Trenkler, Carsten

Working Paper

Determining p-values for Systems Cointegration Tests With a Prior Adjustment for Deterministic Terms

Papers / Humboldt-Universität Berlin, Center for Applied Statistics and Economics (CASE), No. 2004,37

Provided in cooperation with:
Humboldt-Universität Berlin

Suggested citation: Trenkler, Carsten (2004) : Determining p-values for Systems Cointegration Tests With a Prior Adjustment for Deterministic Terms, Papers / Humboldt-Universität Berlin, Center for Applied Statistics and Economics (CASE), No. 2004,37, http://hdl.handle.net/10419/22210

Determining p-values for Systems Cointegration Tests With a Prior Adjustment for Deterministic Terms

by

Carsten Trenkler
Humboldt-Universität zu Berlin
School of Business and Economics
Institute of Statistics and Econometrics
Spandauer Str- 1
D-10178 berlin
Germany
Tel.: +49-30-2093-5705
Fax.: +49-30-2093-5712
email: trenkler@wiwi.hu-berlin.de

Abstract

In this paper I present a procedure to approximate the asymptotic distributions of systems cointegration tests with a prior adjustment for deterministic terms suggested by Lütkepohl, Saikkonen & Trenkler (2004), Saikkonen & Lütkepohl (2000a, 2000b, 2000c), and Saikkonen & Luukkonen (1997). The asymptotic distributions are approximated by the Gamma distribution and the parameters necessary to fit the Gamma distributions are obtained from response surfaces which I describe in this paper. The approximation can be easily used to derive arbitrary p-values or percentiles.

Keywords: p-values, systems cointegration tests, response surface
JEL classification: C15

*Part of the research for this paper has been done while I was a Jean Monnet Fellow at the European University Institute (EUI). I would like to thank the EUI for the award of the Fellowship and its hospitality. Moreover, I am grateful to Helmut Lütkepohl and Pennti Saikkonen for helpful comments and suggestions.

1 Introduction

In this paper I present a procedure to approximate the asymptotic distributions of systems cointegration tests with a prior adjustment for deterministic terms suggested by Lütkepohl et al. (2004), Saikkonen & Lütkepohl (2000a, 2000b, 2000c) and Saikkonen & Luukkonen (1997). The asymptotic distributions are approximated by the Gamma distribution and the parameters necessary to fit the Gamma distributions are obtained from response surfaces which I describe in this paper. The approximation can be easily used to derive arbitrary p-values or percentiles.

The mentioned cointegration tests differ with respect to the deterministic terms they allow for. The procedures of Lütkepohl et al. (2004) and Saikkonen & Lütkepohl (2000b) are the most general ones by taking account of shifts in the level of the time series. The specific feature in Lütkepohl et al. (2004) is that they allow for a level shift at unknown time. For both test setups there exist one test version which assumes a linear trend and another one which excludes it. The test by Saikkonen & Lütkepohl (2000c) is the corresponding procedure without level shifts. It originally incorporates a linear trend but can be adjusted in order to rule out a trend explicitly. A similar test with a mean term only is due to Saikkonen & Luukkonen (1997). Finally, Saikkonen & Lütkepohl (2000a) consider a linear trend that is orthogonal to the cointegration space. Furthermore, seasonal dummy variables can be incorporated into all procedures without changing the asymptotic results. The idea of all tests is to estimate the deterministic terms in a first step and to adjust the original time series by these estimated terms. Then, a likelihood ratio type test like in Johansen (1988) is applied to the adjusted data. The resulting asymptotic distributions are nonstandard and functions of Brownian motions. However, their percentiles can be obtained by simulation. The most recent and extensive set of percentiles is given in Trenkler (2003).

One important problem related to these tabulated sets is that they are restricted in terms of the number of percentiles which are covered. Accordingly, it is not possible to obtain p-values. However, p-values are preferable in certain simulation setups and empirical applications in order to assess the tests' outcome easily. The cointegrating rank tests based on Johansen (1988) are affected by the same problem. Their asymptotic distributions are also functions of Brownian motions. To compute p-values and percentiles for these asymptotic distributions MacKinnon, Haug & Michelis (1999) determine 221 percentiles by simulation

methods. Then, arbitrary p-values and percentiles can be derived by interpolation. However, the method of MacKinnon et al. (1999) involves the computation of more than 20,000 coefficients per test. A more efficient way is to follow a procedure suggested by Doornik (1998) and Johansen, Mosconi & Nielsen (2000). The basic idea is to approximate the asymptotic distributions of the Johansen type tests by Gamma distributions. These approximations work rather well as shown by Doornik (1998). Since the parameters of the Gamma distribution only depend on the first two moments it suffices to determine approximations of the mean and variance of the asymptotic test distribution. This can be easily done by using a response surface. The response surfaces derived in Doornik (1998) and Johansen et al. (2000) are based on less than 150 estimated coefficients. Having obtain an estimate for the mean and variance one can fit a Gamma distribution and compute arbitrary p-values or percentiles. Furthermore, the use of the response surfaces provides much better approximations to the asymptotic distributions than the standard simulations on which tables of critical values are usually based.

In line with the discussion I have computed response surfaces according to Doornik (1998) and Johansen et al. (2000) for the cointegration tests with a prior adjustment for deterministic terms. The approximations to the corresponding asymptotic distributions using the Gamma distribution are also very good for these tests including the asymptotic distribution of the test versions by Saikkonen & Lütkepohl (2000b) and Saikkonen & Lütkepohl (2000c) which consists of Brownian Bridges. It has turned out that the response surfaces following Doornik (1998) produce slightly better fits than the ones by Johansen et al. (2000).

