Diskussionsschriften

**A Comparison of Johansen's, Bierens
and the Subspace Algorithm Method
for Cointegration Analysis**

Martin Wagner

02-10

October 2002

A Comparison of Johansen's, Bierens' and the Subspace Algorithm Method for Cointegration Analysis

Martin Wagner [*]
Department of Economics
University of Bern
Gesellschaftsstrasse 49
CH-3012 Bern

Abstract

The methods listed in the title are compared by means of a simulation study and a real world application. The aspects compared in the simulations are: The performance of the tests of the different methods for the dimension of the cointegrating space and the quality of the estimated cointegrating space. It turns out that the subspace algorithm method, formulated in the state space framework and thus applicable for ARMA processes, performs at least comparable to the Johansen procedure and both perform significantly better than Bierens' method.

The real world application is an investigation of the long-run properties of the neoclassical growth model for Austria. It turns out that the results do not fully support the theoretical predictions and that they are very versatile across the employed methods. The degree of versatility depends strongly upon the number of variables. For the case of 6 variables and about 100 observations huge differences occur, which lead us to conclude that the results of this typical situation in the applied literature should be interpreted with more caution than is commonly done.

JEL Classification: C13, C15, C32, E20

Keywords: Cointegration, State Space Models, Subspace Algorithms, Simulation, Neoclassical Growth Model

1   Introduction

The majority of cointegration studies is performed in the context of the vector autoregressive (VAR) model using the methods developed over the years by Johansen, see his monograph (Johansen, 1995) for a detailed description. Johansen derives the maximum likelihood estimate for Gaussian cointegrated VAR processes. Within this framework a variety of questions, e.g. testing for hypotheses on the cointegrating vectors, the extension to higher integration orders, have also been addressed, contributing further to the popularity of the method.

Several authors have analyzed the behavior of Johansen's method an a VAR approximation in case the data are generated by a vector ARMA process, see e.g. Saikkonen (1992) or Saikkonen and Luukkonen (1997). These authors show that, if the lag length of the autoregressive approximation is increased sufficiently with the sample size, consistency of the method carries over to the (vector) ARMA case. This result is essentially a generalization of the result of Said and Dickey (1984) concerning the applicability of the augmented Dickey-Fuller tests to ARMA processes to the multivariate case.

The available asymptotic results may however not be the best guide for the finite sample sensitivity of the Johansen procedure to the fact that the data may be generated by an ARMA processes. It is thus interesting to see whether some of the methods developed explicitly for ARMA processes deliver, for small samples at least, better results, in terms of e.g. the performance of the respective tests for cointegration, than the Johansen procedure. If this is the case, then the results obtained by applying these other methods can be used at least to back-up or cast doubt on results obtained by applying the Johansen procedure. Given that the additional results are confirmative, one can proceed in the econometric analysis as usual, i.e. one can continue to use well developed VAR techniques (like testing hypotheses on the cointegrating space or testing for exogeneity). If not, then one should resort directly to ARMA based techniques. Note that we only focus on the presence of ARMA errors in the DGP, and neglect other important issues in model specification like outliers, heteroskedasticity, non-normality or structural breaks.

In this paper we compare the results obtained by applying the Johansen procedure with the results obtained from two procedures designed explicitly for ARMA processes. The comparison is both by means of simulations and by comparing the results obtained with the various methods in a real world application. The two methods that we use for comparison are Bierens'

(1997) nonparametric cointegration analysis and the subspace algorithm cointegration analysis put forward in Bauer and Wagner (2002a). We choose these two methods out of the menu of available methods for cointegration analysis in the context of ARMA processes (see e.g. Boswijk et al., 1999, Phillips, 1991, 1995 or Yap and Reinsel, 1995), because of their low computational load, compared to e.g. the nonlinear optimization problem that has to be solved in pseudo ML estimation for ARMA processes. Their low computational load is however the only common feature between the two methods. Bierens' method, due to its nonparametric nature, abstains from estimating the parameters of the underlying rational process and delivers only estimates of the cointegrating space and a test for its dimension. Subspace algorithm cointegration analysis on the other hand delivers estimates of all parameters. It does so, however, in two respects different from usual in the literature. First, it is based and formulated in the state space framework and secondly it uses *subspace algorithms* (see the description in Section 2.3). In a series of papers Bauer and Wagner (2002a-d) develop the structure theory of cointegrated state space models, as well as estimators and tests. The state space framework is an equivalent way to represent ARMA processes that turns out to have some advantages for cointegration analysis, in terms of e.g. simplicity of the representation especially for higher order integrated systems (see Bauer and Wagner, 2002b). By applying subspace algorithms it is possible to estimate all parameters in a computationally extremely cheap way by performing OLS regressions and singular value decompositions. In Bauer and Wagner (2002a) one of these algorithms, which have up to now only been used in a stationary context and which originate in the engineering literature, has been modified to obtain consistent estimates also in the I(1) case.

We compare the results, as already indicated above, on simulated ARMA data and on a real world application (see the next paragraph). In the simulations we are interested in two aspects. First the performance of the tests corresponding to the different applied methods. There are in total 9 different tests: 6 subspace, 2 Johansen and 1 Bierens. The performance measure employed is the acceptance frequency of the correct dimension of the cointegrating space as a result of the testing sequence. The second issue investigated is the quality of the estimated cointegrating space. As a measure of quality we employ the Hausdorff distance (for a definition see Section 3) between the estimated and the true cointegrating space. The results are mixed. For the simulated 2-dimensional ARMA processes for small samples the subspace algorithm cointegration analysis shows better performance (in terms of e.g. the tests) and

for the 3-dimensional processes the Johansen procedure on a VAR approximation leads in many cases to better test results for small samples. The estimated cointegrating spaces are essentially equally well estimated by these two methods. Across all simulations the Bierens method shows the worst performance, both for the tests and the cointegrating spaces.

The real world application is an investigation of the long-run implications of the neoclassical one-sector growth model for the Austrian economy. This model has under some assumptions, see Section 4, clear implications concerning cointegration. None of the methods lends support to the full predictions of the theoretical model, but the even more interesting observation in our context is the fact that the methods lead to quite versatile results, especially concerning the estimated cointegrating spaces (measured by the mutual Hausdorff distances between the cointegrating spaces estimated with the different methods and the distances to the theoretically formulated cointegrating spaces). If one investigates in a bivariate set-up the consumption-output relation the results, at least for the estimated cointegrating space, become more similar between the methods. Thus, from the application we observe that for the usual sample size of about 100 observations the variability of results across different methods depends strongly upon the number of variables. For 5 or 6 variables the differences are enormous. Hence, the results of these exercises typical in the literature should probably be interpreted with more caution than is common practice. In the appendix a couple of further results like stability of the cointegrating space and detailed hypotheses testing results on the cointegrating space are reported.

The paper is organized as follows: In Section 2 the applied methods are discussed, where only the subspace algorithm cointegration analysis is presented in detail. In Section 3 the results of the simulation study are presented. In Section 4 the application to the neoclassical growth model is discussed and Section 5 briefly summarizes and concludes. In the appendix some additional empirical results derived from further investigations of the application discussed in Section 4 are presented.

2 A Description of the Applied Econometric Methods

In this section we describe the applied methods. The descriptions of Johansen's and Bierens' method are very sketchy, as especially the first method is most widely known and used throughout the profession and also the second method has been described at various places. Thus, for these two methods only the main arguments will briefly be summarized to set the

stage. For the third method, the subspace algorithm cointegration analysis, the description is more detailed. There are, we think, two main reasons for doing so: First of all, this method is formulated in the state space framework. The state space framework offers an equivalent and alternative way for representing ARMA processes, which turns out to be very convenient for cointegration analysis (see e.g. Bauer and Wagner, 2002b). Despite its potential advantages, it is not in widespread use in the cointegration literature, from which we conclude that a few comments on this framework might be interesting for some readers unacquainted with this approach. After that, we describe the method (developed in Bauer and Wagner, 2002a). Both the estimation of the system matrices and the tests for the cointegrating rank are discussed. As subspace algorithm are hardly used in an econometric context, contrary to their increased usage in the control and engineering literature, it might also be interesting to present the ideas of these algorithms in some detail as well. The idea of this type of algorithms is simple, however the description turns out to require some space. Note at this point only the fact that with the investigated adapted subspace algorithm it is possible to estimate the parameters of cointegrated vector ARMA models by OLS regressions and singular value decompositions. Readers acquainted with some or all of the methods can skip the respective subsections without loss.

2.1 Johansen VAR Cointegration Analysis

The method developed and extended in a variety of ways by Johansen over a couple of years and nicely summarized in his monograph (Johansen, 1995) is the workhorse in cointegration analysis and therefore constitutes a natural benchmark to compare other methods with. As just mentioned above, we abstain from yet another description of the method and only want to describe the basic set-up and a few issues relevant for our study.

Johansen's approach is developed in the context of the I(1) vector autoregressive model. Assuming normality for the innovations ε_t, maximum likelihood estimates for all parameters and likelihood ratio tests for the dimension of the cointegrating space are derived. The estimation problem is formulated in the *error correction representation*, which is, ignoring deterministic components for notational simplicity, given by

$$\Delta y_t = \alpha \beta' y_{t-1} + \Gamma_1 \Delta y_{t-1} + \ldots + \Gamma_{p-1} \Delta y_{t-p+1} + \varepsilon_t$$

where $y_t \in \mathbb{R}^s$ is the time series observed for $t = 1, \ldots, T$, $\alpha, \beta \in \mathbb{R}^{s \times r}$, $\Gamma_i \in \mathbb{R}^{s \times s}$ and ε_t is i.i.d. normally distributed with variance matrix Σ. The columns of the matrix $\beta \in \mathbb{R}^{s \times r}$, with

5

$0 \leq r \leq s$, span the cointegrating space. As is well known, estimation and testing proceeds by first regressing Δy_t and y_{t-1} on the lagged differences and then the residual moment matrices of these two regressions are the input in a generalized eigenvalue problem. Hence, the cointegration testing problem is based on the canonical correlations between Δy_t and y_{t-1} corrected for the short-run dynamics. The estimate for the r-dimensional cointegrating space is given by the canonical variates, i.e. the corresponding generalized eigenvectors, to the r largest canonical correlations.

The approach has been extended in many directions. See e.g. Boswijk et al. (1999) for an overview and a simulation study of the effects of relaxing the assumption of normality. In case that the innovations are not normally distributed, the approach can be extended to account for the non-normality both in the estimation and in the testing step, see also Lucas (1998). More interesting for our study are however the available results concerning robustness of the Johansen approach to ARMA data generating processes. As already mentioned in the introduction, Saikkonen (1992), see also Saikkonen and Luukkonen (1997), shows that the method applied to an autoregressive approximation, where the lag length is growing sufficiently with the sample size, results in consistent estimates of all parameters. Furthermore also the null distribution of the test statistics remains unchanged compared to the case that the DGP is indeed a VAR process. Wagner (1999) shows that the cointegrating space is estimated consistently for ARMA date generating processes even in a fixed order VAR approximation, in this case however no distributional results are available. These robustness results with respect to ARMA DGPs, notwithstanding the formulation in the popular VAR framework, form, to our understanding, one additional important basis for the popularity of the method. However, the asymptotic results are not necessarily a guarantee for good finite sample approximation properties. Hence, it may be interesting to compare the results concerning cointegration obtained with the Johansen procedure, based on a VAR approximation, with the results obtained from methods that are designed for ARMA processes, which is done in the simulation study presented in Section 3.

2.2   Bierens Nonparametric Cointegration Analysis

The method developed by Bierens (1995, 1997a, 1997b) is developed to test for cointegration in I(1) ARMA processes. The method only aims at testing for the dimension of the cointegrating space and at obtaining an estimate of the basis of the cointegrating space. No other parameters

of the underlying data generating process are estimated. For detailed descriptions of the method the reader is referred to the metnioned papers of Bierens. Here we only want to briefly present the main idea of the approach. Let $\{F_k, k = 1, 2, \ldots\}$ be a sequence of functions on $[0, 1]$ that satisfy $\int_0^1 F_k(z)dz = 0$ and $\int_0^1 F_j(z)F_k(z)dz = 0$ for $j \neq k$. In particular two choices are proposed by Bierens, namely $F_k(z) = \cos(2k\pi z)$ and $F_k(z) = \cos(2k\pi(1 - \frac{1}{2T}))$, where T denotes again the sample size. These choices of the function sequence F_k maximize the power function of the resulting test. The second variant, $F_k(z) = \cos(2k\pi(1 - \frac{1}{2T}))$, makes the cointegration test invariant with respect to the presence of deterministic trends in the data. Based on the above function sequences, next define for an observed time series $y_t, t = 1, \ldots, T$ the quantity

$$M_k(y) = \frac{1}{T} \sum_{t=1}^{T} F_k(\frac{t}{T})y_t$$

Now, if y_t is I(1), it can be shown that $M_k(y)$ is $O_p(T^{1/2})$. For both, the first differences Δy_t and for stationary linear combinations, $\beta' y_t$, say, it holds that $M_K(\Delta y)$ and $M_k(\beta' y)$ are $O_p(T^{-1/2})$. This difference in the orders of magnitude between the stationary and the nonstationary directions is exploited in the construction of the test statistics.

