding to the remaining $n - r$ eigenvalues diverge. Since the test statistic only involves the smallest r eigenvalues, the estimated nuisance parameters do not affect the null distribution.

5 Using More Efficient Estimates

In Johansen's ML estimation procedure, the eigenvectors corresponding to the r largest eigenvalues are T-consistent estimates for some suitably normalized cointegration vectors. For a CCA between y_t and Y_{t-1} a similar result can be obtained. The eigenvalue λ_j for the problem (14) can be written as

$$\lambda_j = \frac{b_j' X' Z (Z'Z)^{-1} Z' X b_j}{b_j' b_j} \tag{16}$$

and the corresponding eigenvector b_j can be decomposed as

$$b_j = \beta p_j + \beta_\perp q_j \ ,$$

where p_j and q_j are $r \times 1$ and $(n - r) \times 1$ vectors. In the transformed system, the cointegration matrix is $\beta = [I_r, 0]'$ and the orthogonal complement is given by $\beta_\perp = [0, I_{n-r}]'$. Since the r smallest eigenvalues are $O_p(1)$, it follows that the vector q_j must converge to zero with the rate $O(T^{-1})$ and, thus, the eigenvectors are T-consistent estimates for the respective cointegration vectors βp_j. In the proof of Theorem 4.1 it is shown that by normalizing the matrix of the eigenvectors as $\tilde{\beta}_T = [I_r, -\tilde{\Phi}_T']'$ the submatrix $\tilde{\Phi}_T$ is asymptotically equivalent to an instrumental variable (IV) estimator of Φ in the model

$$x_{1t} = \Phi' x_{2t} + \nu_t \tag{17}$$

with

$$\tilde{\Phi}_T = [X_2' Z (Z'Z)^{-1} Z' X_2]^{-1} X_2' Z (Z'Z)^{-1} Z' X_1 + O_p(T^{-3}).$$

Recall that in the rotated system $x_{1t} = \beta^{*\prime} y_t$ is stationary and $x_{2t} = \gamma^{*\prime} y_t$ is nonstationary so that $\widetilde{\Phi}_T$ converges to zero as $T \to \infty$.

A useful instrument w_t for estimating (17) should obey two conditions

$$\begin{aligned} T^{-\delta} \sum_t w_t \nu_t' &\Rightarrow 0 \\ T^{-\delta} \sum_t w_t x_{2t}' &\Rightarrow A \neq 0 \end{aligned}$$

for some $\delta > 0$. It is easy to verify that X_{1t} satisfies these conditions for $\delta = 2$ and X_{2t} satisfies these conditions for $\delta = 3$. However, in addition x_{2t} is a useful instrument implying, which can be seen by setting $\delta = 2$. Hence, the IV estimator can be improved by adding x_{2t} to the set of instruments. This can be done by considering the eigenvalues of the problem:

$$|\lambda^* I_n - X' Z^* (Z^{*\prime} Z^*)^{-1} Z^{*\prime} X| = 0 \ , \tag{18}$$

where $z_t^* = [X_{1t}', x_{2t}', X_{2t}']'$ and $Z^* = [z_1^*, \ldots, z_{T-1}^*]'$.

For estimating the nuisance parameters, the covariance matrices are computed as

$$\Gamma_i^* = T^{-1} \sum_{t=1}^{T-i} y_t^* y_{t+i}^{*\prime} \ ,$$

where $y_t^* = [y_t', \Delta x_{2t}']'$. The differences of x_{2t} are used because this term is known to be $I(1)$ under both the null and under the alternative. If it is unknown how to construct x_{2t}, one may use the $n - r$ eigenvectors corresponding to the zero eigenvalues of Johansen's estimation procedure to construct an estimated version of the nonstationary components. It is easy to verify that the asymptotic distribution is not affected by using consistent estimates of the nonstationary components. The asymptotic null distribution of the test statistic is given the following Theorem.

Theorem 5.1: *Let y_t be generated as in (4), where $1 \leq r \leq n-1$ and $\{u_t\}$ obeys Assumption 3.1. Furthermore $\varphi_r^a = \sum_{j=1}^r \lambda_j^*$, where $\lambda_1^* \leq \cdots \leq \lambda_n^*$ are the eigenvalues of the problem (18). Then, as T $\to \infty$:*

$$\varphi_r^{iv} \ \Rightarrow \ tr \left[\left(\int dV_r \xi_n^{*\prime} \right) \left(\int \xi_n^* \xi_n^{*\prime} \right)^{-1} \left(\int \xi_n^* dV_r' \right) \right]$$

where

$$dV_r = dW_r \quad - \quad \left\{ \left[\int dW_r \xi^*_{2n-r}{}' \left(\int \xi^*_{2n-r} \xi^*_{2n-r}{}' \right)^{-1} \int \xi^*_{2n-r} W'_{n-r} \right] \right.$$
$$\times \quad \left. \left[\int W_{n-r} \xi^*_{2n-r}{}' \left(\int \xi^*_{2n-r} \xi^*_{2n-r}{}' \right)^{-1} \int \xi^*_{2n-r} W'_{n-r} \right]^{-1} \right\} W_{n-r},$$

$\xi^*_{2n-r} = [W'_r, W'_{n-r}, \int W'_{n-r}]'$, W_r *and* W_{n-r} *are* r *and* $(n-r)$ *dimensional standard Brownian motions.*

Critical values resulting from this limiting distribution are presented in Appendix B.

Another possibility is to use the efficient "Fully-modified" estimator of Phillips and Hansen (1990) or the projection estimator of Saikkonen (1991) as in Shin (1994). Assume that the time series vector can be partitioned as $y_t = [y'_{1t}, y'_{2t}]'$ where y_{2t} is assumed to be strongly exogenous. Furthermore we assume that the cointegration matrix can be normalized as $\beta = [I, -\Psi']$. In this case an efficient estimate of the cointegration matrix can be obtained from a regression of y_{1t} on y_{2t}. A test statistic corresponding to the sum of the r smallest eigenvalues is obtained as

$$\varphi^e_r = tr\left[\widehat{\beta}' y' Y (Y'Y)^{-1} Y' y \widehat{\beta}\right],$$

where $\widehat{\beta} = [I_r, -\widehat{\Psi}'_e]'$ and $\widehat{\Psi}_e$ is an asymptotically efficient estimator for the cointegration regression $y_{1t} = \Psi' y_{2t} + u_t$. As in Shin (1994) the regression includes leads and lags of Δy_{2t} if y_{2t} is endogenous. Alternatively, the "fully-modified" system estimator of Phillips (1995) may be used (see Harris and Inder 1994). The following theorem gives the asymptotic null distribution of the resulting test statistic.