The paper is organized as follows. The next section describes the model framework, the test procedures, and the limiting distributions. In Section 3 I explain how to use the Gamma distribution to approximate the asymptotic distributions. The response surfaces are presented in Section 4. Section 5 contains a comparison of different approaches to compute response surfaces and, finally, Section 6 summarizes and concludes.

2 Model Framework and Test Procedures

Let us consider a n-dimensional times series $y_t = (y_{1t}, \ldots, y_{nt})'$ $(t = 1, \ldots, T)$ which is generated by

$$y_t = \mu_t + x_t, \quad t = 1, 2, \ldots, \tag{2.1}$$

where μ_t contains the deterministic terms depending on the tests' assumptions. The term x_t is an unobservable stochastic error process which is assumed to follow a vector autoregressive process of order p (VAR(p)). The corresponding vector error correction model (VECM) has the form

$$\Delta x_t = \Pi x_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta x_{t-j} + \varepsilon_t, \quad t = 1, 2, \ldots, \tag{2.2}$$

where Π and Γ_j $(j = 1, \ldots, p-1)$ are $(n \times n)$ unknown parameter matrices. The error term ε_t is assumed to be a martingal sequence such that $E(\varepsilon_t | \varepsilon_s, s < t) = 0$, $E(\varepsilon_t \varepsilon_t' | \varepsilon_s,\ s < t) = \Omega$ is a non-stochastic positive definite matrix and the fourth moments are bounded. For the validity of the limiting distributions it suffices that the initial values x_t $(t = -p+1, \ldots, 0)$ have a fixed distribution which does not depend on the sample size. Furthermore, it is assumed that x_t is at most integrated of order one and cointegrated with a rank r implying the same properties for y_t. Moreover, it follows that the matrix Π can be written as $\Pi = \alpha\beta'$, where α and β are $(n \times r)$ matrices of full column rank. When determining the number of cointegration relations one tests for the rank of the matrix Π. I consider the trace and maximum eigenvalue test versions, i.e. the pairs of hypotheses

$$H_0(r_0) : \mathrm{rk}(\Pi) = r_0 \quad vs. \quad H_1(r_0) : \mathrm{rk}(\Pi) > r_0. \tag{2.3}$$

and

$$H_0(r_0) : \mathrm{rk}(\Pi) = r_0 \quad vs. \quad H_1(r_0) : \mathrm{rk}(\Pi) = r_0 + 1. \tag{2.4}$$

are tested respectively. Note that the maximum eigenvalue version has not been considered for all of the tests explicitly. One can derive, however, asymptotic distributions free of nuisance parameters for the relevant cases as discussed later on.

For the proposal of Saikkonen & Lütkepohl (2000b) we have $\mu_t = \delta d_t + \mu_0 + \mu_1 t$ allowing for a linear trend and assuming one level shift only. The shift is modelled by the dummy variable d_t which is one for all $t \geq T_1$ and zero otherwise where T_1 is the break point. The case of several level shifts can be treated in the same way by defining further shift dummies. Obviously, the remaining two quantities μ_0 and $\mu_1 t$ represent the mean and linear trend terms. The unknown $(n \times 1)$ parameter vectors δ, μ_0 and μ_1 are estimated by a GLS procedure. To obtain feasible GLS estimators $\hat{\delta}$, $\hat{\mu}_0$, and $\hat{\mu}_1$ the model (2.1) is transformed

accordingly by using first stage estimators from a reduced rank (RR) regression of

$$\Delta y_t = \nu + \alpha(\beta' y_{t-1} + \phi d_{t-1} + \tau(t-1)) + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + \sum_{j=0}^{p-1} \gamma_j \Delta d_{t-j} + \varepsilon_t, \quad t = p+1, p+2, \ldots, \tag{2.5}$$

with $\nu = -\Pi\mu_0 + (I_n - \sum_{j=1}^{p-1} \Gamma_j)\mu_1$, $\phi = -\beta'\delta$, $\tau = -\beta'\mu_1$, $\gamma_j = \delta$ for $j = 0$, and $\gamma_j = \Gamma_j\delta$ for $j = 1, \ldots, p-1$. This VECM model for y_t is derived from (2.1) using (2.2) and the aforementioned definition of μ_t. For the RR regression the rank r_0 specified under H_0 is applied since the transformation of (2.1) considers the structure of x_t under the null hypothesis. Having estimated the deterministic terms one can adjust y_t and compute $\hat{x}_t = y_t - \hat{\delta} d_t - \hat{\mu}_0 - \hat{\mu}_1 t$. Then, a Johansen-type test is performed on $\hat{x}_t$. Since $\hat{x}_t$ is adjusted, the test version in Johansen (1988) assuming no deterministic terms is applied. Hence, we have to solve a generalized eigenvalue problem. Using the resulting eigenvalues $\hat{\lambda}_1 \geq \cdots \geq \hat{\lambda}_n$ the trace test statistic for the pair of hypotheses in (2.3) is[1]

$$TR^{\delta}_{tr}(r_0) = -T \sum_{j=r_0+1}^{n} \log(1 - \hat{\lambda}_j) \tag{2.6}$$

and the maximum eigenvalue statistic regarding (2.4) is

$$ME^{\delta}_{tr}(r_0) = -T \log(1 - \hat{\lambda}_{r_0+1}). \tag{2.7}$$

If no linear trend is present one proceeds in the same way by making the necessary adjustments. In line with Saikkonen & Lütkepohl (2000b), μ_1 is set to zero, i.e. $\mu_t = \delta d_t + \mu_0$ in (2.1), and the intercept term in (2.5) is restricted to the cointegration relations since $\nu = \Pi\mu_0$ and $\tau = 0$ in this case. The resulting test statistics are denoted by $TR^{\delta}_{mean}(r_0)$ and $ME^{\delta}_{mean}(r_0)$.