In order to perform the test, a generalized eigenvalue problem invoking weighted sums of products of $M_k(y)$ and of $M_k(\Delta y)$ has to be solved. Besides the test for the dimension of the cointegrating space, Bierens proposes also an estimate for its dimension. The estimates of a basis of the cointegrating space are obtained as generalized eigenvectors of a second generalized eigenvalue problem that involves only weighted sums of products $M_k(y)$.

2.3 Subspace Algorithm Cointegration Analysis

In this subsection we describe and discuss the subspace algorithm cointegration analysis introduced in Bauer and Wagner (2002a).[1] The state space framework is an equivalent framework of representing and modelling (vector) ARMA processes, that is not widely used in the cointegration literature, despite some convenient features. Thus, we think that a relatively more detailed description of the approach might be helpful to some readers. The underlying structure theory (including the development of a canonical form, a parameterization and their topological properties) is developed in Bauer and Wagner (2002b) for the class of ARMA processes with unit roots with integer integration orders at any point on the unit circle. The

---

[1]For details including proofs concerning the proposed method see Bauer and Wagner (2002a). A very detailed description of all computational aspects is also contained in Bauer and Wagner (2002e).

discussed estimation method for the I(1) case rests upon these general results.

For a detailed discussion of the state space framework see e.g. Aoki (1990) or Hannan and Deistler (1988). If y_t follows an ARMA process, then a *minimal* state space representation of the following form exists:

$$
\begin{aligned}
x_{t+1} &= Ax_t + K\varepsilon_t \\
y_t &= Cx_t + E\varepsilon_t
\end{aligned}
\tag{1}
$$

where y_t denotes the s-dimensional output series observed for $t = 1,\ldots,T$. ε_t denotes an s-dimensional white noise sequence. $A \in \mathbb{R}^{n\times n}$, $K \in \mathbb{R}^{n\times s}$, $C \in \mathbb{R}^{s\times n}$, $E \in \mathbb{R}^{s\times s}$, E nonsingular, and $x_t \in \mathbb{R}^n$ denotes the n-dimensional unobserved state sequence. The noise sequence ε_t is assumed to be a strict martingale difference sequence, the precise conditions are formulated in Bauer and Wagner (2002a). *Minimality* refers to the fact that there exists no other state space representation with a smaller state dimension, and it is the state space analogue to left coprimeness. We restrict ourselves to systems that are strictly minimum-phase, i.e. to systems where the eigenvalues of $(A - KE^{-1}C)$ are smaller than 1 in absolute value. The eigenvalues of A determine the integration properties of the solution y_t of the state space equations. This can directly be seen by solving the state space equations, assuming $x_0 = 0$ for notational simplicity:

$$
y_t = E\varepsilon_t + \sum_{j=1}^{t-1} CA^{j-1}K\varepsilon_{t-j}
$$

The sum of the matrix products $CA^{j-1}K$ converges if and only if all eigenvalues of A are smaller than one in absolute value, eigenvalues on the unit circle correspond to (seasonal) integration of y_t. It is shown in Bauer and Wagner (2002b) that the process y_t is I(1) if and only if the eigenvalues of A are either smaller than one in absolute value or equal to one, with Jordan blocks all of size equal to one.[2] In this case, the system can equivalently also be written, after a suitable coordinate change in the state, as:

$$
y_t = \begin{bmatrix} C_1 & C_{st} \end{bmatrix} x_t + \varepsilon_t
\tag{2}
$$

$$
\begin{bmatrix} x_{t+1,1} \\ x_{t+1,st} \end{bmatrix} = \begin{bmatrix} I_c & 0 \\ 0 & A_{st} \end{bmatrix} \begin{bmatrix} x_{t,1} \\ x_{t,st} \end{bmatrix} + \begin{bmatrix} K_1 \\ K_{st} \end{bmatrix} \varepsilon_t
\tag{3}
$$

where $x_{t,1} \in \mathbb{R}^c$ denotes the nonstationary part of the state and $x_{t,st} \in \mathbb{R}^{n-c}$ denotes the stationary part of the state and $x_t = \begin{bmatrix} x_{t,1}', x_{t,st}' \end{bmatrix}'$. By c we denote the number of common

---

[2]It can be shown that also for higher integration orders the eigenvalue structure of A in a minimal representation determines the integration structure of y_t, as well as that of x_t. For instance it can be shown that the integration order corresponding to any unit root equals the length of the longest Jordan chain corresponding to the respective eigenvalue of A on the unit circle.

trends. In a minimal state space representation, c common trends are present in both the state and the observations. It can easily be shown (see Bauer and Wagner, 2002a) that for a minimal state space representation $c \leq s$ has to hold, irrespective of the system order n. From the structure of the state space representation one obtains,

$$
\begin{aligned}
y_t &= C_1 x_{t,1} + C_{st} x_{t,st} + E\varepsilon_t \\
&= C_1 K_1 \sum_{j=1}^{t-1} \varepsilon_{t-j} + k_{st}(z)\varepsilon_t
\end{aligned}
\tag{4}
$$

where $k_{st}(z) = E + zC_{st}(I_{n-c} - A_{st})^{-1}K_{st}$. Due to the fact that all eigenvalues of A_{st} are inside the unit circle, $k_{st}(z)$ is a stable transfer function. z is used to denote both the backward shift operator as well as a complex variable. Thus, the above representation (4) coincides with Granger's, with the first term corresponding to the common trends.

It is clear that in the above representation only the product $C_1 K_1$ is identified. The number of common trends present in y_t is given by the rank of $C_1 K_1$, which is less or equal to s.[3] In a minimal representation the ranks of both C_1 and K_1 are equal to c. Now it immediately follows, that if r denotes the number of cointegrating relationships for y_t, the equality $c = s - r$ holds and c is – as already stated – the number of common trends present in both the state x_t and the observations y_t. This relationship just shows from a state space perspective the well known fact that for I(1) processes the sum of the number of common trends and the number of cointegrating relationships is equal to the dimension of y_t.

To achieve identification of $C_1 K_1$ further restrictions have to be imposed on the system representation. In the canonical form presented in Bauer and Wagner (2002b), this is achieved by choosing C_1 to be part of an orthonormal matrix, i.e. $C_1 \in \mathbb{R}^{s \times c}$ and $C_1'C_1 = I_c$ is assumed.[4] Therefore there exists a matrix $C_1^\perp$ with $(C_1^\perp)'C_1^\perp = I_r$ and $(C_1^\perp)'C_1 = 0$, i.e. $C_1^\perp$ spans the orthogonal complement of C_1. By pre-multiplying equation (4) with $(C_1^\perp)'$, it immediately follows that the columns of $C_1^\perp$ span the cointegrating space of dimension r. Note that due to the block-diagonal set-up (due to the use of the Jordan normal form) the stationary part, $k_{st}(z)$ can be investigated independently of the nonstationary part. Thus, for the stationary part any developed canonical representation can be chosen.

After having briefly discussed the set-up, let us now turn to a description of the estimation approach. Subspace algorithms, up to now only used for stationary processes, originated

---

[3]This reflects once again the fact that at most s stochastic trends can be present in y_t.

[4]These restrictions are not sufficient for identifiability in the general case and some further restrictions are needed. However, the additional restrictions are not relevant for the present setting and thus we refer to Bauer and Wagner (2002b) for details. The asymptotically stationary part can be parameterized using any of the known parameterizations for stationary processes, like echelon parameters.

in the engineering literature in the 1980ies and provide a computationally cheap alternative to e.g. (pseudo) maximum likelihood estimation procedures. In the subspace algorithm as described below, first estimates of the unobservable state x_t are obtained and then estimates of the system matrices are obtained. This is different to e.g. ML estimation, where first estimates of the system matrices are obtained and based on these an estimate of the state is obtained.

The basic idea of subspace algorithms lies in the interpretation of the state, it can be shown that the best linear prediction of y_{t+j} given $y_t, \ldots, y_1$ and x_1 is given by $CA^j x_t$. Thus, the state is a basis for the predictor space for the whole future of y_t, i.e. for $y_{t+j}, j \geq 0$ and is contained in the past of the time series, $y_{t-j}, j \geq 1$. This fact is exploited in the construction of the algorithm as follows. Choose two integers f and p, both larger or equal than n, and define $Y_{t,f}^+ = [y_t', y_{t+1}', \ldots, y_{t+f-1}']'$ and $Y_{t,p}^- = [y_{t-1}', y_{t-2}', \ldots, y_{t-p}']'$. Further let $E_{t,f}^+ = [\varepsilon_t', \varepsilon_{t+1}', \ldots, \varepsilon_{t+f-1}']'$. Let $\mathcal{O}_f = [C', A'C', \ldots, (A^{f-1})'C']'$ and $\mathcal{K}_p = [K, (A - KE^{-1}C)K, \ldots, (A - KE^{-1}C)^{p-1}K]$. Finally define $\mathcal{E}_f$ as the matrix, whose i-th block row is equal to the matrix $[CA^{i-1}K, \cdots, CK, E, 0]$. Then it follows from the system equations (1), that

$$Y_{t,f}^+ = \mathcal{O}_f \mathcal{K}_p Y_{t,p}^- + \mathcal{O}_f (A - KE^{-1}C)^p x_{t-p} + \mathcal{E}_f E_{t,f}^+$$

Noting that for $p \to \infty$ the term $(A - KE^{-1}C)^p$ vanishes, the above observations lead to the following procedure:

1) In a first step regress $Y_{t,f}^+$ on $Y_{t,p}^-$ to obtain an estimate $\hat{\beta}_{f,p}$ of $\mathcal{O}_f \mathcal{K}_p$.[5]

2) Typically $\hat{\beta}_{f,p}$ has full rank, whereas $\mathcal{O}_f \mathcal{K}_p$ has rank n for $f, p \geq n$, where n again denotes the true system order. Thus, for an appropriate estimate of n, $\hat{n}$ say (see below for one way of obtaining a consistent order estimate), approximate $\hat{\beta}_{f,p}$ by a rank n approximation with decomposition $\hat{\mathcal{O}}_f \hat{\mathcal{K}}_p$.

3) Use the derived estimate $\hat{\mathcal{K}}_p$ to obtain an estimate of the state $\hat{x}_t = \hat{\mathcal{K}}_p Y_{t,p}^-$.

4) Given the estimated state, the system equations (1) can be used to obtain estimates $(\hat{A}, \hat{K}, \hat{C}, \hat{E})$ of the system matrices (A, K, C, E) via OLS regressions.

The rank n approximation in step 2 of the procedure outlined above is not performed on $\hat{\beta}_{f,p}$, but is based on a singular value decomposition (SVD) of a transformed matrix $\hat{W}_f^+ \hat{\beta}_{f,p} \hat{W}_p^-$.

---

[5]The construction of the variables $Y_{t,f}^+$ and $Y_{t,p}^-$ implies that the sample range in the regressions is $t = p+1, \ldots, T-f+1$. We denote the effective sample size by $T_{f,p} = T - f - p + 1$.