Theorem 5.2: *Let* y_t *be generated as in (4), where* $1 \leq r \leq n-1$ *and* $\{u_t\}$ *obeys Assumption 3.1. Let* $\widehat{\beta} = [I_r, -\widehat{\Phi}'_e]'$ *and* $\widehat{\Phi}_e$ *is an asymptotically efficient estimator of the cointegration matrix normalized as* $\beta = [I_r, -\Phi']'$. *Then, as* $T \to \infty$*:*

$$\varphi^e_r \;\Rightarrow\; tr\left[\left(\int dV_r \xi'_n\right)\left(\int \xi_n \xi'_n\right)^{-1}\left(\int \xi_n dV'_r\right)\right],$$

where

$$dV_r = dW_r \quad - \quad \left[\int dW_r W_{n-r}' \left(\int W_{n-r} W_{n-r}' \right)^{-1} \right] W_{n-r} \ ,$$

and W_r *and* W_{n-r} *are* r *and* $(n-r)$ *dimensional standard Brownian motions.*

The attractive feature of this approach is that such a test uses an efficient estimate for the cointegration matrix. However, in practice it is not clear whether the chosen normalization is valid. In particular for large dimensions r, there is a serious danger that the normalization fails which may have serious effects on the distribution of the test statistic. Therefore, the CCA approach or a test based on principal components (Harris 1997, Snell 1998) is favorable in practice.

6 A Small Sample Refinement

From KPSS type of tests it is known that the correction for nuisance parameters reduce the power of the test considerably (e.g. KPSS 1992, Leybourne and McCabe 1994b). Although the local power of the test is unaffected, the power in finite samples depends crucially on the truncation lag of the estimates (cf Breitung 1995). Leybourne and McCabe (1994b) therefore suggest to adopt a parametric model to correct for nuisance parameters. However, such an approach requires to estimate an ARMA model with r MA unit roots by exact maximum likelihood which would be fairly complicated task in a multivariate framework. We therefore adopt a simpler approach suggested in Breitung (1995).

The principle is easily explained in a univariate context. Assume that a univariate time series y_t (without deterministics) is tested for stationarity by using the test suggested by KPSS (1992). Let Y_t denote the partial sum of y_t and $\tau_T = T^{-2} \sum Y_t^2 / \bar{\sigma}_y^2$ is the KPSS statistic, where $\bar{\sigma}_y^2$ is the estimated "long run variance" of y_t.

Now, consider the autoregression

$$y_t = \phi y_{t-1} + v_t \ .$$

If y_t is $I(1)$, then the OLS estimator of ϕ converges to one at rate T and the residuals are approximately the difference of y_t. The next step is to form the partial sum $V_t = \sum_{j=2}^{t} v_j$ and run the regression

$$y_t = \gamma V_{t-1} + e_t. \tag{19}$$

If y_t is stationary, then the OLS estimator of γ should be close to zero, because the partial sum V_{t-1} cannot explain a stationary variable. In contrast, if y_t is $I(1)$, then $V_{t-1} = y_{t-1}$ and we therefore expect that $\hat{\gamma}$ is close to one. Accordingly, for the residuals of (19) we have $\hat{e}_t \approx y_t$ for a stationary series and $\hat{e}_t \approx \Delta y_t$ if y_t is $I(1)$. This reasoning suggest that the residuals of (19) behave like a stationary series no matter whether y_t is $I(0)$ or $I(1)$.

Unfortunately, this reasoning is only valid if v_t is observable. If v_t is replaced the residual and $\widehat{V}_t = \sum_{i=2}^{t} \hat{v}_i$ is used instead of V_t, the estimate of γ does not converge to one under the alternative (cf Breitung 1995). Nevertheless, under the null hypothesis that y_t is $I(0)$, it can be shown that the estimate of γ indeed converge to zero at a sufficient rate, so that estimating the nuisance parameter using the residuals $\hat{e}_t$ instead of y_t does not affect the limiting distribution of the test.

Notwithstanding the asymptotic failure under the alternative hypothesis, it is reasonable to expect that our intuitive reasoning is helpful in small samples. Since the regression minimize the variances of the residuals, the regression will render a residual series that resembles a stationary series as much as possible and, thus, produces a correction term which is usually smaller than the one computed from the original series. Thus, the loss in power is usually smaller by using $\hat{e}_t$ instead of y_t when estimating the nuisance parameters.

This approach can be straightforwardly adopted to the multivariate case. For convenience we will consider the rotated system x_t. Since the CCA is invariant with respect to such transformations this does not imply any loss of generality. The first auxiliary regression is

$$\Delta x_t = \Pi x_{t-1} + v_t \ . \tag{20}$$

The second auxiliary regression is of the type

$$x_t = \Xi \widehat{V}_{t-1} + e_t \ , \tag{21}$$

where $\widehat{V}_t$ is the multivariate partial sum given by $\widehat{V}_t = \sum_{i=2}^{t} \hat{v}_t$ and $\hat{v}_t$ denotes the residual from (20). Following Breitung (1995) it is straightforward to show that using the residuals of (20) instead of x_t for estimating the nuisance parameters does not affect the asymptotic null distribution.