Lütkepohl et al. (2004) use a similar setup as Saikkonen & Lütkepohl (2000b) but they make the assumption that the breakpoint is unknown. Lütkepohl et al. (2004) have proposed estimators to locate the shift date based on an unrestricted VAR model. Given the estimated break point the tests work as described above. Let us denote the corresponding test statistics as $TR^{un}_{tr}(r_0)$, $ME^{un}_{tr}(r_0)$, $TR^{un}_{mean}(r_0)$, and $ME^{un}_{mean}(r_0)$.

[1]Note that the generalized eigenvalue problem in Saikkonen & Lütkepohl (2000b) is formulated in a different way than in Johansen (1988). Therefore, the obtained eigenvalues and the specific form of the test statistic differ. However, the two kinds of eigenvalue problems can be transformed into each other by an appropriate redefinition of the respective eigenvalues. Thus, the test statistics based on the two different sets of eigenvalues are identical apart from minor numerical differences.

Table 1. Summary of Test Statistics

Deterministic Terms	Test Statistics	References
Linear Trend (μ_1 arbitrary)		
Level Shift (δ arbitrary)		
Unknown Shift Date	$TR^{un}_{tr}(r_0), ME^{un}_{tr}(r_0)$	Lütkepohl et al. (2004)
Known Shift Date	$TR^{\delta}_{tr}(r_0), ME^{\delta}_{tr}(r_0)$	Saikkonen & Lütkepohl (2000b)
No Level Shift ($\delta = 0$)	$TR_{tr}(r_0), ME_{tr}(r_0)$	Saikkonen & Lütkepohl (2000c)
Linear Trend orthogonal to cointegration space ($\mu_1 \neq 0$, $\beta'\mu_1 = 0$, $\delta = 0$)	$TR_{ort}(r_0), ME_{ort}(r_0)$	Saikkonen & Lütkepohl (2000a)
Mean Term only ($\mu_1 = 0$, μ_0 arbitrary)		
Level Shift (δ arbitrary)		
Unknown Shift Date	$TR^{un}_{mean}(r_0), ME^{un}_{mean}(r_0)$	Lütkepohl et al. (2004)
Known Shift Date	$TR^{\delta}_{mean}(r_0), ME^{\delta}_{mean}(r_0)$	Saikkonen & Lütkepohl (2000b)
No Level Shift ($\delta = 0$)	$TR_{mean}(r_0), ME_{mean}(r_0)$	Saikkonen & Lütkepohl (2000c)

In case of no level shifts one can set up cointegration tests as before by setting all terms associated with the level shift in (2.1) and (2.5) to zero. These tests are due to Saikkonen & Lütkepohl (2000c) and the test statistics are abbreviated as $TR_{tr}(r_0)$, $ME_{tr}(r_0)$, $TR_{mean}(r_0)$, and $ME_{mean}(r_0)$. For the situation of a mean term only, Saikkonen & Luukkonen (1997) have proposed to estimate μ_0 by a GLS procedure based on first-stage estimators from a VAR model for y_t imposing no rank restriction. Lütkepohl, Saikkonen & Trenkler (2001), however, have pointed out that using a GLS procedure as suggested by Saikkonen & Lütkepohl (2000c) results in better size properties in small samples. Note that the alternative way of estimating μ_0 does not change the asymptotic null distributions of $TR_{mean}(r_0)$ and $ME_{mean}(r_0)$.

The test version of Saikkonen & Lütkepohl (2000a) is similar to the one in Saikkonen & Lütkepohl (2000c) with a linear trend but considers the restriction that the linear trend is orthogonal to the cointegration space. Therefore, the restriction $\tau = \beta'\mu_1 = 0$ is imposed within the RR regression and secondly, the adjustment of the data occurs according to the

Table 2. Asymptotic Distribution of Test Statistics

Deterministic Terms	Test Statistics	Distributions
Linear Trend	$TR_{tr}^{un}(r_0),\ TR_{tr}^{\delta}(r_0),\ TR_{tr}(r_0)$	$\mathrm{tr}(D_{tr})$
(μ_1 arbitrary)	$ME_{tr}^{un}(r_0),\ ME_{tr}^{\delta}(r_0),\ ME_{tr}(r_0)$	$\lambda_{max}(D_{tr})$
Linear Trend orthogonal	$TR_{ort}(r_0)$	$\mathrm{tr}(D_{ort})$
to cointegration space	$ME_{ort}(r_0)$	$\lambda_{max}(D_{ort})$
($\mu_1 \neq 0,\ \beta'\mu_1 = 0$)		
Mean Term only	$TR_{mean}^{un}(r_0),\ TR_{mean}^{\delta}(r_0),\ TR_{mean}(r_0)$	$\mathrm{tr}(D_{mean})$
($\mu_1 = 0,\ \mu_0$ arbitrary)	$ME_{mean}^{un}(r_0),\ ME_{mean}^{\delta}(r_0),\ ME_{mean}(r_0)$	$\lambda_{max}(D_{mean})$

model

$$\Delta y_t - \mu_1 = \Pi(y_{t-1} - \mu_0) + \sum_{j=1}^{p-1} \Gamma_j(\Delta y_{t-j} - \mu_1) + \varepsilon_t, \quad t = p+1, p+2, \ldots, \tag{2.8}$$

which is obtained from (2.5) by applying $\tau = \beta'\mu_1 = 0$ and $\delta = 0$. Otherwise the test-setup is the same. The corresponding test statistics are denoted as $TR_{ort}(r_0)$ and $ME_{ort}(r_0)$. All described test versions and test statistics are summarized in Table 1.