Variants of existing subspace algorithms differ i.a. in their choice of respective weighting matrices. In the employed algorithm they are given by $\hat{W}_f^+ = (\hat{\Gamma}_f^+)^{-1/2}$ and $\hat{W}_p^- = (\hat{\Gamma}_p^-)^{1/2}$, where $\hat{\Gamma}_f^+ = \frac{1}{T-f-p+1} \sum_{t=p+1}^{T-f} Y_{t,f}^+(Y_{t,f}^+)'$ and $\hat{\Gamma}_p^- = \frac{1}{T-f-p+1} \sum_{t=p+1}^{T-f} Y_{t,p}^-(Y_{t,p}^-)'$. Thus, in this algorithm the canonical correlations between $Y_{t,f}^+$ and $Y_{t,p}^-$ are estimated. Fix n for the moment, then the SVD is decomposed in two parts:

$$\hat{W}_f^+ \hat{\beta}_{f,p} \hat{W}_p^- = \hat{U}\hat{\Sigma}\hat{V}' = \hat{U}_n\hat{\Sigma}_n\hat{V}_n' + \hat{R}$$

where $\hat{U}_n \in \mathbb{R}^{fs \times n}, \hat{V}_n \in \mathbb{R}^{ps \times n}$ and $\hat{\Sigma}_n \in \mathbb{R}^{n \times n}$. Here $\hat{\Sigma}_n = \mathrm{diag}(\hat{\sigma}_1, \ldots, \hat{\sigma}_n)$ contains the n dominant singular values ordered decreasing in size, i.e. $1 \geq \hat{\sigma}_1 \geq \ldots \geq \hat{\sigma}_n > 0$. The matrices $\hat{U}_n$ and $\hat{V}_n$ contain the corresponding left and right singular vectors. The remaining singular values and vectors are attributed to $\hat{R}$ and are neglected. The rank n approximation of $\hat{\beta}_{f,p}$ is now given by $\hat{\mathcal{O}}_f \hat{\mathcal{K}}_p = [(\hat{W}_f^+)^{-1}\hat{U}_n][\hat{\Sigma}_n \hat{V}_n'(\hat{W}_p^-)^{-1}]$ and thus $\hat{\mathcal{K}}_p = \hat{\Sigma}_n \hat{V}_n'(\hat{W}_p^-)^{-1}$. Concerning now finally the choice of n, an order estimation criterion is based on the size of the first neglected singular value in $\hat{\Sigma}$. This rests (as the other arguments above, for details see Bauer and Wagner, 2002a) upon the asymptotic argument that in $\hat{\Sigma} = \mathrm{diag}(\hat{\sigma}_1, \ldots, \hat{\sigma}_{min(f,p)s})$ for a system of order n, the n largest singular values have positive limits and the remaining singular values converge to 0. Hence, a consistent order estimate, $\hat{n}$ say, is obtained by minimizing the following criterion SVC:

$$\hat{n} = \mathrm{argmin}_{1 \leq n < \min(f,p) \times s} SVC(n) = \mathrm{argmin}_{1 \leq n < \min(f,p) \times s} (\hat{\sigma}_{n+1}^2 + 2nsH_T/T) \qquad (5)$$

Here $H_T > 0, H_T/T \to 0$ denotes a penalty term, which determines the asymptotic properties of the estimated order.

This standard algorithm (consistent for stationary processes) as just described, has to be *adapted* in order to result in consistent estimates also for I(1) processes. The modification is necessary to separate the stationary and the nonstationary components of y_t in order to establish consistency for the stationary part of the transfer function, $k_{st}(z)$. However, a description of the standard version of the algorithm is useful and necessary, as for the adapted version estimates of C_1 obtained from the standard algorithm are employed. For correctly specified c, the standard subspace estimate $\hat{C}_1$ of C_1 is super-consistent. Let, as before denote by r the true cointegrating rank[6], then $c = s - r$ common trends drive the system. Denote with $\hat{\bar{C}} = [\hat{C}_1, \hat{C}_1^{\perp}]'$, where $\hat{C}_1^{\perp} \in \mathbb{R}^{s \times r}, \hat{C}_1'\hat{C}_1^{\perp} = 0$ and $(\hat{C}_1^{\perp})'\hat{C}_1^{\perp} = I_r$.

---

[6]Tests for the cointegrating rank will be presented after the discussion of the adapted subspace algorithm.

Define a new weighting matrix, with $\otimes$ denoting the Kronecker product, $W^{\hat{+}}_{f,C_1} = [(I \otimes \hat{\tilde{C}})\frac{1}{T_{f,p}}\sum_{t=1}^{T} Y^+_{t,f}(Y^+_{t,f})'(I \otimes \hat{\tilde{C}})']^{-1/2}(I \otimes \hat{\tilde{C}})$, using again the Cholesky decomposition as the square root of a matrix. In combination with the modified weighting matrix also the estimate for $\hat{\mathcal{K}}_p$ has to be modified. For any choice of weighting matrices, the estimated matrix $\hat{\mathcal{K}}_p = \hat{\Sigma}_n\hat{V}'_n(\hat{W}^-_p)^{-1}$ can alternatively be written as $\hat{\mathcal{K}}_p = \hat{U}'_n\hat{W}^+_f\hat{\beta}_{f,p}$. Now, if the modified weighting matrix $W^{\hat{+}}_{f,C_1}$ is used, the corresponding matrix of left singular vectors $\hat{U}_n$ has to be changed to $\hat{U}_{n,c}$, where

$$\hat{U}_{n,c} = \left[\begin{array}{cc} I_c & 0^{c\times(n-c)} \\ 0^{(fs-c)\times c} & \hat{U}(2,2) \end{array} \right],$$

which provides the required separation of the non-stationary components from the stationary components. $\hat{U}(2,2)$ denotes the $(2,2)$-block of the matrix $\hat{U}_n$.[7] Thus, under the assumption of a correctly specified number of common trends, c, the subspace procedure can be modified as follows:

1) Perform steps 1) to 4) of the standard subspace algorithm as described above.

2) Use the estimate $\hat{C}_1$ to construct the modified weighting matrix $W^{\hat{+}}_{f,C_1}$.

3) Generate the adapted estimate of $\hat{\mathcal{K}}_{p,C_1} = \hat{U}'_{n,c}W^{\hat{+}}_{f,C_1}\hat{\beta}_{f,p}$.

4) Use the adapted estimate $\hat{\mathcal{K}}_{p,C_1}$ to obtain the adapted estimate of the state vector $\hat{x}_{t,C_1} = \hat{\mathcal{K}}_{p,C_1}Y^-_{t,p}$.

5) Use, as in item 4) of the standard algorithm, the system equations, to obtain estimates $(\hat{A}_{C_1}, \hat{K}_{C_1}, \hat{C}_{C_1}, \hat{E}_{C_1})$ of the system matrices via OLS regressions.

For stationary processes, i.e. when $r = s$ and thus $c = 0$, the adapted procedure coincides with the standard procedure. Note that both, the initial estimate $\hat{C}^\perp_1$ obtained from the standard algorithm, as well as $\hat{C}^\perp_{\hat{C}_1,1}$ are consistent estimates of the cointegrating space. Note also that the above OLS approach can be modified to a reduced rank regression in order to incorporate the integration structure also in the estimation of A, for details on this reduced rank regression approach see again Bauer and Wagner (2002a).

In the above discussion, the number of common trends is assumed to be known or correctly specified. In order to make the approach useful, we next need to discuss tests for the number of common trends. There are two possibilities to construct tests. One idea is to base tests

[7]From a theoretical point of view, to achieve consistency only the $(2,1)$-block of the matrix $\hat{U}_n$ has to be replaced by a 0-block.

Test Nr.	I	II	III	IV
Test Stat.	$T\mathrm{re}(\hat{\mu}_c)$	$T\sum_{i=1}^{c}\mathrm{re}(\hat{\mu}_i)$	$T\mathrm{abs}(\hat{\mu}_c)$	$T\sum_{i=1}^{c}\mathrm{abs}(\hat{\mu}_i)$

Table 1: The 4 tests based on the eigenvalues $\hat{\mu}$ of the matrix $\hat{A}_{C_1} - I_n$. Under the null hypothesis of c common trends, the first c columns of the standard subspace estimate of C are chosen as C_1 and used for the construction of the modified weighting matrix $W_{f,C_1}^{\hat{+}}$. re denotes the real part of a (possibly) complex number and abs denotes the absolute value. Critical values for the tests are given in Bauer and Wagner (2002e).

on the estimated singular values in $\hat{\Sigma}$ noting that the first c singular values converge to 1, whereas the other singular values have limits smaller than 1 (and 0 respectively). See Bauer and Wagner (2000, 2002a) for a description, an application and a performance evaluation of this test.[8] Similarly in spirit to the order estimation criterion, where a distinction has to be made between singular values significantly different from zero, an estimate of the number of common trends can be obtained by determining the number of singular values statistically not distinguishable from 1. The results from this approach in combination with sub-sequent testing are often found to have good performance, however the properties depend upon the chosen penalty term, for details see Bauer and Wagner (2002a).

The other idea is to base tests on the eigenvalues of $\hat{A}$, remembering that in the I(1) case the number of common trends equals the number of eigenvalues of A equal to 1. In fact we base the tests on the eigenvalues of the matrix $\hat{A} - I_n$ and compare their magnitude with 0. The reason for doing this is that in this formulation the asymptotic distribution is directly of the standard integral of Brownian motions type. Corresponding tests for the null hypothesis of c common trends can either be based on the c-th largest eigenvalue alone or on the c largest eigenvalues, furthermore one can also use the real parts of the estimated eigenvalues or their absolute values. This leads to 4 different tests, see Table 1. The number of common trends c is determined in a recursive testing sequence, usually starting with an initial null of a maximal possible number of common trends s. Note that in each step of the sequence the system has to be re-estimated using the matrix $\hat{C}_1$ from the standard algorithm with the number of columns corresponding to the number of common trends under the null hypothesis at this step of the sequence.

Two more tests are discussed in Bauer and Wagner (2002d), where the Johansen procedure

---

[8]The drawback of this approach is that the singular value based test depends upon nuisance parameters. Therefore in this study this test is not investigated further.

is replicated on the estimated state $\hat{x}_t$. Note again that the state follows an AR(1) equation. Hence, the Johansen procedure is very simple in that case and just amounts to a computation of the canonical correlations between $\hat{\Delta}x_t$ and $\hat{x}_{t-1}$. Thus, for the n-dimensional state, the null hypothesis of c common trends can also be tested by performing a Johansen cointegration test on the state equation with the null hypothesis of $(n-c)$ linearly independent cointegrating relationships. This observation gives rise to two additional tests, replicating the Johansen trace test (test number V) and the Johansen max test (test number VI). Again the testing is performed sequentially, where a difference to a standard Johansen application in a VAR model is that after each step of the testing sequence the system has to be re-estimated, compare the description of the sequence for tests I to IV.

3   A Simulation Study

In this section we compare the performance of the described methods on simulated ARMA(1,1) processes for sample sizes $T = 100, 200, 300, 400$ and 1000. We investigate two aspects, the acceptance frequencies of the correct decisions for the dimension of the cointegrating spaces of the discussed tests and the approximation quality of the estimated cointegrating spaces to the true cointegrating spaces.

The lag order of an autoregressive approximation of the ARMA processes that is needed in the Johansen procedure is determined by minimizing AIC. The indices f, p required in the subspace algorithm are chosen to equal twice these values, see Bauer and Wagner (2002a) for a discussion on this choice. Note at this point that the results for the Johansen procedure depend upon the fact how well a (low order) VAR can approximate the underlying ARMA model. This depends upon the zeros of the transfer function, i.e. upon the zeros of $\det b(z)$ in an ARMA representation or equivalently upon the eigenvalues of $(A - KE^{-1}C)$ in a state space representation. More precisely, the closer the roots of $\det b(z)$ are to the unit circle, equivalently the closer the eigenvalues of $(A - KE^{-1}C)$ are to the unit circle, the more lags have to be included in a VAR approximation in order to approximate the underlying transfer function well. The minimum absolute value of the roots of $\det b(z)$ is henceforth denoted by ρ_0.

With respect to the tests we compare in total 9 different tests (6 subspace, 2 Johansen, 1 Bierens) and are interested in the relative performance of these tests depending upon the sample size and upon the extent of autocorrelation of the errors. All test results reported,

14

present acceptance frequencies of the correct dimension of the cointegrating space, with each step of the test sequence carried out at a nominal size of 5%.