Theorem 6.1: *Let $\hat{e}_t = x_t - \widehat{\Xi}\widehat{V}_{t-1}$ denote the residuals of (21), where $\widehat{\Xi}$ is the least-squares estimator of Ξ. If Ψ^x and Ω^x are estimated as in Assumption 3.2 but using $\hat{e}_t$ instead of x_t, then the the resulting test statistic has the same asymptotic distribution as $\tilde{\varphi}_r(k)$ in Theorem 4.2.*

Although the modification does not affect the asymptotic size of the test, it may have an important fact on the power of the test. Assume that we estimate Ξ in (21) by using

$$\begin{aligned} V_{t-1} &= \sum_{t=2}^{t-1} \begin{bmatrix} \Delta x_{1t} \\ \Delta x_{2t} \end{bmatrix} - \begin{bmatrix} \Pi_{11} & 0 \\ \Pi_{21} & 0 \end{bmatrix} \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} \\ &= \begin{bmatrix} x_{1t} - \Pi_{11} X_{1t} \\ x_{2t} - \Pi_{21} X_{1t} \end{bmatrix} \end{aligned}$$

instead of $\widehat{V}_{t-1}$. It is not difficult to see that in this case the least-squares estimator of Ξ converges to the matrix

$$\Xi^* = \begin{bmatrix} 0 & 0 \\ \Pi_{21}\Pi_{11}^{-1} & I_{n-r} \end{bmatrix}$$

and, thus, we have in the limit

$$e_t^* = y_t - \Xi^* V_{t-1} = \begin{bmatrix} x_{1t} \\ \Delta x_{2t} + \eta_t \end{bmatrix}$$

where $\eta_t = \Pi_{21}\Pi_{11}^{-1} x_{1t}$ is stationary. Obviously, e_t^* has the desired properties for estimating the nuisance parameters Ψ^x and Ω^x because the resulting estimates converge in probability to a fixed limit as $T \to \infty$. Unfortunately, this reasoning is no longer valid if V_{t-1} is replaced by $\widehat{V}_{t-1}$. Nevertheless, we may hope that $\widehat{V}_{t-1}$ resembles V_{t-1} so that the power of the test may be improved substantially when using $\hat{e}_t$ instead of y_t.

7 Simulation Results

To compare the properties of the new tests with the test suggested by Shin we consider a bivariate model given by the two equations

$$y_{1t} = y_{1,t-1} + \varepsilon_t \quad (22)$$

$$\Delta y_{2t} = \gamma \Delta y_{1t} + v_t - \phi v_{t-1} \ , \quad (23)$$

where ε_t and v_t are mutually uncorrelated white noise with unit variance. If $\phi = 1$, the difference operator drops out and (23) defines the cointegrating relationship $y_{2t} - \gamma y_{1t} = v_t$. On the other hand, an integration of equation (23) shows that there is no cointegration between y_{1t} and y_{2t} for $|\gamma| < 1$. Besides ϕ, the power of the test depends on parameter γ, so we present results for different values of ϕ and γ.

First, we use the test statistic suggested in Section 4 to test the hypothesis $r = 1$ ($\phi = 1$) against $r = 0$ ($|\phi| < 1$). Two different truncation lags $k = 4$ and $k = 8$ are used. The corresponding test statistics are indicated by CCA(k). The respective test statistics using a the modified estimates of the nuisance parameters suggested in Section 6 is labeled as CCA*(k).

The CCA statistic using the augmented set of instrumental variables are indicated by CCA$_a$. Two versions of this test statistic are computed. First, y_{1t} is used as additional instrument. By construction, this variable is $I(1)$ and therefore is a valid instrument for estimating Φ in (17). The respective statistic with the modified estimator of the nuisance parameters (see Section 6) is labeled as CCA$_a^*$(k). Second, the nonstationary linear combination is estimated using the eigenvectors corresponding to the nonstationary eigenvalues of Johansen's ML estimation procedure. The respective test statistic is denoted by CCA$_{\hat{a}}^*$(k).

For the test problem considered here, the test suggested by Shin can be applied and will be used as a benchmark for testing the power of the new statistics. The test is based on Saikkonen's (1991) approach, estimating the equation

$$y_{2t} = \gamma y_{1t} + \sum_{j=-m}^{m} \Delta y_{1,t+j} + \nu_t \ , \quad (24)$$

where $m = 2$ is used in our simulations. To estimate the long-run variance of ν_t a Bartlett kernel with truncation lag $k = 4$ is used. The respective test statistic is denoted by Shin$(2, 4)$. Note that for $r = 1$ this test is asymptotically equivalent to Harris' (1997) test and, thus, we expect that our results apply to the latter test as well. Table 2 reports the rejection frequencies computed from 10.000 samples generated from the model (22) – (23) with sample size $T = 200$. The following conclusions can be drawn from the simulation results. The tests using y_t for computing the nuisance parameters tend to be conservative. On the other hand, if $\hat{e}_t$ is used to compute the nuisance parameters as suggested in Section 6, the actual size is much closer to the nominal one, although now the test tend to be slightly liberal.

The original CCA statistic is less powerful than Shin's test although the modification for estimating the nuisance parameters suggested in Section 6 improves the power substantially. For $k = 8$ there is a considerable loss in power compared to a truncation lag of $k = 4$. A similar finding was reported for the KPSS statistic by KPSS (1992) and Breitung (1995). The inclusion of the nonstationary linear combination y_{1t} leads to a substantial gain in power and the resulting test has roughly the same power as Shin's test. For γ close to one, Shin's test is slightly more powerful, whereas for γ close to zero, $\text{CCA}_a^*(k)$ and $\text{CCA}_{\hat{a}}^*(k)$ perform slightly better.

Next we investigate the impact of γ on the size of the test. For the Shin test we assume that the model is (inappropriately) specified as

$$y_{1t} = (1/\gamma)y_{2t} + \sum_{j=-m}^{m} \Delta y_{2,t+j} + \nu_t^* \ .$$

In this formulation of the model y_{2t} is correlated with ν_t^* and for $\gamma = 0$ (i.e. y_{2t} is stationary) Shin's test is invalid because a nonstationary variable (y_{1t}) is regressed on a stationary variable (y_{2t}) and there is no value of γ rendering a stationary error process. From the simulation results presented in Table 3 it is seen that Shin's test is seriously biased when γ is close to zero. These findings clearly demonstrate the problems with Shin's test if the normalization

Table 2: Rejection frequencies for different values of ϕ ($\gamma = 1$)

Test statistic	1	0.95	0.9	0.8	0.5	0
CCA(4)	0.033	0.176	0.308	0.387	0.418	0.421
$\mathrm{CCA}^*(4)$	0.059	0.249	0.397	0.480	0.524	0.575
CCA(8)	0.018	0.091	0.170	0.223	0.242	0.243
$\mathrm{CCA}^*(8)$	0.060	0.211	0.315	0.378	0.411	0.466
$\mathrm{CCA}^*_a(4)$	0.046	0.401	0.620	0.742	0.797	0.865
$\mathrm{CCA}^*_{\hat{a}}(4)$	0.045	0.397	0.613	0.725	0.781	0.833
Shin(2,4)	0.049	0.478	0.619	0.705	0.741	0.744

Note: Entries report the rejection frequencies computed from 10.000 replications of model (22) – (23) with sample size $T = 200$.