We now turn to the asymptotic distributions of the considered test statistics. Let $\mathbf{B}_p(s) = (B_1(s), \ldots, B_p(s))'$ be a p-dimensional standard Brownian motion,

$$\begin{aligned}
D_{tr} &= \left(\int_0^1 \mathbf{B}_* \mathrm{d}\mathbf{B}_*'\right)' \left(\int_0^1 \mathbf{B}_* \mathbf{B}_*' \mathrm{d}s\right)^{-1} \left(\int_0^1 \mathbf{B}_* \mathrm{d}\mathbf{B}_*'\right), \\
D_{ort} &= \left(\int_0^1 \bar{\mathbf{B}}^s \mathrm{d}\mathbf{B}_{(n-r_0)}'\right)' \left(\int_0^1 \mathbf{B}^s \mathbf{B}^{s'} \mathrm{d}s\right)^{-1} \left(\int_0^1 \bar{\mathbf{B}}^s \mathrm{d}\mathbf{B}_{(n-r_0)}'\right), \\
D_{mean} &= \left(\int_0^1 \mathbf{B}_{(n-r_0)} \mathrm{d}\mathbf{B}_{(n-r_0)}'\right)' \left(\int_0^1 \mathbf{B}_{(n-r_0)} \mathbf{B}_{(n-r_0)}' \mathrm{d}s\right)^{-1} \left(\int_0^1 \mathbf{B}_{(n-r_0)} \mathrm{d}\mathbf{B}_{(n-r_0)}'\right),
\end{aligned} \tag{2.9}$$

where $\mathbf{B}_*(s) = \mathbf{B}_{(n-r_0)}(s) - s\mathbf{B}_{(n-r_0)}(1)$ is an $(n-r_0)$-dimensional Brownian bridge, $\mathrm{d}\mathbf{B}_*(s) = \mathrm{d}\mathbf{B}_{(n-r_0)}(s) - \mathrm{d}s\mathbf{B}_{(n-r_0)}$, $\bar{\mathbf{B}}^s(s) = \mathbf{B}^s(s) - \int_0^1 \mathbf{B}^s(u)du$, and $\mathbf{B}^s(s) = [\mathbf{B}_{(n-r_0-1)}(s)' : s]'$. Then, the test statistics have the limiting distributions given in Table 2 where $\mathrm{tr}(A)$ and $\lambda_{max}(A)$ denote the trace and the maximum eigenvalue of matrix A respectively.

Obviously, the null distributions depend on $n - r_0$, not on n and r_0 separately. They are independent of the actual values of μ_0 and, regarding the test versions allowing for a linear trend (μ_1 arbitrary) and level shifts (δ arbitrary), also independent of μ_1 and δ respectively.

Furthermore, only the treatment of the linear trend term is important for distinguishing the distributions. In contrast, the inclusion of level shifts with known shift dates or one level shift at unknown time does not change the asymptotic properties of the tests in case of a linear trend (μ_1 arbitrary) and a mean term only ($\mu_1 = 0$, μ_0 arbitrary). To be precise, allowing for level shifts does not affect the asymptotic distributions. Thus, $TR_{tr}^{un}(r_0)$, $TR_{tr}^{\delta}(r_0)$, $ME_{tr}^{un}(r_0)$, and $ME_{tr}^{\delta}(r_0)$ have the same limiting distributions as $TR_{tr}(r_0)$ and $ME_{tr}(r_0)$ respectively. Interestingly, $TR_{mean}(r_0)$ and $ME_{mean}(r_0)$ follow the same limiting distributions as the cointegration test statistics proposed by Johansen (1988) which assume no deterministic terms ($\mu_0 = \mu_1 = 0$).

Keep in mind that $ME_{tr}^{un}(r_0)$, $ME_{mean}^{un}(r_0)$, $ME_{tr}^{\delta}(r_0)$, $ME_{mean}^{\delta}(r_0)$, and $ME_{ort}(r_0)$ have not been considered in the respective references. Nevertheless, the limiting distributions of $ME_{tr}^{\delta}(r_0)$, and $ME_{mean}^{\delta}(r_0)$ can be easily obtained. As pointed out in the Appendix of Saikkonen & Lütkepohl (2000b) one can reduce the problem of deriving asymptotic results for the trace version to the case of no level shifts. For the latter setup a maximum eigenvalue version exists. Referring to the results and remarks in the Appendix of Saikkonen & Lütkepohl (2000c) one can derive the limiting distributions of $ME_{tr}^{\delta}(r_0)$ and $ME_{mean}^{\delta}(r_0)$ accordingly. The same holds for $ME_{tr}^{un}(r_0)$ and $ME_{mean}^{un}(r_0)$ using the results in the supplement[2] which accompanies Lütkepohl et al. (2004). One can see that a derivation of the limiting distributions is possible similar to the case of a known shift date. By contrast, the possibility of deriving the asymptotic distribution of $ME_{ort}(r_0)$ given in Table 2 is less obvious. Therefore, the Appendix shows how one can obtain the stated limiting distribution.

3 Gamma Distribution

As mentioned the asymptotic distributions are approximated by the Gamma distribution

$$\Gamma(y; a, b) = \int_0^y \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx} dx, \qquad y > 0, a > 0, b > 0. \tag{3.1}$$

Doornik (1998) has demonstrated that Gamma distributions provide a good approximation for the asymptotic distributions of the cointegration rank tests based on Johansen (1988). These distributions are functions of Brownian motions like the asymptotic null distributions given above. Accordingly, the use of the Gamma distribution works also very well for the

[2]The supplement is available in the internet at http://www.iue.it/Personal/Luetkepohl/Welcome.html.