The estimation of the cointegrating space itself is a prime issue in cointegration analysis and the properties of methods of cointegration analysis in concern to the estimation of the cointegrating space are clearly very important. All methods, the subspace algorithm, Bierens' method and the Johansen approach estimate the cointegrating space at rate T, however in small samples the approximation may still be poor. Hence, to investigate this issue we compute the Hausdorff distance between the true and the estimated cointegrating spaces. The Hausdorff distance is defined as follows: Let M and N denote two linear subspaces of $\mathbb{R}^s$, then the Hausdorff distance $d_H(M, N)$ is given by

$$d_H(M, N) = \max(\sup_{x \in M, \|x\|=1} \| (I - Q)x \|, \sup_{x \in N, \|x\|=1} \| (I - P)x \|)$$

where Q denotes the orthogonal projection onto N, P the orthogonal projection onto M and $\| x \|$ denotes the Euclidean norm on $\mathbb{R}^s$. The Hausdorff distance is between 0 and 1, and is equal to 1 for instance for spaces of different dimensions. The investigation concerning the Hausdorff distance is separated from the testing problem, i.e. the Hausdorff distance is computed between the correct number of estimated cointegrating vectors and the true cointegrating space. Note that the Hausdorff distance is equal to one for spaces of different dimensions. Four different Hausdorff distances are computed: The distance between the initial subspace estimate and the true cointegrating space, the distance between the adapted subspace estimate and the true cointegrating space, the distance between the VAR Johansen estimate and the true cointegrating space and the distance between the Bierens estimate and the true cointegrating space. The results are displayed in graphical form, we plot densities over the replications of the log Hausdorff distances. The logarithm is taken to increase variability, as due to consistency for all methods the estimated cointegrating spaces approach the true cointegrating spaces and hence the Hausdorff distances tend to 0.

Additional results concerning e.g. the minimizing values of the information criteria, the results of the order estimation, or descriptive statistics concerning the distributions of the Hausdorff distances are available upon request.[9] The `MATLAB` code used in producing these results can be obtained from the author. Note finally that in the simulations the processes are all started with zero initial values and that the results are based upon 5000 replications.

---

[9]In Bauer and Wagner (2002e) furthermore the forecasting performance of cointegrated systems estimated with the presented subspace algorithm is discussed.

	1	2	3	4	5	6	7	8
γ_1	-0.8	-0.5	-0.2	0	0.2	0.5	0.8	1
γ_2	-0.8	-0.5	-0.2	0	0.2	0.5	0.8	0.8
ρ_0	1.25	2	5	-	5	2	1.25	1

Table 2: Parameter values γ_1 and γ_2 used in the simulations. ρ_o denotes the minimum absolute value of the roots of $\det b(z)$.

3.1 2-dimensional systems

The first set of models is adopted from Hargreaves (1994) and consists of 2-dimensional ARMA(2,1) systems with one cointegrating vector given by $(1, -3)$.

$$
\begin{bmatrix} 1 & -2 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} = \begin{bmatrix} 1.5 & 0 \\ 0 & 0.5 \end{bmatrix} \begin{bmatrix} u_{1t-1} \\ u_{2t-1} \end{bmatrix} +
$$
$$
+ \begin{bmatrix} -0.5 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_{1t-2} \\ u_{2t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} + \begin{bmatrix} \gamma_1 & 0 \\ 0 & \gamma_2 \end{bmatrix} \begin{bmatrix} \varepsilon_{1t-1} \\ \varepsilon_{2t-1} \end{bmatrix} \tag{6}
$$

The parameter values for γ_i in the MA polynomials can be seen in Table 2. Note that increasing values of γ_i correspond to increasing autocorrelation in the MA errors. The ε_t are i.i.d. normally distributed with variance 2 and covariance $1/2$. The first point investigated is the test performance. In Table 3 the acceptance frequencies for the correct dimension of the cointegrating space are displayed for all systems and all sample sizes. There is strong evidence (mainly for the smaller sample sizes) that the test performance depends upon the correlation in the moving average errors of the process. For systems 1 to 8, with the corresponding parameters γ_i increasing monotonously, for $T = 100$ the following behavior occurs: The acceptance frequency of the subspace tests I to IV is increasing with γ_i, and is close to the 95 % from system 3 onwards. For systems 1 and 2 the subspace tests V and VI, replicating Johansen's procedure on the state equation, and the Johansen procedure applied on a VAR approximation have better performance, with the best performance given by the subspace Johansen type tests. From system 3 onwards, including the pure autoregressive system 4, the eigenvalue based subspace tests show the best performance. With larger positive values of γ_i the performance of the Johansen procedure and even more the performance of the subspace tests V and VI deteriorates. This effect is present not only for $T = 100$ but also for $T = 200$ and marginally for $T = 300$. It is worth noting that even for system 4, the pure AR process, the performance of the subspace tests I to IV is better than the Johansen results

System	Sample	Subspace Tests I to VI						Johansen Tests		Bierens
1	100	0.421	0.421	0.409	0.409	0.870	0.855	0.824	0.807	0.834
	200	0.533	0.533	0.521	0.521	0.908	0.897	0.883	0.862	0.935
	300	0.553	0.553	0.542	0.542	0.899	0.885	0.882	0.861	0.923
	400	0.579	0.579	0.566	0.566	0.898	0.888	0.896	0.877	0.924
	1000	0.750	0.750	0.742	0.742	0.953	0.945	0.906	0.893	0.928
2	100	0.884	0.884	0.880	0.880	0.990	0.994	0.903	0.894	0.606
	200	0.930	0.930	0.925	0.925	0.998	0.998	0.946	0.934	0.732
	300	0.928	0.928	0.922	0.922	1.000	1.000	0.943	0.931	0.808
	400	0.948	0.948	0.943	0.943	0.999	0.998	0.947	0.935	0.862
	1000	0.944	0.944	0.941	0.941	1.000	1.000	0.947	0.941	0.937
3	100	0.945	0.943	0.936	0.935	0.841	0.948	0.887	0.880	0.426
	200	0.946	0.946	0.941	0.941	0.995	0.997	0.930	0.927	0.561
	300	0.963	0.963	0.960	0.960	0.999	0.999	0.949	0.939	0.637
	400	0.964	0.964	0.963	0.963	0.999	0.999	0.955	0.942	0.707
	1000	0.950	0.950	0.943	0.943	0.999	0.999	0.952	0.937	0.866
4	100	0.963	0.964	0.951	0.950	0.714	0.873	0.869	0.858	0.364
	200	0.966	0.966	0.963	0.963	0.999	1.000	0.948	0.936	0.467
	300	0.977	0.977	0.970	0.970	0.997	0.997	0.944	0.937	0.567
	400	0.969	0.969	0.965	0.965	0.998	0.997	0.954	0.949	0.605
	1000	0.946	0.946	0.942	0.942	1.000	1.000	0.943	0.937	0.804
5	100	0.962	0.963	0.960	0.962	0.608	0.802	0.827	0.815	0.340
	200	0.968	0.968	0.965	0.965	0.993	0.996	0.935	0.925	0.431
	300	0.962	0.961	0.960	0.960	0.996	0.995	0.946	0.939	0.500
	400	0.961	0.961	0.957	0.957	0.999	0.999	0.951	0.943	0.525
	1000	0.948	0.948	0.943	0.943	0.999	0.999	0.949	0.939	0.733
6	100	0.946	0.956	0.944	0.952	0.509	0.702	0.725	0.719	0.282
	200	0.963	0.963	0.962	0.962	0.985	0.989	0.927	0.912	0.349
	300	0.957	0.958	0.957	0.958	0.987	0.990	0.955	0.948	0.425
	400	0.942	0.942	0.940	0.940	0.991	0.993	0.939	0.924	0.428
	1000	0.955	0.955	0.949	0.949	0.998	0.997	0.956	0.939	0.626
7	100	0.959	0.957	0.956	0.960	0.324	0.528	0.590	0.577	0.241
	200	0.964	0.965	0.959	0.961	0.959	0.973	0.877	0.876	0.269
	300	0.949	0.949	0.948	0.949	0.959	0.967	0.937	0.925	0.344
	400	0.947	0.948	0.947	0.947	0.969	0.978	0.949	0.936	0.388
	1000	0.955	0.955	0.950	0.950	0.989	0.993	0.955	0.948	0.545
8	100	0.955	0.955	0.954	0.960	0.258	0.464	0.562	0.553	0.232
	200	0.976	0.976	0.971	0.971	0.971	0.985	0.886	0.881	0.301
	300	0.971	0.969	0.966	0.966	0.960	0.963	0.937	0.928	0.329
	400	0.953	0.952	0.954	0.953	0.946	0.955	0.950	0.941	0.381
	1000	0.954	0.952	0.951	0.951	0.958	0.972	0.958	0.954	0.546

Table 3: Frequencies of correct decisions of the 9 discussed tests for cointegration for all 8 systems and all sample sizes $T = 100, \ldots, 400$ and 1000. The subspace tests are ordered as indicated in the text. The Johansen VAR based tests are ordered as first the *trace* and second the *max* test. The individual steps of the testing sequences are carried out at 5 % nominal size.

17

for $T = 100$.[10] The performance of Bierens' test is deteriorating strongly from system 1 to system 8. From system 3 or 4 onwards even for $T = 1000$ the results are quite bad, and the acceptance frequency is far off the 95 % value.

From the 2-dimensional systems we thus conclude that the subspace tests I to IV are outperforming the Johansen VAR results for systems with no or positive autocorrelation of the MA errors. Tests V and VI have deteriorating performance for increasing autocorrelation, a feature that is to a lesser extent shared by Johansen's procedure.

In Figures 1 and 2 we display the densities of the log Hausdorff distances between the true and the estimated cointegrating spaces. As indicated already above, four distances are computed. The results can be summarized as follows: Except for system 1 it turns out that the performance of the subspace estimates and the Johansen VAR estimate of the cointegrating space are of roughly the same quality. The Bierens estimates are worse throughout, and are therefore seen to be placed right of the other densities in the pictures. For system 1 the Johansen procedure dominates the other approaches and it is remarkable that for this system the initial subspace estimate of the cointegrating space is better than the adapted estimate. For the other systems the adapted estimate is only slightly better than the initial estimate, and for the bigger sample sizes there is basically no difference between the initial and the adapted estimates.

3.2 3-dimensional systems

Also 3-dimensional ARMA(2,1) processes with a 2-dimensional cointegrating space have been simulated:

$$
\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{bmatrix} = \begin{bmatrix} u_{1t} \\ u_{2t} \\ u_{3t} \end{bmatrix} = \begin{bmatrix} 0.8 & 0 & 0 \\ 0 & 1.2 & 0 \\ 0 & 0 & 1.5 \end{bmatrix} \begin{bmatrix} u_{1t-1} \\ u_{2t-1} \\ u_{3t-1} \end{bmatrix} +
$$

$$
+ \begin{bmatrix} 0 & 0 & 0 \\ 0 & -0.7 & 0 \\ 0 & 0 & -0.5 \end{bmatrix} \begin{bmatrix} u_{1t-2} \\ u_{2t-2} \\ u_{3t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix} + \begin{bmatrix} \gamma_1 & 0 & 0 \\ 0 & \gamma_2 & 0 \\ 0 & 0 & \gamma3 \end{bmatrix} \begin{bmatrix} \varepsilon_{1t-1} \\ \varepsilon_{2t-1} \\ \varepsilon_{3t-1} \end{bmatrix} \tag{7}
$$

Again the ε_t are i.i.d. normally distributed.[11] The cointegrating space is equal to the space of the spanned by the first two rows of the first matrix in the above equation. The parameters

---

[10]Note at this point that all test results are based on asymptotic tables, where no small sample correction has been performed for any of the tests.