Table 3: Rejection frequencies for different values of γ ($\phi = 1$)

Test statistic	1.00	0.50	0.10	0.05	0.01	0.00
$\mathrm{CCA}^*(4)$	0.059	0.059	0.059	0.059	0.059	0.059
$\mathrm{CCA}^*_a(4)$	0.042	0.034	0.015	0.014	0.013	0.013
$\mathrm{CCA}^*_{\hat{a}}(4)$	0.045	0.045	0.046	0.046	0.046	0.046
Shin(2,4)	0.073	0.118	0.562	0.781	0.899	0.902

Note: see Table 2.

of the cointegration vector is invalid.[2] On the other hand, the CCA statistics perform well in this situation. For $\gamma = 0$, the statistic $\mathrm{CCA}^*_a(4)$ is also based on a wrong normalization as it uses y_{2t} as additional instrument. However, it can be shown that in this case the $\mathrm{CCA}^*_a(k)$ statistic has the same asymptotic distribution as the original $\mathrm{CCA}^*(k)$ statistic. Since the critical values of the latter statistic are lower than those of $\mathrm{CCA}^*_a(k)$, the test using $\mathrm{CCA}^*_a(4)$ with an invalid normalization tends to be conservative. In contrast, the actual size of the test using the estimated nonstationary linear combination from the Johansen procedure ($\mathrm{CCA}^*_{\hat{a}}(4)$) is close to the nominal size of 0.05.

In the final Monte Carlo experiment we investigate the potential of the new tests to select the cointegration rank. A four-dimensional cointegrated VAR(1)

[2]See Boswijk (1996) and Saikkonen (1996) for a further discussion of the problems related to the normalization of the cointegration vectors.

Table 4: Rank Selection with Alternative Test Statistics

Test Statistic	$r = 0$	$r = 1$	$r = 2$	$r = 3$	$r = 4$
			$\alpha_1 = 0.4$		
LR	0.000	0.000	0.956	0.042	0.002
$\mathrm{CCA}^*(8)$	0.016	0.013	0.830	0.141	0.000
$\mathrm{CCA}^*_a(8)$	0.000	0.048	0.903	0.049	0.000
$\mathrm{CCA}^*_{\hat{a}}(8)$	0.000	0.023	0.909	0.068	0.000
			$\alpha_1 = 0.3$		
LR	0.000	0.000	0.948	0.046	0.006
$\mathrm{CCA}^*(8)$	0.016	0.012	0.835	0.136	0.001
$\mathrm{CCA}^*_a(8)$	0.000	0.060	0.898	0.041	0.001
$\mathrm{CCA}^*_{\hat{a}}(8)$	0.000	0.026	0.918	0.055	0.001
			$\alpha_1 = 0.2$		
LR	0.000	0.000	0.948	0.048	0.004
$\mathrm{CCA}^*(8)$	0.015	0.019	0.802	0.163	0.001
$\mathrm{CCA}^*_a(8)$	0.000	0.056	0.883	0.060	0.001
$\mathrm{CCA}^*_{\hat{a}}(8)$	0.002	0.039	0.888	0.070	0.001
			$\alpha_1 = 0.1$		
LR	0.000	0.000	0.935	0.062	0.003
$\mathrm{CCA}^*(8)$	0.026	0.046	0.793	0.134	0.001
$\mathrm{CCA}^*_a(8)$	0.003	0.191	0.770	0.035	0.001
$\mathrm{CCA}^*_{\hat{a}}(8)$	0.004	0.180	0.759	0.056	0.001

Note: The entries report the relative frequencies of selecting the indicated rank computed from 1.000 replication of a four-dimensional VAR(1) system with sample size T=200.

model is used, where the matrices β and α are specified as follows:

$$\beta = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad \alpha = \begin{bmatrix} -\alpha_1 & 0 \\ 0 & -0.5 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

that is, the first error correction term enters the first equation with the coefficient $-\alpha_1$ and the second error correction term enters the second equation with the coefficient -0.5. The innovations of the model are mutually uncorrelated Gaussian white noise with unit variances. If α_1 approaches zero, the test procedures will have difficulties to decide whether the cointegration rank is $r = 1$ or $r = 2$.

To assess the ability of the new tests to determine the cointegration rank, a "bottom-up" strategy is used (see Section 2). For the tests using a reversed sequence of hypotheses the truncation lag is $k = 8$, although for higher values of α_1 (such as $\alpha_1 = 0.4$) a smaller truncation is sufficient and leads to a more powerful test procedure. For Johansen's LR test, we assume that the VAR order is known to be one. Of course, assuming the correct parametric model to be known, whereas the other tests adopt a semiparametric approach to estimate the nuisance parameters favors the parametric LR procedure of Johansen. However, it is not the intention here to investigate which procedure performs better. Rather we use the LR test as a benchmark against which we are able to assess the potential of the new tests for the selection of the cointegration rank.

Table 4 reports the observed relative frequencies of selecting a particular rank $r \in \{0, 1, \ldots, 4\}$. The frequencies are based on 1000 simulated samples. From the results it turns out that for small values of α_1, Johansen's LR test performs better than the CCA statistics. For substantial values of α_1, however, the relative performance of the CCA statistics is similar to the LR test procedure.

8 An Empirical Application

In this section the application of the new tests is illustrated by using a dataset of interest yields with different time to maturity. The expectation hypothesis of the term structure implies that interest yields with different time to maturity are mutually cointegrated. Therefore, we expect three cointegration relationships between four k-month interest rates, where $k = 1, 3, 6, 12$. The data are monthly observations running from 1982(1) to 1996(12) and were taken from the database of the German *Bundesbank*.

For the Johansen procedure a VAR(12) model is used and for the CCA statistics a truncation lag of 16 is applied. Whereas the results from the Johansen test do not change very much for different lag orders, the CCA statistics are quite sensitive to the choice of the truncation lag. Since an underspecification of the truncation lag may produce a considerable size bias, a fairly large value of k is chosen. Furthermore, we allow for a constant mean in the data. The results for the cointegration rank tests are presented in Table 5.