Table 3. Comparison of Percentiles for $\mathrm{tr}(D_{tr})$

d	1	2	3	5	10
0.90 Response Surface Steps I-III	5.48	13.88	26.07	62.45	223.43
0.90 Gamma Approximation	5.47	13.79	25.88	61.90	220.86
0.90 Tabulated in Trenkler (2003)	5.42	13.78	25.93	61.92	220.92
0.95 Response Surface Steps I-III	6.79	15.76	28.52	66.13	230.24
0.95 Gamma Approximation	6.78	15.64	28.31	65.54	227.62
0.95 Tabulated in Trenkler (2003)	6.79	15.83	28.46	65.66	227.99
0.99 Response Surface Steps I-III	9.73	19.71	33.50	73.42	243.36
0.99 Gamma Approximation	9.74	19.54	33.26	72.73	240.68
0.99 Tabulated in Trenkler (2003)	10.04	19.85	33.76	73.12	241.80

cointegration tests with a prior adjustment of deterministic terms as shown later on. This also true for the limiting distributions of the test variants with a general linear trend (μ_1 arbitrary) which consist of Brownian bridges.

Obviously, the Gamma distribution depends only on the two parameters a and b. Furthermore, a random variable Y following a Gamma distribution has a mean $E(Y) = \mu = a/b$ and a variance $E(Y-\mu)^2 = \sigma^2 = a/b^2$ which both depend on a and b. Hence, given the mean m and variance v of the distribution of interest we can obtain an approximating Gamma distribution via the relationships $a = \mu^2/\sigma^2$ and $b = \mu/\sigma$ by setting $\mu = m$ and $\sigma^2 = v$.

In practice, we have to simulate the first two moments of the relevant distribution. Denoting the simulated moments by $\hat{m}$ and $\hat{v}$ the parameters of the approximating Gamma distribution are

$$\hat{a} = \hat{m}^2/\hat{v}, \qquad \hat{b} = \hat{m}/\hat{v}. \tag{3.2}$$

Since the asymptotic distributions of the cointegration test statistics differ with respect to the specification of the deterministic terms and the dimensions $d = n - r_0$ one has to determine different sets of the first two moments. In other words, one has to fit different Gamma distributions. The fitted Gamma distributions can then be used to compute p-values or percentiles.

For the practical implementation of the approximation one can also refer to χ^2 distribution with non-integer degrees of freedom instead of the Gamma distribution. The relationship

between both distributions is given by $2bY \sim \chi^2(2a)$ where Y has a Gamma distribution.

In line with the argumentation we have to distinguish two issues with respect to obtaining a reliable and applicable approximation. The first question is whether the Gamma distribution is a good approximation to the type of asymptotic distributions we consider. The second problem is to get reasonable estimates of the means and variances of the asymptotic distributions.

Since the asymptotic distributions are nonstandard and continuous we can evaluate the approximating quality of the Gamma distribution only for some simulated finite and discrete realizations. Such evaluation is presented in Tables 3 and 4. Table 3 shows the 0.90, 0.95, and 0.99 percentiles[3] for $\mathrm{tr}(D_{tr})$ for different dimensions d tabulated in Trenkler (2003). The percentiles are based on standard simulation methods for deriving critical values. These simulations also provide the mean and variance of the simulated asymptotic distributions with respect to the dimension d.[4] I have used the means and variances in order to fit Gamma distributions. The resulting percentiles are stated in the line *Gamma Approximation*. Obviously, the percentiles from the Tables in Trenkler (2003) and the Gamma approximation are rather close.

A second possibility to evaluate the approximation is to compute the p-values of the tabulated percentiles (critical values) using the fitted Gamma distributions. In Table 4, which matches Table 2 in Doornik (1998), the frequencies of the absolute errors made by using the Gamma approximation for dimensions 1 to 15 with respect to $\mathrm{tr}(D_{tr})$ is given. It can be seen that the fit for $\mathrm{tr}(D_{tr})$ is equally good as in Doornik (1998). In only one case for the 0.90 and 0.95 percentiles the p-values are outside the 0.0975-0.1025 and 0.0475-0.0525 ranges respectively. Thus, the Gamma approximation works also very well for the trace distribution which consist of Brownian Bridges. The results for the other trace distributions are similar.

As pointed out by Doornik (1998) the fit for the maximum eigenvalue distributions is worse because they resemble the distribution of the smallest order statistic. Nevertheless, the approximation may be still acceptable as shown in Table 4 for $\lambda_{max}(D_{tr})$. The accuracy of the fit regarding the $\lambda_{max}(D_{ort})$ and $\lambda_{max}(D_{mean})$ is again similar.

[3]Obviously, these percentiles correspond to the 10%, 5%, and 1% critical values.

[4]The means and variances for the different dimensions are not given in Trenkler (2003) but can be obtained from the author.

Table 4. Errors made by using Gamma approximation for $\mathrm{tr}(D_{tr})$ and $\lambda_{max}(D_{tr})$ with dimensions 1 to 15

	Frequencies for p-values					
	$\mathrm{tr}(D_{tr})$			$\lambda_{max}(D_{tr})$		
Error	0.10	0.05	0.01	0.10	0.05	0.01
< 0.0010	10	5	9	5	1	0
0.0010-0.0025	4	9	6	7	0	4
0.0025-0.0050	1	1	0	3	10	11
0.0050-0.0075	0	0	0	0	4	0
> 0.0075	0	0	0	0	0	0

The tabulated percentiles (critical values) for $\mathrm{tr}(D_{tr})$ from Trenkler (2003) and from corresponding simulations for $\lambda_{max}(D_{tr})$ are taken as a basis for computing the p-values using the Gamma approximation. The mean and variance for each dimension $d = 1, \ldots, 15$ from these simulations are taken to fit the Gamma distributions.

As mentioned before, given the result that the Gamma distribution provides a good approximation the next problem is to obtain reasonable estimates of the means and variances of the asymptotic distributions. These asymptotic moments will be derived from response surfaces described in the next section.

4 Response Surfaces

The aim of the response surfaces is to relate the means and variances of the asymptotic distributions to the dimensions d. That allows to obtain values of $\hat{m}$ and $\hat{v}$ for any d. This is done separately for the different specifications of the deterministic terms, i.e. for the three groups in Table 2 and the Trace and Maximum eigenvalues test variants. Hence, we have to derive six response surfaces for the means and variances respectively.