[11]The covariance matrix is given by $\begin{pmatrix} 0.47 & 0.20 & 0.18 \\ 0.20 & 0.32 & 0.27 \\ 0.18 & 0.27 & 0.30 \end{pmatrix}$ and is taken from Saikkonen and Luukkonnen (1997).
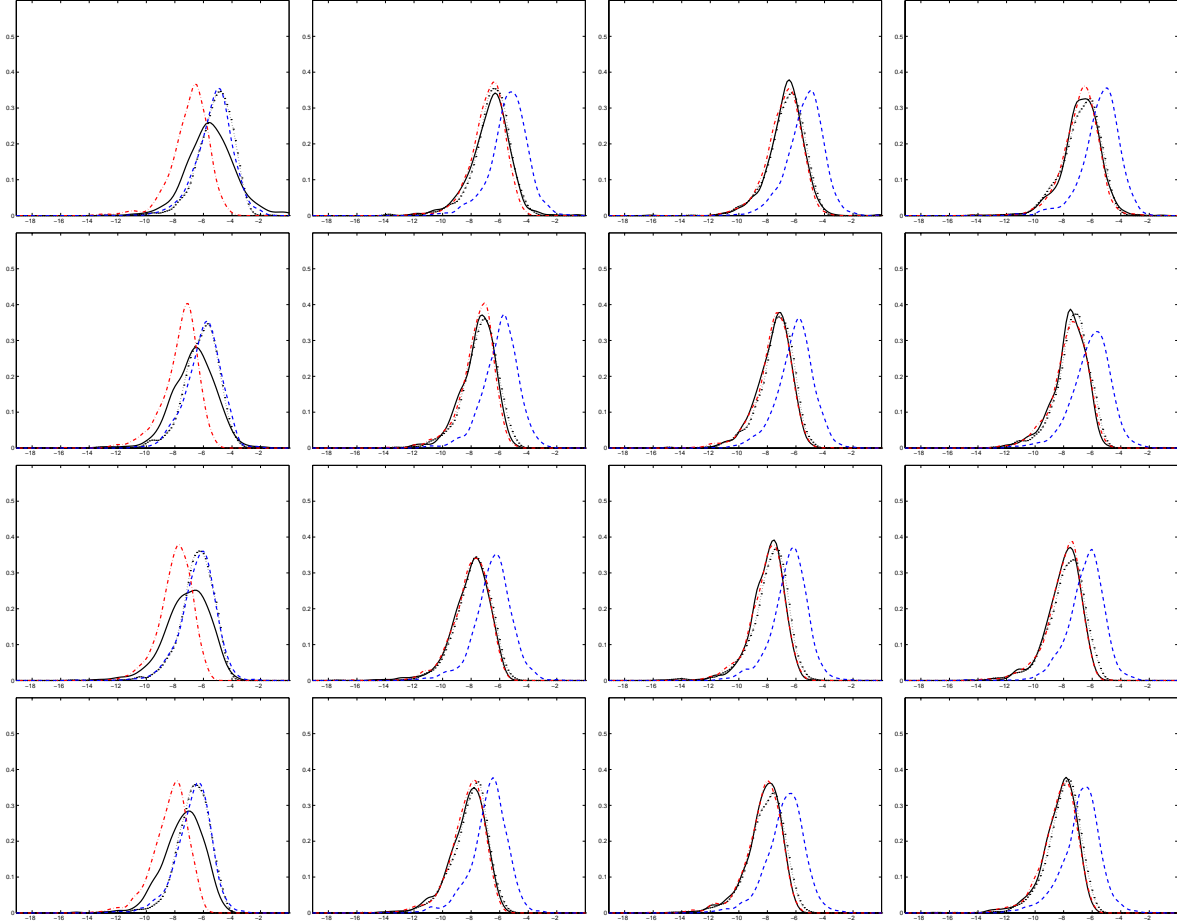
18

Figure 1: Density plots of Log Hausdorff distances between true and estimated cointegrating spaces for systems 1 to 4. Ordered from system 1 to 4 along columns and along rows for $T = 100, \ldots, 400$. The solid lines correspond to the *initial* subspace estimates, the dotted lines correspond to the *adapted* subspace estimates, the dash-dotted lines correspond to the Johansen estimates derived from a VAR approximation and the dashed lines correspond to the Bierens estimates.
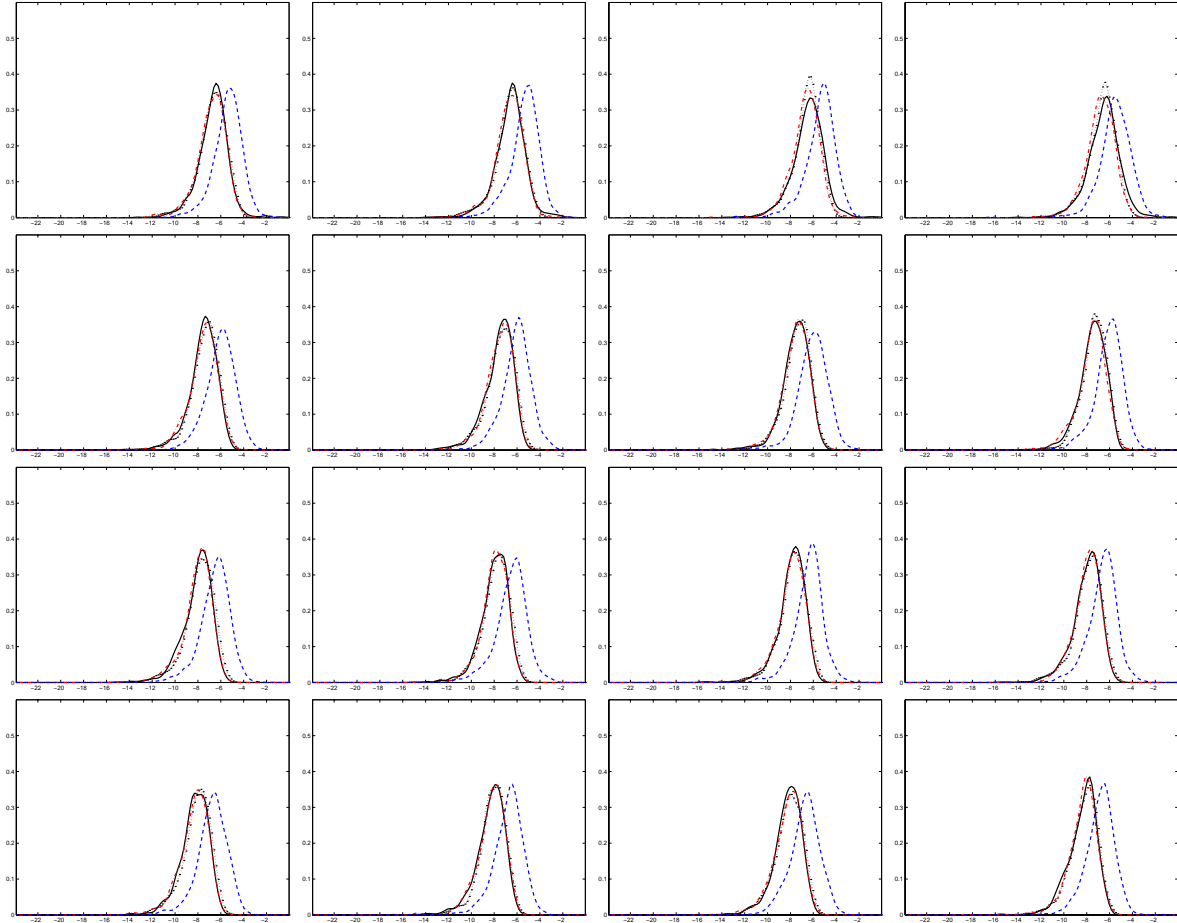
Figure 2: Density plots of Log Hausdorff distances between true and estimated cointegrating spaces for systems 5 to 8. Ordered from system 5 to 8 along columns and along rows for $T = 100, \ldots, 400$. The solid lines correspond to the *initial* subspace estimates, the dotted lines correspond to the *adapted* subspace estimates, the dash-dotted lines correspond to the Johansen estimates derived from a VAR approximation and the dashed lines correspond to the Bierens estimates.

	1	2	3	4	5	6
γ_1	-1	-0.5	0	0.6	0.8	1
γ_2	-0.9	-0.5	0	0.6	0.8	0.9
γ_3	-0.9	-0.5	0	0.6	0.8	0.9
ρ_0	1	2	-	1.66	1.25	1

Table 4: Parameter values γ_1, γ_2 and γ_3 used in the simulations. ρ_0 denotes the minimum absolute value of the roots of $\det b(z)$.

γ_i used in the simulations are given in Table 4.

The test results presented in Table 5 are comparable to the results obtained with the 2-dimensional processes. Look at $T = 100$ first: For system 1 with the most negative parameters γ_i, the Johansen procedure and the Johansen type subspace tests (V and VI) outperform the other tests. For system 2, the subspace tests I to VI all lead to correct decisions in around 65 % of the replications, whereas Johansen's tests lead to a correct decision in only about 55 % of the replications. From system 3 onwards for this sample size the subspace tests are inferior to the Johansen results, which remain at roughly the same level for the rest of the systems. For $T \geq 200$ all subspace tests as well as the Johansen tests exhibit good performance, with some variability across the tests for $T = 200$. The Bierens procedure on the other hand delivers inacceptable results even for $T = 1000$, with the exception of system 1, where for $T = 200, \ldots, 400$ in around 70 % of the replications the correct decision is made. Further investigation of the results of Bierens' procedure shows that across most of the replications a too small number of cointegrating relationships is decided for.

Hence we note that for these 3-dimensional systems, the test performance of the Johansen procedure on these VAR approximations is not outperformed by the subspace cointegration tests, and both approaches are significantly better than the Bierens method. The latter exhibits a strong tendency to deliver a too low dimensional cointegrating space.

The second issue investigated is again the quality of the estimation of the cointegrating space, see Figure 3. Again the Bierens procedure performs worse than either the subspace estimates of the Johansen VAR estimates for all systems and sample sizes. The other two methods yield again quite comparable results. It has to be noted that for the two systems with negative γ_i's, the Johansen procedure performs best. For systems 3 to 6 with $\gamma_i \geq 0$ the performance of the Johansen and the subspace algorithm estimates are again comparable, where in some cases a slightly better performance of the Johansen procedure is observed.

System	Sample	Subspace Tests I to VI						Johansen Tests		Bierens
1	100	0.757	0.758	0.746	0.746	0.861	0.944	0.893	0.882	0.378
	200	0.818	0.818	0.812	0.812	0.967	0.966	0.917	0.903	0.714
	300	0.869	0.869	0.864	0.864	0.982	0.982	0.928	0.916	0.702
	400	0.882	0.882	0.877	0.877	0.989	0.988	0.907	0.895	0.720
	1000	0.957	0.957	0.951	0.951	0.999	0.999	0.944	0.937	0.525
2	100	0.645	0.651	0.641	0.642	0.637	0.702	0.575	0.543	0.041
	200	0.863	0.864	0.857	0.856	0.892	0.945	0.928	0.911	0.043
	300	0.901	0.901	0.896	0.896	0.970	0.993	0.942	0.932	0.056
	400	0.930	0.930	0.923	0.923	0.999	1.000	0.939	0.931	0.047
	1000	0.946	0.946	0.941	0.941	1.000	1.000	0.955	0.944	0.145
3	100	0.722	0.740	0.758	0.762	0.454	0.634	0.883	0.870	0.026
	200	0.960	0.965	0.962	0.961	0.979	0.995	0.973	0.969	0.044
	300	0.977	0.977	0.976	0.976	0.997	0.996	0.977	0.972	0.069
	400	0.975	0.975	0.974	0.974	0.999	1.000	0.979	0.978	0.122
	1000	0.981	0.981	0.976	0.976	1.000	1.000	0.979	0.974	0.248
4	100	0.343	0.477	0.408	0.507	0.500	0.726	0.887	0.863	0.013
	200	0.817	0.834	0.817	0.822	0.999	0.998	0.938	0.924	0.034
	300	0.905	0.905	0.902	0.902	0.995	0.996	0.943	0.930	0.043
	400	0.924	0.923	0.914	0.914	0.995	0.999	0.942	0.934	0.048
	1000	0.947	0.947	0.943	0.943	1.000	0.999	0.951	0.939	0.121
5	100	0.417	0.594	0.492	0.640	0.232	0.396	0.754	0.712	0.018
	200	0.982	0.982	0.981	0.981	0.984	0.996	0.950	0.942	0.022
	300	0.992	0.992	0.992	0.992	0.995	0.994	0.982	0.978	0.035
	400	0.983	0.982	0.982	0.980	0.992	0.993	0.984	0.980	0.040
	1000	0.993	0.993	0.991	0.991	0.996	1.000	0.990	0.986	0.087
6	100	0.302	0.494	0.374	0.541	0.466	0.639	0.824	0.784	0.009
	200	0.960	0.967	0.956	0.960	0.985	0.997	0.950	0.940	0.020
	300	0.968	0.968	0.963	0.963	0.998	0.997	0.953	0.942	0.028
	400	0.942	0.942	0.939	0.939	0.997	0.998	0.956	0.947	0.031
	1000	0.931	0.932	0.923	0.924	0.984	0.988	0.931	0.923	0.094