The sequence of LR tests suggests that $r = 2$. However, it may be that the

Table 5: Cointegration Rank Statistics for Interest Yields

	LR	CCA*(16)	$\text{CCA}^*_a(16)$	$\text{CCA}^*_{\hat{a}}(16)$
$r = 0$	52.961*	n.a.	n.a.	n.a.
$r = 1$	33.195*	0.464	9.652	9.639
$r = 2$	17.350	9.502	19.987	21.522
$r = 3$	2.879	28.206	38.035	38.849
$r = 4$	n.a.	92.587**	n.a.	n.a

Note: "LR" denotes Johansen's LR trace statistic from a VAR(12) model with constant term. * and ** indicate significance at the levels 0.05 and 0.01, respectively.

power of the test is not sufficient to reject $r = 2$ and $r = 3$, so that the rank may well be three or four. Therefore, it is useful to apply the reverse sequence of tests.

Applying the CCA statistics, the picture is quite clear. Since the hypothesis $r = 4$ is rejected, while the hypotheses $r \leq 3$ are accepted by all versions of the CCA statistic, both tests together suggest that the rank is either two or three.

9 Conclusions

In this paper a CCA approach is adopted to test for cointegration using a reverse sequence of hypotheses. Together with Johansen's LR tests, such a test may give useful additional and allows the construction of a confidence set for the cointegration rank.

As for univariate time series, it is shown that similar principles can be adopted for testing the opposite hypotheses. However several differences remain. First, it is difficult to adopt a parametric framework like the VAR model for Johansen's tests. We therefore make use of nonparametric corrections for nuisance parameters in the tradition of Phillips (1987). Second, whereas the power of Johansen's LR test does not seem to depend sensitively on the correction for short run dynamics, the power of the CCA statistics (similar as the KPSS statistic) is highly sensitive to the choice of the truncation lag. This is an undesirable property of the tests because the power of the tests can be made arbitrarily small by choosing a truncation lag sufficiently high. Third, the asymptotic theory is a little bit more complicated and the critical values have to be tabulated for both dimensions r and n.

Since there does not appear to exist a theoretical reason for choosing among the principal components and the CCA approach it might be interesting to compare the performance of the different tests in an extensive Monte Carlo study as was done by Gonzalo (1994), Haug (1997) and Hubrich et al. (1998) for the test of $r = r_0$ against $r > r_0$.

Appendix A

Proof of Theorem 3.1:

From standard asymptotic results we get:

$$\begin{aligned} T^{-1}\sum Y_{t-1}y_t' &\Rightarrow \int \Omega^{1/2}W_n dW_n'\Omega^{1/2} + \Psi \\ T^{-2}\sum Y_{t-1}Y_{t-1}' &\Rightarrow \int \Omega^{1/2}W_n W_n'\Omega^{1/2} \\ T^{-1}\sum y_t y_t' &\Rightarrow \Gamma_0 \ , \end{aligned}$$

where W_n is an n-dimensional Brownian motion. Using these results, the normalized sum of the eigenvalues results as

$$T\sum_{j=1}^{n}\lambda_j \Rightarrow tr\left[\left(\int W_n W_n'\right)^{-1}\left(\int W_n dW_n' + \Upsilon\right)\Omega_y^{1/2}\Gamma_0^{-1}\Omega_y^{1/2}\left(\int dW_n W_n' + \Upsilon'\right)\right].$$

Proof of Lemma 4.1

From (5) we have for every cointegration matrix β and a linearly independent matrix γ:

$$\begin{aligned} \beta' y_t &= \beta' C^*(L)\varepsilon_t \\ &= \beta' C^*(1)\varepsilon_t + \beta' C^{**}(L)\Delta\varepsilon_t \\ \gamma' y_t &= \gamma' C\sum_{i=1}^{t}\varepsilon_i + C^*(L)\varepsilon_t \ , \end{aligned}$$

where $C^{**}(L) = [C^*(L) - C^*(1)](1-L)^{-1}$ has all roots outside the complex unit circle. Let R be a block diagonal matrix such that

$$R = \begin{bmatrix} R_{11} & 0 \\ R_{21} & R_{22} \end{bmatrix}$$

and

$$\begin{bmatrix} \beta' C^*(1) \\ \gamma' C \end{bmatrix} \Sigma \begin{bmatrix} \beta' C^*(1) \\ \gamma' C \end{bmatrix}' = RR'.$$

Then, by using

$$Q = R^{-1}\begin{bmatrix} \beta' \\ \gamma' \end{bmatrix} = \begin{bmatrix} \beta^{*\prime} \\ \gamma^{*\prime} \end{bmatrix}$$

it readily follows that $T^{-1/2}\sum_{i=1}^{[aT]} x_{1i}$ and $T^{-1/2}x_{2,[aT]}$ converge weakly to the standard Brownian motions W_r and W_{n-r}, respectively.

Proof of Theorem 4.1:

It is convenient to normalize the matrix of the first r eigenvectors as $B_1 = [b_1, \ldots, b_r] = [I_r, -\Phi_T']'$. Consider the j'th eigenvector ($j = 1, \ldots, r$), which is determined by the equation

$$[\lambda_j I_n - X'Z(Z'Z)^{-1}Z'X]b_j = 0.$$

The lower $n - r$ equations of this system can be written as

$$\lambda_j \phi_j - X_2'Z(Z'Z)^{-1}Z'(X_{(j)} - X_2\phi_j) = 0$$

where ϕ_j is the j'th column of Φ_T, $X_{(j)}$ is the j'th column of X and $X_2 = [x_{21}, \ldots, x_{2T}]'$. Since ϕ_j is $O_p(T^{-1})$, $\lambda_j = O_p(1)$ it follows that

$$X_2'Z(Z'Z)^{-1}Z'(X_{(j)} - X_2\phi_j) = O_p(T^{-1}).$$

Using $X_2'Z = O_p(T^3)$ and $Z'Z = O_p(T^4)$ we get

$$\phi_j = [X_2'Z(Z'Z)^{-1}Z'X_2]^{-1}X_2'Z(Z'Z)^{-1}Z'X_{(j)} + O_p(T^{-3}) \qquad (25)$$

or, by collecting the results for $\phi_1, \ldots, \phi_r$:

$$\Phi_T = [X_2'Z(Z'Z)^{-1}Z'X_2]^{-1}X_2'Z(Z'Z)^{-1}Z'X_1 + O_p(T^{-3}),$$

where X_1 is a submatrix with the first r columns of X.