The response surfaces for the cointegration tests with a prior adjustment of deterministic terms presented in this paper follow closely the method suggested by Doornik (1998). I mention the differences when describing the methodology in the following. I have also derived response surfaces along the lines of Johansen et al. (2000). However, the procedure

Table 5. Simulation Details

Distributions:	$\text{tr}(D_{tr})$, $\text{tr}(D_{ort})$, $\text{tr}(D_{mean})$, $\lambda_{max}(D_{tr})$, $\lambda_{max}(D_{ort})$, $\lambda_{max}(D_{mean})$
Replications N:	100,000
Dimensions d:	1,2,...,15
	2,...,15 for $\text{tr}(D_{ort})$ and $\lambda_{max}(D_{ort})$
Sample Size T:	50, 75, 100, 150, 200, 250, 500, 1000, 2500, 5000

according to Doornik (1998) produces a slightly better fit for the moments. In Section 5 I briefly describe the procedure due the Johansen et al. (2000) and comment on the comparison between both approaches.

The response surfaces are obtained in three steps.

Step I

As for the usual simulation of critical values the components of the asymptotic distributions in (2.9) are replaced by their discrete counterparts. This is practically done by using T-step random walks for the Brownian motions. Based on the discrete quantities one can compute the asymptotic expressions. This is repeated a large number of times for different sets of dimensions d and sample lengths T. Table 5 summarizes the simulations details. Since we consider 15 dimensions and 10 sample lengths we have 150 simulation sets with 100,000 discrete realization of each of the corresponding asymptotic distribution.

Note that for $\text{tr}(D_{ort})$ and $\lambda_{max}(D_{ort})$ we start with dimension $d = 2$. The reason is that the cointegrating ranks specified both under the null and the alternative hypothesis have to be smaller than the system dimension due to the orthogonal trend (see Saikkonen & Lütkepohl 2000a). Accordingly, we obtain 140 simulation sets for these two distributions.

The simulation setup is in general identical to Doornik (1998). The only difference is that we have used 100,000 replications throughout all simulation experiments.

The simulations are done with GAUSS V5. In order to obtain independent standard normal variates for the generation of the the random walks I have used the Monster-KISS random number generator implemented in GAUSS V5 (compare Marsaglia 2000, Ford & Ford 2001). For further details on the simulation of the asymptotic distributions see Trenkler (2003).

Step II

The mean and variance of each of the 150 (140 for $\mathrm{tr}(D_{ort})$ and $\lambda_{max}(D_{ort})$) simulation sets for all of the six distributions are computed. Then, the aim is to obtain estimates of the asymptotic means and variances for each of the 15 (14) considered dimensions d with respect to the six distributions. To this end, I regress for each dimension d the different means regarding the sample lengths T on a constant, $1/T$, and $1/T^2$. This is done separately for all distributions. To capture specific effects of smaller sample lengths I also include $1/T^3$ and a dummy variable for $T = 50$ if they are significant. Furthermore, $1/T^2$ is dropped if not significant. In this case, I also omit the term $1/T$ if it is insignificant. Significance refers in all cases to the 5% significance level. By letting $T \to \infty$, the estimated intercept terms in these 15 (14) regressions provide estimates of the asymptotic means for each of the six distributions depending on the dimension d. Estimates for the 15 (14) asymptotic variances are obtained accordingly.

In contrast to Doornik (1998) I also consider the term $1/T^3$ besides the dummy variable for $T = 50$. The term $1/T^3$ can be regarded as a smooth dummy variable for small sample sizes. It often gives a better regression fit than the simple dummy variable. This is in line with a similar discussion in Johansen et al. (2000) regarding the dimension of the distributions. Note in this respect that if $1/T^3$ and the dummy for $T = 50$ are only separately significant I keep the variable with the higher t-ratio. In most of the cases $1/T^3$ is the preferable regressor.

Step III

To obtain the response surfaces the estimated asymptotic means and variances are regressed separately for each of the six distributions on polynomials of d. To be specific, we consider the terms d^2, d, $\sqrt{d}$, 1 (constant), and two dummy variables for $d = 1$ and $d = 2$. These regressors, however, are only included if they are significant at the 10% level. The estimated coefficients with respect to the six considered distributions are given in Tables 6 and 7. The coefficients represent the 12 response surfaces which can be used to compute estimates of the asymptotic mean and variance for any dimension d.

Finally, the computed means and variances can be used to obtain an approximating Gamma distribution for the asymptotic distribution of interest. By the help of this Gamma distribution it is possible to calculate p-values for any value of the test statistic or arbitrary percentiles.

Table 6. Response Surfaces for Mean and Variance of Asymptotic Distributions of the Trace Test Statistics

Mean	$\text{tr}(D_{tr})$	$\text{tr}(D_{ort})$	$\text{tr}(D_{mean})$
d^2	1.9996	2.0008	2.0000
d	0	−2.0990	−1.0134
$\sqrt{d}$	0	0.4463	0
1 (constant)	1.0365	0	0.1309
Dummy for $d = 1$	−0.3469	0	0.0218
Dummy for $d = 2$	−0.1112	−0.0503	0
Variance	$\text{tr}(D_{tr})$	$\text{tr}(D_{ort})$	$\text{tr}(D_{mean})$
d^2	2.9715	3.0152	2.9778
d	0	−3.0099	0
$\sqrt{d}$	0	2.1117	0
1 (constant)	1.4089	0	−1.7144
Dummy for $d = 1$	0	0	0.9507
Dummy for $d = 2$	0.4297	−0.8004	0.4259

Note that the response surfaces for $\text{tr}(D_{ort})$ are only valid for $d > 1$.