Table 5: Frequencies of correct decisions of the 9 discussed tests for cointegration for all 6 systems and all sample sizes $T = 100, \ldots, 400$ and 1000. The subspace tests are ordered as indicated in the text. The Johansen VAR based tests are ordered as first the *trace* and second the *max* test. The individual steps of the testing sequences are carried out at 5 % nominal size.
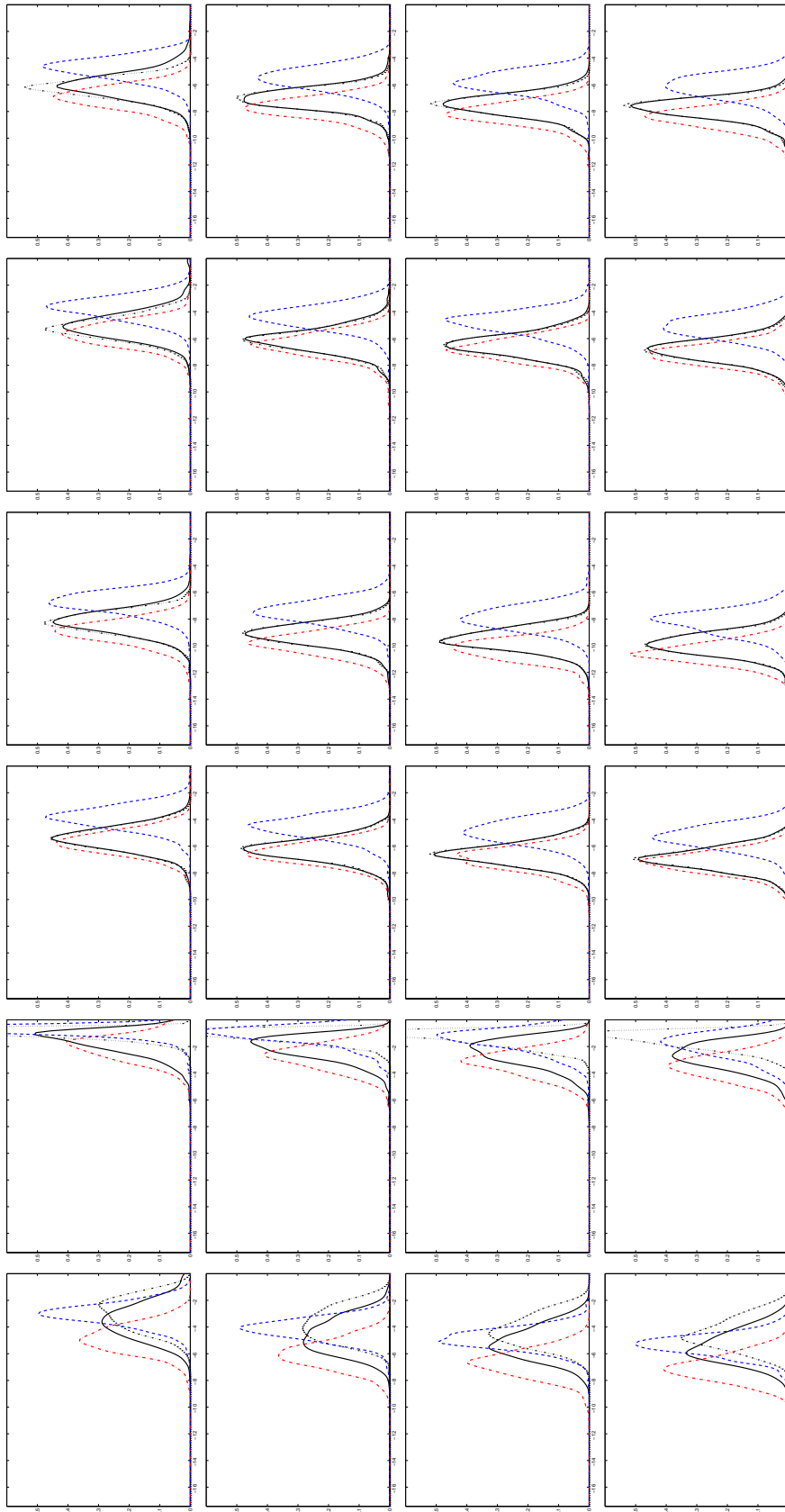
Figure 3: Density plots of Log Hausdorff distances between true and estimated cointegrating spaces for systems 1 to 6. Ordered from system 1 to 6 along columns and along rows for $T = 100, \ldots, 400$. The solid lines correspond to the *initial* subspace estimates, the dotted lines correspond to the *adapted* subspace estimates, the dash-dotted lines correspond to the Johansen estimates derived from a VAR approximation and the dashed lines correspond to the Bierens estimates.

The results are different for the systems with negative coefficients γ_i and especially for system 1 with a unit root in the moving average polynomial. For these systems, the subspace estimates are worse than the Johansen estimate. One reason for this may of course be the fact that the subspace algorithm is not designed for systems with unit roots in the moving average polynomial, as the strict minimum phase assumption is used in the consistency proofs. Hence, some problems could have been expected for the systems with a (near) unit root in the moving average polynomial. It is however remarkable that the tests, based on the eigenvalues of the estimated system matrix $\hat{A}$, seem to be unaffected.

4   An Application: The Neoclassical Growth Model

In this section we apply the discussed methods for cointegration analysis to investigate the long-run properties of the Austrian economy through the mirror of the (stochastic) one-sector neoclassical growth model.[12] The one-sector neoclassical growth model is one of the most widely used models in empirical macroeconomics, see King, Plosser and Rebelo (1988) for an early discussion. This exercise serves two purposes: The first is to see whether the results obtained by applying the different methods are as versatile as those obtained on simulated data, or maybe even more versatile. The second point, mainly relegated to the appendix is to re-assess in detail the long-run growth properties of the Austrian economy. This updates in a certain sense an earlier study undertaken by Kunst and Neusser (1990). In the Appendix additional empirical results, concerning the data, the cointegrating space, and the stability of the results are collected. In a separate subsection we focus in a bivariate analysis on the consumption-output, or consumption-GDP, relationship, which is another prime candidate for a cointegrating relationship.

As in Kunst and Neusser (1990) the set of variables consists of the following (seasonally adjusted) real variables: GDP, labelled also output, private consumption, gross investment, exports, total wage sum and the (ex-post) real interest rate. The frequency of observations is quarterly and the time range is from 1977(IV) to 1998(III). For this set of variables, the underlying model has some clear predictions concerning cointegration: Although (the logarithm of) technical progress is assumed to be an I(1) process, the log consumption-output ratio (CR), the log investment-output ratio (IR), the log wage sum-output ratio (WR) and the real

---

[12]The computations reported in this section and in the appendix have been performed using `MATLAB`, `EVIEWS`, `EASYREG` and `CATS`.

interest rate are predicted to be stationary on theoretical grounds. Exports are included in the variable set to account for the fact that Austria is a small open economy, with export shares in the 1990ies at above 50 % of GDP.[13] The first step in the analysis is to perform unit root tests on the individual series. These lead us to conclude that all series except for the real interest rate can be sufficiently well described as random walks (integrated of order 1) with drifts, see also Figure 6 in the appendix.

Before proceeding to a multivariate cointegration analysis, let us first inspect the *great ratios*, i.e. CR, IR and WR, predicted to be stationary by the theory, graphically in Figure 4. Graphical inspection leads us to tentatively conclude stationarity of *CR*, nonstationarity of
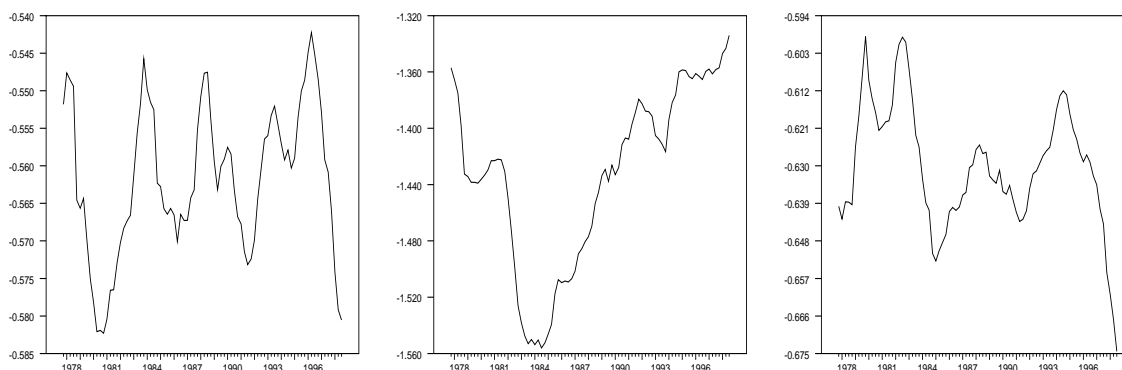
Figure 4: Starting from left this figure displays the log consumption-output ratio (CR), the log investment-output ratio (IR) and the log wage sum-output ratio (WR).

IR and an unclear picture for *WR*. Also for the real interest rate, see again Figure 6 in the Appendix, the decision is not completely clear. Graphical analysis is followed by statistical analysis, of course: Unit root tests are performed on these great ratios (and on the real interest rate), as a direct device of testing for cointegration. The results of these tests are quite in line with the graphical impression: Stationarity of CR, nonstationarity of IR and both WR and the real interest rate are borderline cases, where the test results are sensitive with respect to the specification of the test equation.

Concerning the investment-output ratio, during the first half of the 1980ies a strong decline in investment activity can be observed, most likely triggered by the oil-shock in 1979, 1980. Then, from about 1985 to 1991 investment expanded. During the years 1991 to 1993 the

---

[13]King et al. (1991) augment the basic model to include money and prices. In that context the additional cointegrating relationships that are predicted by the theoretical model are a cointegrating relationship between nominal interest rates and inflation and a cointegrating relationship between real balances, output and nominal interest rates.

Thresh.	I	II	III	IV	V	VI	Trace	Max	Bierens
3	0	0	0	0	3	3	3	2	0/3

Table 6: Test results for the dimension of the cointegrating space in the 6 variable system. For Bierens' method both the test and the estimation result are presented.

effects of the last recession observed in Austria can be seen clearly.

The fall of WR since 1995 has already come to the attention of Austrian macroeconometric modellers, and causes problems in finding a stable wage equation. Since about 1995 wage growth is smaller than GDP growth in Austria, reflecting a possible change in the Austrian wage setting behavior or a change of other institutional characteristics present on the Austrian labor market. However, structural stability tests lead to the conclusion of no structural change, this may stem from the fact that the possible changepoint is at the very end of the sample, thus the power of these tests is rather low.

We are now ready to apply the discussed systems cointegration analysis methods. For the Johansen procedure information criteria lead to an optimal lag length of 2 for a VAR model and for the deterministic part an unrestricted intercept vector turns out to be the most appropriate choice (after testing for the appropriate deterministic components). The results concerning the dimension of the cointegrating space delivered by the nine different tests and two estimates from the three different methods are given in Table 6. From the results one observation is quite striking: The subspace algorithm tests I to IV and the Bierens test all lead to the conclusion of no cointegration, i.e. of 6 random walks driving the system. The other tests, subspace V and VI, the Johansen trace test, as well as the subspace algorithm estimate (labelled Thresh. in Table 6) and the Bierens estimate of the dimension of the cointegrating space are equal to 3 and Johansen's max test leads to a conclusion of a 2-dimensional cointegrating space. A result of a 3-dimensional cointegrating space is in line with the univariate investigations, no cointegration at all seems to be highly implausible. Thus, we first conclude that the results obtained on real world data depend quite strongly on the method used. We also conclude that for the small sample size and 6 variables the imprecise estimation of the system matrices and hence the eigenvalues of A detrimentally influences the properties of the subspace tests, at least for the eigenvalue based tests I to IV.

But not only the test results are quite differing between the methods, the differences become even more striking when one looks at the estimated 3-dimensional cointegrating spaces

	Theor.	Joh.	Bie.	Sub.	Adap.
Theor.	–	0.999	0.999	0.957	1
Joh.	-0.001	–	0.807	0.965	0.985
Bie.	-0.001	-0.215	–	0.779	0.850
Sub.	-0.044	-0.035	-0.250	–	0.324
Adap.	0	-0.015	-0.016	-1.126	–

Table 7: Pairwise Hausdorff distances between the 3-dimensional theoretically predicted cointegrating space spanned by CR, WR and the real interest rate and the 3-dimensional cointegrating spaces estimated with the different methods. The upper half of the table displays the Hausdorff distances, the lower half displays the logarithms of the Hausdorff distances.

themselves. In particular we compute in Table 7 the Hausdorff (and log Hausdorff) distances between the estimates obtained from the different methods and also the distance to the 3-dimensional space spanned by CR, WR and the real interest rate. This 3-dimensional space is the theoretical space that is, from the univariate investigations, the most plausible candidate to compare with, but the results are essentially unchanged if any other 3-dimensional subspace from the theoretically predicted 4-dimensional cointegrating space is used in computing the distances. The results are rather disillusioning. Not only are the distances to the theoretically predicted space almost one, which might as well indicate a non-adequacy of the one-sector neoclassical growth model to describe the Austrian economy, the results are also extremely different between the methods. The Hausdorff distances differ between 0.324 and 0.985 for the different methods. Noting again that one is the upper bound for the Hausdorff distance, the variability of the results becomes evident, and raises some doubt about the approximation quality of cointegrating spaces for the sample sizes (about 100 observations) and system dimensions (up to 5 or 6 variables) usually at hand in macroeconometric applications.