Finally, let $D_T =$diag$[I_r, T^{-1}I_{n-r}]$. Then, using standard asymptotic results for unit root processes:

$$\begin{aligned} T^{-2}X_2'ZD_T &\Rightarrow \int W_{n-r}\xi_n' \\ T^{-2}D_TZ'ZD_T &\Rightarrow \int \xi_n\xi_n' \\ T^{-2}D_TZ'X_1 &\Rightarrow \int \xi_n dW_r' \end{aligned}$$

the limiting distribution follows immediately.

Proof of Theorem 4.2:

The eigenvalue problem is equivalent to

$$\left|\widetilde{\lambda}\widehat{\Omega}^x(k) - [X'Z - T\widehat{\Psi}^x(k)](Z'Z)^{-1}[Z'X - T\widehat{\Psi}^x(k)']\right| = 0 \qquad (26)$$

yielding

$$\widetilde{\lambda}_j = \frac{\tilde{b}_j'[X'Z - T\widehat{\Psi}^x(k)](Z'Z)^{-1}[Z'X - T\widehat{\Psi}^x(k)']\tilde{b}_j}{\tilde{b}_j'\widehat{\Omega}^x(k)\tilde{b}_j}$$

To show that the eigenvalues only depend on the nuisance parameters Ψ_{11} we consider the numerator and denominator separatly.

Let $D_T = \text{diag}[I_r, T^{-1}I_{n-r}]$. Then, using

$$T^{-2}\sum_{t=2}^{T} x_{1t}\left(\sum_{i=1}^{t-1} x_{2i}'\right) = -T^{-2}\sum_{t=1}^{T}\left(\sum_{i=1}^{t} x_{1i}\right)x_{2t}' + o_p(1) \;\Rightarrow\; -\int W_r W_{n-r}'$$

we have as $k/T \to 0$

$$\begin{aligned} T^{-1}D_T X'ZD_T - D_T\Psi^x(k)D_T &= T^{-1}D_T X'ZD_T - \begin{bmatrix} \Psi_{11} + o_p(1) & O_p(k/T) \\ O_p(k/T) & O_p(k/T) \end{bmatrix} \qquad (27) \\ &\Rightarrow \begin{bmatrix} \int dW_r W_r' - \Psi_{11} & -\int W_r W_{n-r}' \\ -\int W_{n-r}W_r' & \int (W_{n-r}\int W_{n-r}') \end{bmatrix}. \end{aligned}$$

Furthermore, we have

$$T^{-2}D_T Z'ZD_T \;\Rightarrow\; \begin{bmatrix} \int W_r W_r' & \int (W_r \int W_{n-r}' \\ \int(\int W_{n-r})W_r' & \int(\int W_{n-r})(\int W_{n-r}') \end{bmatrix}.$$

Next we show that the eigenvector $\tilde{b}_j$ $(j = 1, \ldots, r)$ does not depend on nuisance parameters, asymptotically. From Assumption 4.2 it follows that the lower $n - r$ equations of eigenvalue problem (26) we have

$$X_2'Z(Z'Z)^{-1}Z'X\tilde{b}_j = O_p(k)$$

and, thus,

$$\tilde{\phi}_j = [X_2'Z(Z'Z)^{-1}Z'X_2]^{-1}X_2'Z(Z'Z)^{-1}Z'X_{(j)} + O_p(k/T^2)$$

where $\tilde{\phi}_j$ is defined as the lower $n - r$ subvector of b_j. Since all variables in X_2 and Z are $I(1)$ or $I(2)$ variables, the asymptotic distribution of $\tilde{\phi}_j$ only depends on their long run covariance matrices, which are unity by construction. Thus, for $k/T \to 0$ the asymptotic distribution of $\tilde{\phi}_j$ is the same as the asymptotic distribution of ϕ_j in (25).

It remains to show that $b_j'\widehat{\Omega}^x(k)b_j$ does not depend on nuisance parameters. Since $\tilde{\phi}_j = O_p(T^{-1})$ it follows from Assumption 4.2 that

$$b_j'\widehat{\Omega}^x(k)b_j = 1 + O_p(k/T)$$

and thus, for $k/T \to 0$ all terms that enter the expression for $\tilde{\lambda}_j$ are free from nuisance parameters and yield the same asymptotic distribution as the in Theorem 4.1.

Proof of Theorem 5.1

Let $D_T^* = [I_n, T^{-1}I_{n-r}]$. It follows that

$$T^{-1/2}D_T^* z_{[aT]}^* = \begin{bmatrix} T^{-1/2}X_{1,[aT]} \\ T^{-1/2}x_{2,[aT]} \\ T^{-3/2}X_{2,[aT]} \end{bmatrix} \Rightarrow \xi^*$$

and, it is straightforward to show that the limiting distribution results from replacing ξ by ξ^* in Theorem 4.1.

Proof of Theorem 5.2

Following Saikkonen (1991) we find that under appropriate conditions on m the OLS estimator of Φ in the cointegration regression

$$x_{1t} = \Phi' x_{2t} + \sum_{j=-m}^{m} \Delta x_{2,t-j} + \nu_t$$

is asymptotically distributed as

$$T\widehat{\Phi} \Rightarrow \left[\int W_{n-r}W_{n-r}'\right]^{-1} \int W_{n-r} dW_r'$$

Furthermore, using $T^{-1/2}D_T X_{[aT]} \Rightarrow \xi$ the limiting distribution stated in the theorem is easly derived.