Remark I

The comparison of the response surfaces according to Steps I-III and the ones according to Johansen et al. (2000) in Section 5 includes the computation of a goodness of fit measure R_c^2 which can be seen as an analog to the usual R^2. Without discussing the details related to R_c^2 at this place here we may look at the values of R_c^2 for the mean and variance estimations with respect to $\text{tr}(D_{mean})$. They are shown in the last two columns of Table 8. Obviously, the values are close to one and therefore the regression fit is very good. This means our response surfaces describes the simulated moments in dependence on the sample lengths T and dimensions d rather well. Two further observations can be made. First, in line with findings in Johansen et al. (2000) the fit is slightly worse with respect to the variance. Second, the estimation is in general more accurate for higher dimensions. Finally, I have obtained similar results for the other trace and maximum eigenvalue distributions. Note again that a good explaining power regarding the asymptotic moments has to be distinguished from the accuracy issue of the Gamma approximation (compare the discussion in Section 3).

Table 7. Response Surfaces for Mean and Variance of Asymptotic Distributions of Maximum Eigenvalue Test Statistics

Mean	$\lambda_{max}(D_{tr})$	$\lambda_{max}(D_{ort})$	$\lambda_{max}(D_{mean})$
d^2	−0.0039	0	−0.0035
d	6.1600	5.8766	6.1365
$\sqrt{d}$	−3.3281	−1.9791	−3.2161
1 (constant)	−0.5071	−4.8042	−2.3701
Dummy for $d = 1$	0.3725	0	0.5970
Dummy for $d = 2$	0.0850	0	0.1007
Variance	$\lambda_{max}(D_{tr})$	$\lambda_{max}(D_{ort})$	$\lambda_{max}(D_{mean})$
d^2	−0.0418	0	−0.0258
d	3.4915	1.3279	2.6655
$\sqrt{d}$	9.2061	17.6880	12.4462
1 (constant)	−8.9114	1.3279	−13.6992
Dummy for $d = 1$	0.6652	0	0.8563
Dummy for $d = 2$	0	0	0

Note that the response surface for $\lambda_{max}(D_{ort})$ are only valid for $d > 1$.

Remark II

Regarding the Johansen trace test statistics Doornik (1998) has found that the term d^2 enters the response surfaces for the mean and variance with coefficients 2 and 3 respectively. Approximately, this is also the case here. Because we use a 4-digit precision slight differences can be observed.

Since the distribution $\text{tr}(D_{mean})$ is the same like the trace distribution for the Johansen test without any deterministic terms we can compare the corresponding coefficients in more detail. With respect to the mean the results of Doornik (1998) (d^2: 2, d : -1, constant: 0.07, dummy for $d = 1$: 0.07) are very close to ones in Table 6. The inclusion of the term $1/T^3$ when estimating the asymptotic moments may explain the deviations for the constant and the dummy regarding $d = 1$. The same argument can be given with respect to the variance for which the differences are more pronounced. In contrast to Doornik (1998) ($d^2 : 3$, $d : -0.33$, constant: -0.55) I have to consider dummy variables for $d = 1$ and $d = 2$ but do not need the general regressor d. The differences in the coefficients follow then accordingly. Nevertheless,

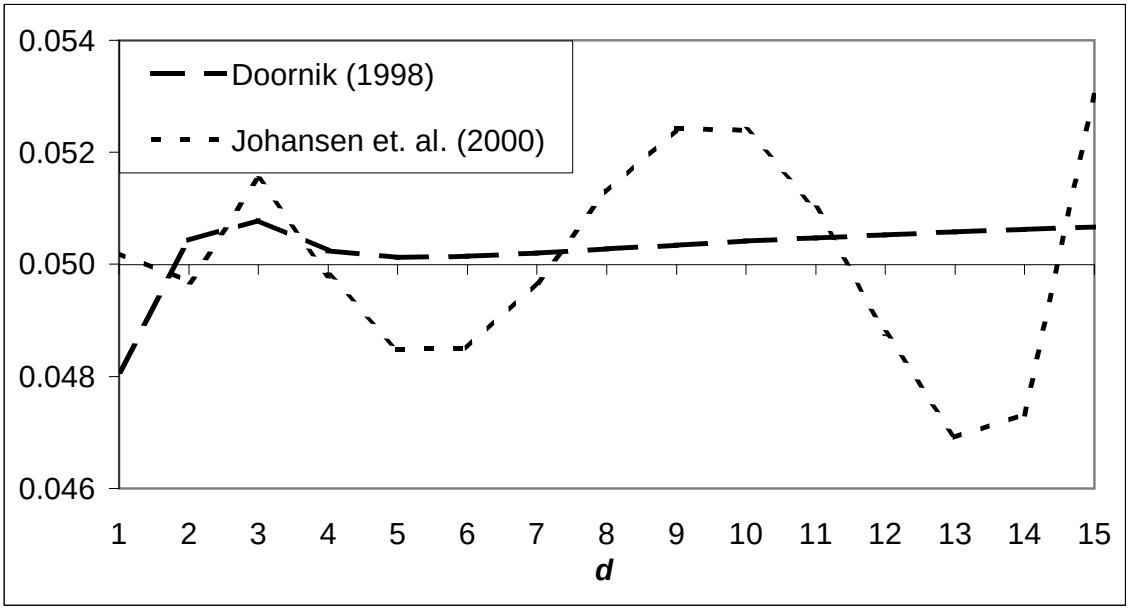

Figure 1. Comparison of p-values from different response surfaces for $\mathrm{tr}(D_{mean})$ based on the 0.95 percentiles resulting from the response surface according to Steps I-III.

the resulting means and variances for the dimensions $d = 1, \ldots, 15$ are very close for both response surfaces given here and in Doornik (1998). Accordingly, the percentiles based on the estimated asymptotic moments are also very similar. As an example Figure 1 shows the p-values resulting from the response surfaces in Doornik (1998) for the 0.95 percentiles from the response surface of this paper. Similar comments can be made for the corresponding maximum eigenvalue distribution $\lambda_{max}(D_{tr})$.