4.1   The Consumption-Output Relationship

In this subsection we take a look at the consumption-output (GDP) relationship, see also Figure 5 below. This relationships is not only from the above discussed neoclassical growth model a prime candidate for cointegration, also the *permanent income hypothesis* predicts a long-run relationship between income and consumption. Of course GDP serves only as a proxy for disposable income, which itself only serves as a proxy for the unobservable permanent income. We investigate this bivariate example also to see the effect of the dimensionality of the system on the differences in the outcome obtained from the employed methods and

Thresh.	I	II	III	IV	V	VI	Trace	Max	Bierens
1	1	1	1	1	1	1	1	1	2/2

Table 8: Test results for the dimension of the cointegrating space in the 2 variable system GDP and consumption. For Bierens' method both the test and the estimation result are presented.
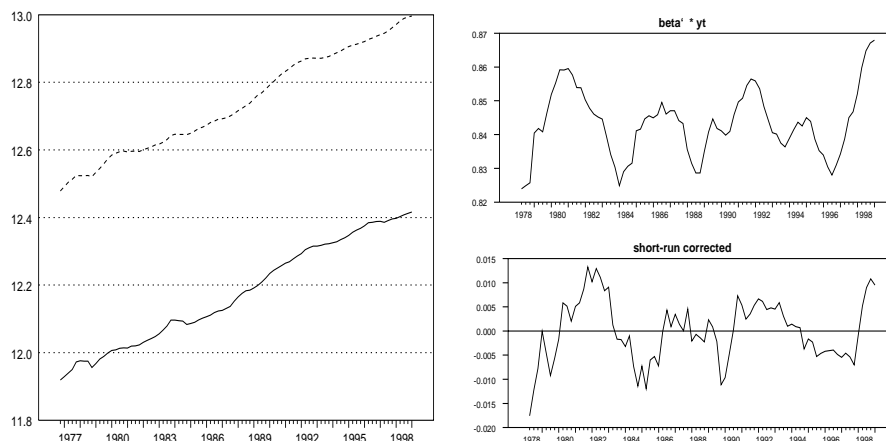

Figure 5: The left plot shows log output (dotted line) and log consumption (solid line). The right plot shows the estimated cointegrating relationship between the two obtained by applying the Johansen methodology. The upper graph displays the actual disequilibrium ($\beta' y_t$ in the notation of Section 2) and the lower graph displays the disequilibrium corrected for the short-run effects, i.e. for the lagged differences.

to check whether in this low-dimensional example the differences are less pronounced. The results concerning the dimension of the cointegrating space are presented in Table 8. The VAR order is again equal to 2 and again an unconstrained constant is included in the VAR model. The order estimated for the state space model is also equal to 2. All tests and the threshold estimate of the dimension of the cointegrating space except for Bierens' test lead to the conclusion of the existence of a cointegrating relationship between the two variables. The result obtained from Bierens' method is implausible, with both the estimate and the test leading to a conclusion of stationarity of both consumption and output. This stands in clear contrast with the left picture in Figure 5. In the right picture of Figure 5 we display the time series of the estimated cointegrating relationship obtained by applying the Johansen procedure. With the variables ordered as GDP and consumption, the estimated cointegrating vector is given by $(1, -0.975)$. Quite similar results emerge for the subspace estimates of the cointegrating relationship, which both coincide to be $(1, -1.0461)$. Bierens' estimate – for

	$(1,-1)$	Joh.	Bie.	Sub./Adap.
$(1,-1)$	–	0.013	0.206	0.023
Joh.	-4.374	–	0.218	0.035
Bie.	-1.580	-1.522	–	0.184
Sub.	-3.794	-3.350	-1.694	–

Table 9: Pairwise Hausdorff distances between the theoretically predicted cointegrating space $(1,-1)$ and the 1-dimensional cointegrating spaces estimated with the different methods. The upper half of the table displays the Hausdorff distances, and the lower half the logarithms of the Hausdorff distances.

a 1-dimensional cointegrating space - is far off these estimates and given by $(1,-1.533)$. In Table 9 we display again the Hausdorff distances between the different estimates and also to the theoretically expected cointegrating vector $(1,-1)$. Compared to the 6-dimensional example, here the results are much closer to each other and all results are also closer to the theoretically specified cointegrating space $(1,-1)$. Hence, as expected the dimensionality of the problem has a big effect on the results obtained with the different methods and it has also an effect on the extent to which the results obtained with the discussed methods differ.

5   Summary and Conclusions

In this paper we compare, both by means of simulations of ARMA processes and by a real world application, the performance of the Johansen VAR method for cointegration analysis with the performance of two methods developed to perform cointegration analysis for ARMA processes. The simulations are performed to assess the finite sample robustness of the Johansen procedure with respect to ARMA data generating processes, which in turn allows to assess the possibility to back-up results obtained by applying Johansen's method with results obtained from methods that are designed for ARMA data generating processes. These two other methods are the nonparametric cointegration analysis of Bierens (1995, 1997a) and the recently developed subspace algorithm cointegration analysis of Bauer and Wagner (2002a). The subspace algorithm cointegration analysis is described in some detail, since it differs from the usual approaches to cointegration analysis in two main aspects. Firstly, it is formulated in the state space framework, which is an equivalent alternative to represent ARMA processes. Secondly, it uses a modified version of a so called subspace algorithm. Subspace algorithms are computationally cheap methods that enjoy increasing popularity in the engineering and control literature. Since, both the state space framework and especially subspace algorithms

are not widely usage in the cointegration literature, these aspects are described in some detail in Section 2, where the applied methods are discussed.

Let us first discuss the evidence gathered via simulations. The simulations are performed with 2- and 3-dimensional ARMA(2,1) processes for $100, \ldots, 400$ and 1000 observations. The clearest message that emerges from the simulations is that Bierens' nonparametric cointegration analysis performs significantly worse than both Johansen's method and the subspace algorithm cointegration analysis. The correct decision frequencies are lower than for the other two methods, and the Hausdorff distances of the cointegrating spaces estimated by Bierens' procedure to the true cointegrating spaces are larger than for the other two methods. The Johansen VAR procedure and the subspace procedure show relatively comparable behavior. For the 2-dimensional systems and small sample sizes the best performing subspace tests outperform the Johansen test results, for the 3-dimensional systems no small sample gain in the test performance is obtained by applying the subspace procedure. Concerning the Hausdorff distances, the results are quite similar for the two methods. This however is to a certain extent expected, as for ARMA processes the cointegrating space depends only upon the autoregressive part, see e.g. Deistler and Wagner (2000). Summarizing, we conclude that the subspace procedure performs equally well and in some cases better than the Johansen procedure on the simulated ARMA(2,1) systems. These results can only be seen as a first investigation and more evidence on a larger variety of systems has to be gathered to better understand the relative performance of the two methods.

Let us finally turn to the cointegration analysis of the neoclassical growth model. Two messages can be reported: First of all, the results obtained by the different methods vary significantly for the system with 6 variables and about 100 observations. This versatility occurs both with respect to the test results for cointegration and also with respect to the estimated cointegrating spaces. In the bivariate application of the methods to the consumption-output relationship more similar results are found across methods. This sensitivity of results with respect to the applied method should be taken more into account than is current practice. The second message to be reported is that the neoclassical growth model does not fully describe the long-run growth properties of the Austrian economy. The dimension of the estimated cointegrating space does not correspond to the theoretical prediction and also the theoretically specified cointegrating relationships are almost all not contained in the estimated cointegrating space. These result confirm to a certain extent previous findings. It is in this context

important to note that the recursive analysis and stability testing for the cointegrating space does not lead to find any tendency that the model becomes a more apt description of the Austrian economy's growth process over the past decade either.

Appendix: Additional Empirical Results

In this appendix we present additional empirical results concerning three items. The first is a discussion of preprocessing the seasonally fluctuating raw data. Here we again discuss also univariate unit root testing. The second is an investigation of the stability of the estimated cointegrating rank and space and the third is hypothesis testing on the cointegrating space. The latter two items are mainly discussed in terms of the Johansen methodology, as this is the most developed method. Compared to that, e.g. for the subspace algorithm cointegration analysis no hypothesis testing on the cointegrating space at all is available so far.

Raw Data, Seasonality and Unit Roots

The Austrian Institute for Economic Research provides real quarterly data classified according to SNA79 for 1977(I) to 1998(III) at 1983 prices. The data are not seasonally adjusted and exhibit clear seasonal variation, most strongly present in the investment series.[14] In order to be able to perform a standard I(1) cointegration analysis, hence the first issue to investigate is the presence of stochastic seasonality. The test proposed by Hylleberg et al. (1990) is applied. The results can be summarized as follows: For the logs of GDP, consumption, investment and wages all 4 seasonal unit roots seem to be present. For the log of exports, the roots at ± 1 cannot be rejected, and for the real interest rate only the root at -1 cannot be rejected. To avoid problems of analyzing series that have been filtered differently we decide to de-seasonalize all 6 series by applying the filter $S(z) = (1 + z + z^2 + z^3)/4$, although strictly speaking only the 4 first mentioned series need to be filtered or *annually averaged* in this way to extract all seasonal unit roots present. The de-seasonalized series are displayed in Figure 6. The applied filter extracts (almost by construction) the observed seasonality reasonably well. The seasonally adjusted series are rather smooth and show a clear trending behavior. The only exception is the real interest rate, where, as expected, no trend seems to be visible, although from about 1980 to 1993 the level of the series seems to be higher than before and afterwards.

The usual battery of unit root tests has been applied to the seasonally differenced series, like augmented Dickey-Fuller, Phillips-Perron, Schmidt-Phillips, the KPSS test with the null of stationarity and also the non-parametric cointegration test of Bierens applied to each series individually. The detailed results are available upon request. The general picture that emerges confirms the quite clear graphical evidence: All series except for the real interest rate seem to be reasonably well described as integrated processes of order one with drift. The real interest rate, theoretically predicted to be stationary, turns out to be a borderline case in the statistical analysis, where the results are sensitive with respect to the precise specification of the test equation.

Stability of Cointegrating Rank and Space

In Section 4 it has been found that the cointegrating space is likely to be of dimension 3. Hence, the full predictions of the neoclassical growth model of a 4-dimensional cointegrating space are not supported. This finding is in line with previous studies of this model using Austrian data, the already mentioned work by Kunst and Neusser (1990) and of Neusser (1991). Neusser (1991) argues, using data from 1964(I) to 1989(IV), that the missing cointegrating relationship is possibly due to the fact that his sample contains several years of reconstruction after World War II. However, we obtain the same result also for our data starting only in 1977. The line of thought of some underlying disturbance

---

[14]The raw data as well as all testing results are available from the author upon request.
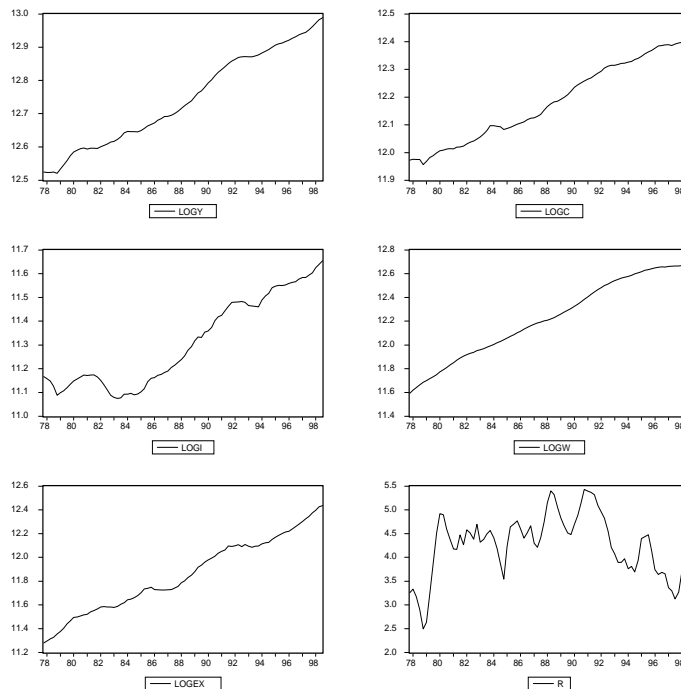
Figure 6: The seasonally adjusted Austrian series, the first 5 series are in logs.

hindering the emergence of underlying long-run relationships might be extended to later periods by noting that the Austrian economy was substantially regulated in certain sectors until a decade or so ago. To investigate this question, a recursive analysis has been performed, in order to see whether the dimension of the cointegrating space changes over time, with a possible tendency to increase. Due to the relatively small sample we start this recursive estimation only in 1993(III), i.e. about a year before Austria becomes a member of the European Union.