Proof of Theorem 6.1

It is not difficult to verify that for the least-squares estimate of Π in (20) we have

$$\widehat{\Pi} = \begin{bmatrix} \Pi_{11} + O_p(T^{-1/2}) & \eta_{1T} \\ \Pi_{21} + O_p(T^{-1/2}) & \eta_{2T} \end{bmatrix}$$

where η_{1T} and η_{2T} are $O_p(T^{-1})$. Accordingly, we get for the partial sum

$$\widehat{V}_t = \begin{bmatrix} \Pi_{11} X_{1,t-1} + \eta_{1T} X_{2,t-1} \\ \Pi_{21} X_{1,t-1} + \eta_{2T} X_{2,t-1} \end{bmatrix} + O_p(1) \ .$$

This gives

$$\sum_{t=2}^{T-1} \widehat{V}_t \widehat{V}_t' = \begin{bmatrix} O_p(T^2) & O_p(T^2) \\ O_p(T^2) & O_p(T^2) \end{bmatrix} \quad \text{and} \quad \sum_{t=2}^{T-1} x_t \widehat{V}_{t-1}' = \begin{bmatrix} O_p(T) & O_p(T) \\ O_p(T^2) & O_p(T^2) \end{bmatrix}$$

and, thus,

$$\widehat{\Xi} = \begin{bmatrix} O_p(T^{-1}) & O_p(T^{-1}) \\ O_p(1) & O_p(1) \end{bmatrix} .$$

It follows that

$$\hat{e}_t = \begin{bmatrix} x_{1t} + o_p(1) \\ x_{2t} + O_p(T^{1/2}) \end{bmatrix}$$

and, thus, estimates of Ω^x and Ψ^x based on $\hat{e}_t$ instead of x_t satisfy Assumption 4.2.

References

Bewley, R. and M. Yang (1995), Tests for Cointegration Based on Canonical Correlation Analysis, Journal of the American Statistical Association, 90, 990–996.

Boswijk, H.P. (1996), Testing Identifiability of Cointegration Vectors, Journal of Business & Economic Statistics, 14, 153–160.

Breitung, J. (1994), Some Simple Tests of the Moving-Average Unit Root Hypothesis, Journal of Time Series Analysis, 15, 351–370.

Breitung, J. (1995), Modified Stationarity Tests With Improved Power in Small Samples, Statistical Papers, 36, 77–95.

Choi, I. (1994), Residual-Based Tests for the Null of Stationarity With Applications to U.S. Macroeconomic Time Series, Econometric Theory, 10, 720–746.

Dickey, D.A. and W.A. Fuller (1979), Distribution of the Estimates for Autoregressive Time Series With a Unit Root, Journal of the American Statistical Association 74, 427–431.

Dickey, D.A. and S.G. Pantula (1989), Determining the Order of Differencing in Autoregressive Processes, Journal of Business and Economic Statistics, 5, 455–461.

Gonzalo, J. (1994), Five Alternative Methods of Estimation Long-run Equilibrium Relationships, Journal of Econometrics, 60, 203–233.

Harries, D. (1997), Principal Components Analysis of Cointegrated Time Series, Econometric Theory, 13,

Harries, D. and B. Inder (1994), A Test of the Null Hypothesis of Cointegration, in: C.P. Hargreaves (ed.), Nonstationary Time Series Analysis and Cointegration, Oxford: Oxford University Press, 133–152.

Haug, A. (1996), Tests for Cointegration: A Monte Carlo Comparison, Journal of Econometrics, 71, 89–115.

Holly, A. (1988), Specification Tests: an overview, in: T.F. Bewley, Fifth World Congress, Vol 1, p. 59–97.

Hubrich, K., H. Lütkepohl, and P. Saikkonen (1998), A Review of System Cointegration Tests, Humboldt-University Berlin, SFB 373, August 1998.

Johansen, S. (1988), Statistical Analysis of Cointegration Vectors, Journal of Economic Dynamics and Control, 12, 231–254.

Johansen, S. (1995), Likelihood-based Inference in Cointegrated Vectorautoregressive Models, Oxford: Oxford Univerity Press.

Kwiatkowski, D., P.C.B. Phillips, P. Schmidt, and Y. Shin (1992), Testing the Null Hypothesis of Stationary Against the Alternative of a Unit Root: How Sure are we that Economic Time Series Have a Unit Root?, Journal of Econometrics 54, 159–178.

Leybourne, S.J. and B.P.M. Mc Cabe (1994a), A Simple Test for Cointegration, Oxford Bulletin of Economics and Statistics, 56, 97–103.

Leybourne, S.J. and B.P.M. Mc Cabe (1994b), A Consistent Test for a Unit Root, Journal of Business & Economic Statistics, 12, 157–166.

Lütkepohl, H. (1991), Introduction to Multiple Time Series Analysis, Berlin: Springer.

Phillips, P.C.B. (1987), Time Series Regression With a Unit Root, Econometrica 55, 277–301.

Phillips, P.C.B. (1995), Fully Modified Least Squares and Vector Autoregressions, Econometrica, 1023–1078.

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

Quintos, C.E. (1998), Fully Modified Vector Autoregressive Inference in Partially Nonstationary Models, Journal of the American Statistical Association, 93, 783–795.

Saikkonen, P. (1991), Asymptotically Efficient Estimation of Cointegration Regressions, Econometric Theory, 7, 1–21.

Saikkonen, P. (1996), Testing Normalization and Overidentification of Cointegration Vectors in Cointegrated Vector Autoregressive Processes, Discussion Paper, University of Helsinki.

Saikkonen, P. and R. Luukkonen (1993), Testing for a Moving Average Unit Root in Autoregressive Integrated Moving Average Models, Journal of the American Statistical Association, 88, 596–601.

Sargan, J.D. and A. Bhargava (1983), Testing residuals from least squares regression for being generated by the Gaussian random walk, Econometrica, 51, 153–174.

Shin, Y. (1994), A Residual-Based Test of the Null of Cointegration Against the Alternative of No Cointegration, Econometric Theory, 10, 91–115.

Snell, A. (1998), Testing for r Versus $r-1$ Cointegrating Vectors, Journal of Econometrics, 88, 151–191.

Stock, J.H. (1994a), Deciding Between I(1) and I(0), Journal of Econometrics, 105–131.

Stock, J.H. (1994b), Unit Roots, Structural Breaks and Trends, in: R.F Engle and D:L McFadden (eds.), Handbook of Econometrics, Vol. 4, Amsterdam: North Holland.

Tanaka, K. (1990), Testing for a Moving Average Root, Econometric Theory 6, 433–444.

Tsay, R.S. (1993), Testing for Noninvertible Models With Applications, Journal of Business & Economic Statistics, 11, 225–233.