The recursive analysis is presented in the context of the VAR model and is based on the Johansen cointegration analysis method. The results are graphically displayed in Figure 7. The different lines are corresponding from top to bottom to the hypotheses that $dim(\beta) \leq i$ for $i = 0, \ldots, 5$. If a line is above the horizontal line at 1, this means that the corresponding hypothesis is rejected. The two graphics corresponding to the so called *Z-model* (left) and the so called *R-model* (right) differ with respect to the estimation of the short-run parameters. In the *Z-model* all the parameters are re-estimated for each sample, whereas in the *R-model* the short-run parameters are considered fixed throughout, and are only estimated for the full sample (see MALCOLM, Mosconi, 1998 for a description). The results are slightly different for the two versions. In the left picture we see that during the years 1996 and 1997 the hypothesis that the cointegrating space is 2-dimensional is at the edge of being accepted during several quarters. But, generally, for the last couple of years there is no tendency for the dimension of the cointegrating space to increase. The *R-model* suggests this increase in dimension, from a 2-to a 3-dimensional space only emerging at the end of the sample period. Hence also in later periods and in a sample that starts later than Neusser's we do not find evidence for the 4-dimensional cointegrating space predicted by the neoclassical growth model.

A second hypothesis investigated within the cointegrated VAR model is the possibility of a changing cointegrating space. In Figure 8 we display the results of the recursive estimation of the spanned cointegrating space. Here for both, the *R-* and the *Z-model* the stability of the space spanned by the estimated cointegrating space cannot be rejected. Again the lines in this figure are normalized in such a way that the critical value is 1, so that the hypothesis of stability is rejected if the lines reporting
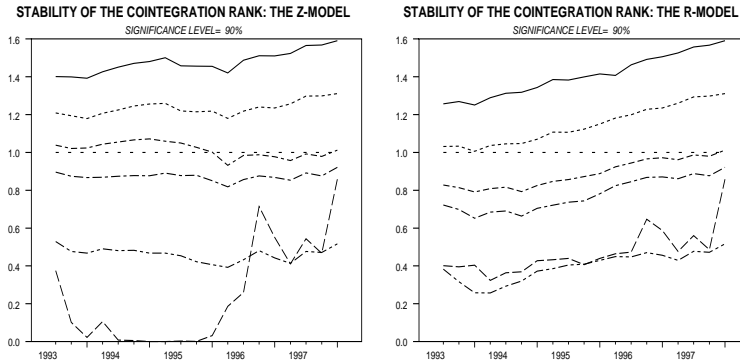
32

Figure 7: Stability of the dimension of the cointegrating space. The left plot corresponds to the *Z-model* and the right plot corresponds to the *R-model*.
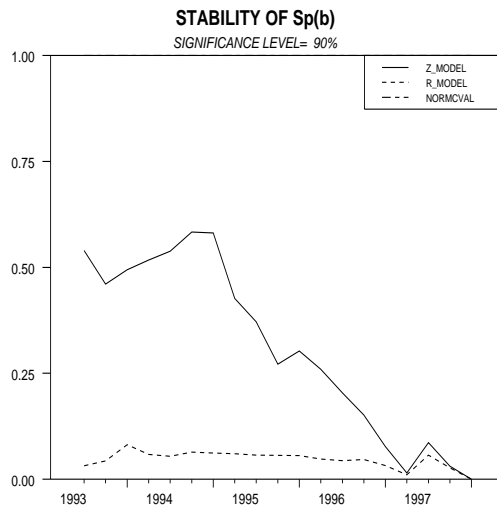

Figure 8: Stability of the cointegrating space.

Null hypothesis	Distribution	Test statistic
$CR \in sp(\beta)$	χ_3^2	9.10^*
$IR \in sp(\beta)$	χ_3^2	15.25^*
$WR \in sp(\beta)$	χ_3^2	5.46
$R \in sp(\beta)$	χ_3^2	4.29
$\{CR,\ IR\} \in sp(\beta)$	χ_6^2	42.73^*
$\{CR,\ WR\} \in sp(\beta)$	χ_6^2	31.86^*
$\{CR,\ R\} \in sp(\beta)$	χ_6^2	16.17^*
$\{IR,\ WR\} \in sp(\beta)$	χ_6^2	32.22^*
$\{IR,\ R\} \in sp(\beta)$	χ_6^2	18.64^*
$\{WR,\ R\} \in sp(\beta)$	χ_6^2	12.44
$\{CR,\ IR,\ WR\} \in sp(\beta)$	χ_9^2	56.87^*
$\{CR,\ IR,\ R\} \in sp(\beta)$	χ_9^2	52.75^*
$\{IR,\ WR,\ R\} \in sp(\beta)$	χ_9^2	47.25^*
$sp(\beta) \in \{CR,\ IR,\ WR,\ R\}$	χ_6^2	20.24^*

Table 10: Results of testing hypotheses on the cointegrating space. * denotes rejection of the null hypothesis at the 10 % critical level.

the (recursive) test statistics have values larger than 1.

There is thus no evidence in favor of instabilities concerning the cointegrating space, neither with regard to the dimension, nor with regard to the subspace of the $\mathbb{R}^6$ that the cointegrating space spans. As indicated above, the validity of these arguments is limited by the small sample size on which the recursive analysis is performed on. Further ongoing work will use other more elaborated methods to detect instabilities in the cointegrating space. Tests to detect a structural change in the cointegrating space are developed in Seo (1998), and tests for a continuous (smooth transition type) structural change in the cointegration space are presented by Ripatti and Saikkonen (1998).

Testing Hypotheses on the Cointegrating Space

The final issue is testing hypotheses on the cointegrating space. Again the results are reported for the Johansen method, as for the Bierens method hypothesis testing is less developed (and all formulated hypotheses have been rejected) and no hypothesis testing is developed for subspace algorithm cointegration analysis.

The tested hypotheses are whether any of the great ratios CR, IR, WR and the real interest rate, or any pair or triple of these is contained in the estimated 3-dimensional cointegrating space. Furthermore also the hypothesis that the estimated 3-dimensional cointegrating space is contained in the theoretically predicted 4-dimensional space is tested. See Table 10 for the results. The results are not always in line with the univariate results presented in the main text, e.g. it is surprising to see the hypothesis that CR is contained in the cointegrating space rejected. After all, CR is the variable that is most clearly found to be stationary if tested univariately. Rejection of IR is in accordance with the univariate results. Two theoretical relationships are found to be contained in the estimated cointegrating space, both individually and jointly, the wage sum output ratio and the real interest rate. Hence, there is at least partial concordance between the univariate results, the systems results and the theoretical predictions.

Two messages emerge: Firstly, there are substantial differences in the results obtained by single equation and systems approaches, which is a well known issue in econometrics. However, also the results obtained by different (multivariate) methods differ to a great extent. Secondly, the applied methods lead us to conclude that the long-run predictions of the neoclassical growth model are only partly

34

reflected in the Austrian data. We observe no changes in that respect due to the deregulations and Austria's joining of the European Union.

References

Aoki, M., 1990. State Space Modelling of Time Series (Springer, New York).

Bauer, D., Wagner, M., 2000. Subspace Algorithm Cointegration Analysis - An Application to Interest Rate Data, in: Nunez-Anton, V., Ferreira, E. eds., Proceedings of the 15th International Workshop in Statistics, 146–151.

Bauer, D., Wagner, M., 2002a. Estimating Cointegrated Systems Using Subspace Algorithms, Journal of Econometrics 111, 47–84.

Bauer, D., Wagner, M., 2002b. A Canonical Form for Unit Root Processes in the State Space Framework. Mimeo.

Bauer, D., Wagner, M., 2002c. Asymptotic Properties of Pseudo Maximum Likelihood Estimates for Multiple Frequency I(1) Processes. Mimeo.

Bauer, D., Wagner, M., 2002d. On Johansen Type Tests in Subspace Algorithm Cointegration Analysis. Mimeo.

Bauer, D., Wagner, M., 2002e. The Performance of Subspace Algorithm Cointegration Analysis: A Simulation Study with New Tests. Mimeo.

Bierens, H.J., 1995. Nonparametric Cointegration Analysis, CentER Discussion Paper, No. 95123, Tilburg.

Bierens, H.J., 1997a. Nonparametric Cointegration Analysis, Journal of Econometrics 77, 379–404.

Bierens, H.J., 1997b. Cointegration Analysis, in: Heij, C., Schumacher, H., Hanzon, B., Praagman, K. eds., System Dynamics in Economic and Financial Models (Wiley, Chichester), 217–238.

Boswijk, P., Lucas, A., Taylor, N., 1999. A Comparison of Parametric, Semi-nonparametric, Adaptive, and Nonparametric Cointegration Tests. Discussion Paper TI99-012/4, Tinbergen Institute, Amsterdam.

Deistler, M., Wagner, M., 2000. On the Structure of Cointegration, 373 – 379, in Dockner, E. et al. eds., Optimization, Dynamics and Economic Analysis (Springer, New York), 373–379

Hannan, E., Deistler, M., 1988. The Statistical Theory of Linear Systems (Wiley, New York).

Hargreaves, C., 1994. A Review of Methods of Estimating Cointegrating Relationships, in: Hargreaves, C. ed., Nonstationary Time Series Analysis and Cointegration (Oxford University Press, Oxford), 87–131.

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

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

King, R.G., Plosser, C.I., Rebelo, S., 1988. Production, Growth and Business Cycles II: New Directions, Journal of Monetary Economics 21, 309–341.

King, R.G., Plosser, C.I., Stock, J.H., Watson, M., 1991. Stochastic Trends and Economic Fluctuations, American Economic Review 81, 819–840.

Kunst, R., Neusser, K., 1990. Cointegration in a Macroeconomic System, Journal of Applied Econometrics 5, 351–365.

Lucas, A., 1998. Inference on Cointegrating Ranks Using LR and LM Tests Based on Pseudo-Likelihoods, Econometric Reviews 17, 185–214.

Mosconi, R., 1998. MALCOLM (Cafoscarina, Venezia).

Neusser, K., 1991. Testing the Long-Run Implications of the Neoclassical Growth Model, Journal of Monetary Economics 27, 3–37.

Phillips, P.C.B., 1991. Optimal Inference in Cointegrated Systems, Econometrica 59, 283–306.

Phillips, P.C.B., 1995. Fully Modified Least Squares and Vector Autoregression, Econometrica 63, 1023–1078.

Ripatti, A., Saikkonen, P., 1998. Cointegrated Vector Autoregressive Process with Continuous Structural Change, Bank of Finland Discussion Papers 29/98.

Said, S.E., Dickey, D.A., 1984. Testing for Unit Roots in Autoregressive- Moving Average Models of Unknown Order, Biometrika 71, 599–607.

Saikkonen, P., 1992. Estimation and Testing of Cointegrated Systems by an Autoregressive Approximation, Econometric Theory 8, 1–27.

Saikkonen, P., Luukkonen, R., 1997. Testing Cointegration in Infinite Order Vector Autoregressive Processes, Journal of Econometrics 81, 93–126.

Seo, B., 1998. Tests for Structural Change in Cointegrated Systems, Econometric Theory 14, 222–259.

Wagner, M., 1999. VAR Cointegration in VARMA Models, Economics Series 65, Institute for Advanced Studies, Vienna.

Yap, S.F., Reinsel, G.C., 1995. Estimating and Testing for Unit Roots in a Partially Nonstationary Vector Autoregressive Moving Average Model, Journal of the American Statistical Association 90, 253–267.