Appendix B: Critical Values

The critical values are computed from 10.000 realizations of the asymptotic expression, where Brownian motions are replaced by random walks with $T = 500$.

Table B.1: Critical Values (zero mean)

	CCA–statistic			CCA_a–statistic		
	0.10	0.05	0.01	0.10	0.05	0.01
$n=2$, $r=1$	4.822	6.372	9.616	8.710	10.503	14.770
$n=2$, $r=2$	10.376	12.220	15.934	n.a.	n.a.	n.a.
$n=3$, $r=1$	6.093	7.947	11.711	14.083	16.330	20.860
$n=3$, $r=2$	13.384	15.639	20.069	19.704	22.383	27.292
$n=3$, $r=3$	21.697	24.272	28.946	n.a.	n.a.	n.a.
$n=4$, $r=1$	7.177	9.350	13.956	18.780	21.543	26.442
$n=4$, $r=2$	16.013	18.479	23.566	28.966	31.959	38.133
$n=4$, $r=3$	26.193	28.774	34.037	34.163	37.286	43.979
$n=4$, $r=4$	36.650	39.665	46.015	n.a.	n.a.	n.a.
$n=5$, $r=1$	8.008	10.588	16.088	23.922	26.507	32.299
$n=5$, $r=2$	18.216	20.799	26.737	38.155	41.519	48.421
$n=5$, $r=3$	30.205	33.322	39.221	47.734	51.458	58.829
$n=5$, $r=4$	42.976	45.976	53.054	53.280	56.926	65.138
$n=5$, $r=5$	55.630	59.283	67.173	n.a.	n.a.	n.a.
$n=6$, $r=1$	8.494	11.231	16.858	28.086	31.150	37.481
$n=6$, $r=2$	20.006	22.944	28.946	46.776	50.605	58.008
$n=6$, $r=3$	33.436	36.580	43.590	60.574	64.932	73.827
$n=6$, $r=4$	48.370	52.044	58.780	70.328	74.647	83.109
$n=6$, $r=5$	63.827	67.888	75.844	76.473	80.603	88.713
$n=6$, $r=6$	78.571	83.178	91.861	n.a.	n.a.	n.a.

Table B.2: Critical Values (mean adjusted)

	CCA–statistic			CCA_a–statistic		
	0.10	0.05	0.01	0.10	0.05	0.01
$n=2,\ r=1$	1.060	1.647	3.453	10.503	12.460	16.436
$n=2,\ r=2$	13.842	15.753	20.122	n.a.	n.a.	n.a.
$n=3,\ r=1$	6.722	8.673	13.174	15.615	17.964	22.778
$n=3,\ r=2$	13.211	15.311	19.484	22.829	25.204	30.066
$n=3,\ r=3$	25.808	28.552	33.581	n.a.	n.a.	n.a.
$n=4,\ r=1$	1.547	2.225	4.005	20.430	22.899	27.985
$n=4,\ r=2$	17.043	19.602	24.702	32.058	34.777	40.607
$n=4,\ r=3$	27.433	30.157	35.572	38.750	41.782	48.296
$n=4,\ r=4$	41.802	44.710	50.524	n.a.	n.a.	n.a.
$n=5,\ r=1$	8.155	10.498	15.436	25.335	28.105	34.411
$n=5,\ r=2$	16.344	18.811	23.556	41.059	44.260	51.320
$n=5,\ r=3$	31.728	34.996	41.465	52.046	55.922	63.246
$n=5,\ r=4$	45.502	49.022	55.514	58.524	62.442	69.403
$n=5,\ r=5$	61.383	65.274	72.566	n.a.	n.a.	n.a.
$n=6,\ r=1$	2.118	2.811	4.713	29.703	32.700	39.091
$n=6,\ r=2$	19.945	22.903	29.395	49.614	53.519	60.683
$n=6,\ r=3$	33.569	36.664	42.972	64.707	68.999	77.761
$n=6,\ r=4$	50.469	54.170	61.740	75.812	80.346	89.405
$n=6,\ r=5$	67.537	71.501	79.508	82.639	86.966	95.530
$n=6,\ r=6$	85.062	89.607	99.066	n.a.	n.a.	n.a.

Table B.3: Critical Values (trend adjusted)

	CCA–statistic			CCA_a–statistic		
	0.10	0.05	0.01	0.10	0.05	0.01
$n = 2,\ r = 1$	1.920	2.688	5.145	13.612	15.852	20.318
$n = 2,\ r = 2$	19.787	22.140	26.959	n.a.	n.a.	n.a.
$n = 3,\ r = 1$	9.572	11.586	16.391	18.314	20.796	26.116
$n = 3,\ r = 2$	16.576	18.839	23.628	28.205	31.063	36.621
$n = 3,\ r = 3$	33.988	36.938	42.589	n.a.	n.a.	n.a.
$n = 4,\ r = 1$	2.414	3.162	5.217	22.866	25.462	31.311
$n = 4,\ r = 2$	21.884	24.766	30.964	36.858	40.287	46.568
$n = 4,\ r = 3$	33.354	36.423	42.830	46.341	49.814	57.097
$n = 4,\ r = 4$	52.309	55.532	62.394	n.a.	n.a.	n.a.
$n = 5,\ r = 1$	10.500	13.144	18.624	27.417	30.199	36.224
$n = 5,\ r = 2$	19.159	21.698	26.779	45.517	48.993	56.940
$n = 5,\ r = 3$	37.694	41.338	49.236	58.744	62.589	70.879
$n = 5,\ r = 4$	53.409	57.326	65.033	68.211	72.418	80.114
$n = 5,\ r = 5$	74.035	78.155	85.836	n.a.	n.a.	n.a.
$n = 6,\ r = 1$	3.092	3.864	5.733	31.781	34.886	41.504
$n = 6,\ r = 2$	24.081	27.390	33.679	54.037	57.635	65.274
$n = 6,\ r = 3$	38.532	41.969	48.534	71.176	75.679	84.773
$n = 6,\ r = 4$	58.544	62.866	71.545	84.523	88.930	99.143
$n = 6,\ r = 5$	77.272	81.477	90.417	93.878	98.423	107.645
$n = 6,\ r = 6$	99.808	104.466	114.275	n.a.	n.a.	n.a.