Remark III

As mentioned above, we obtain, in fact, estimates of the asymptotic moments (moments of the asymptotic distributions) by letting $T \to \infty$. As pointed out by Doornik (1998), letting $T \to \infty$ is the reason why the response surfaces give a better approximation to the asymptotic properties than the usual critical values which have been simulated only with respect to one sample size. Trenkler (2003) uses a rather large sample size of $T = 1000$ for the simulation of the critical values. Therefore, some of his tabulated critical values do not differ much from the corresponding percentiles obtained from the response surfaces. This can be seen for example in Table 3 regarding the distribution $\mathrm{tr}(D_{tr})$. For higher dimensions, however, the differences are much more pronounced resulting in strong deviations of the p-values from the intended ones. This is demonstrated in Figure 2. Here, the p-values for the 5% critical values (0.95 percentiles) of $\mathrm{tr}(D_{tr})$ given in Trenkler (2003) are depicted using

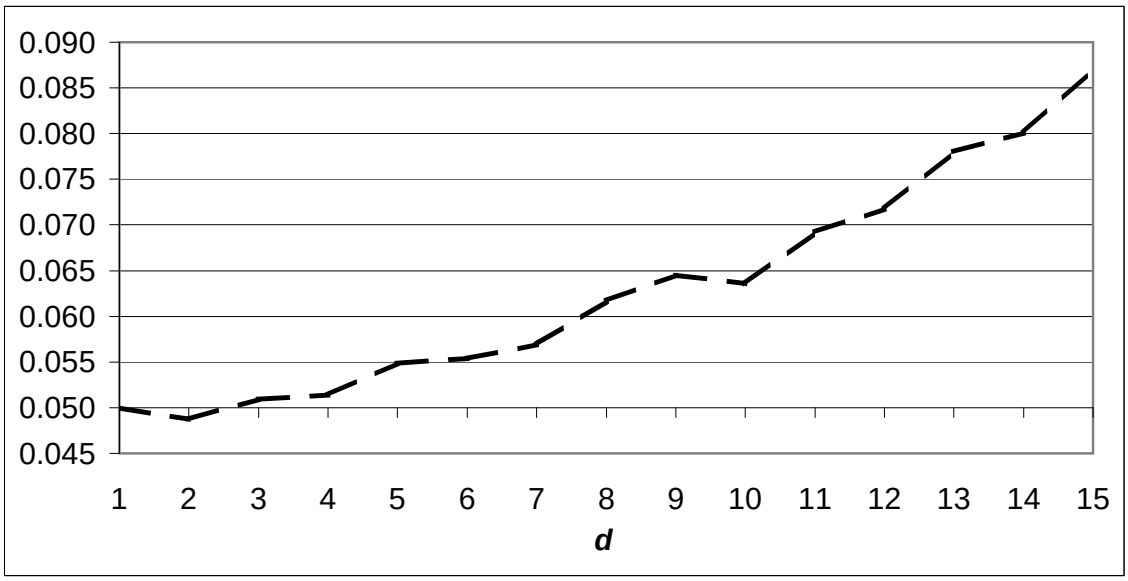

Figure 2. p-values for $\mathrm{tr}(D_{mean})$ regarding the 5% critical values (0.95 percentiles) in Trenkler (2003) using the Gamma distribution and response surface according to Steps I-III.

the Gamma distribution and the response surface of Steps I-III.

5 Comparison of Response Surfaces

Johansen et al. (2000) propose a different design for a response surface. They model the logarithm of the moments in order to deal with hetereoscedasticity in the error terms of the regression. The hetereoscedastictiy is due to the increasing levels of the mean and variance with higher dimension (see Doornik 1998). Therefore, Doornik (1998) performed the regressions separately for each d. The logarithmic specification allows to use only one regression for all d together by relating the log of the moments to a third-order polynomial in d and $1/T$.[5]

I have also derived response surfaces along the lines of Johansen et al. (2000) using the same values for the dimensions d and the sample lengths T as stated in Table 5. When comparing the fit of the two types of response surfaces with respect to the means and variances one faces two problems. First, the occurrence of heteroscedasticity and the separate regressions regarding d does not allow to give an overall R^2 for the response surface applied in this

[5]Johansen et al. (2000) consider tests with breaks in deterministic terms. Since the tests' limiting distributions depend on the break dates Johansen et al. (2000) also include the relative sample lengths of the break periods in their response surface.

Table 8. Comparison of Goodness of Fit (R_c^2) of Response Surfaces for tr(D_{mean})

Dimension	According to Johansen et al. (2000)		According to Steps I-III	
d	Mean	Variance	Mean	Variance
1	0.985432	0.963066	0.983675	0.960561
2	0.993625	0.967177	0.998589	0.981171
3	0.981840	0.977182	0.998318	0.932093
4	0.999305	0.985807	0.999772	0.996067
5	0.996203	0.983667	0.999949	0.985111
6	0.997173	0.991496	0.999778	0.991723
7	0.999738	0.995261	0.999942	0.994541
8	0.999465	0.995157	0.999962	0.980212
9	0.998537	0.991262	0.999993	0.994585
10	0.999022	0.994827	0.999988	0.995594
11	0.999935	0.993021	0.999988	0.994029
12	0.999774	0.996992	0.999997	0.995517
13	0.999089	0.992928	0.999992	0.995999
14	0.999647	0.995427	0.999998	0.998875
15	0.999364	0.995382	0.999997	0.998552

paper in contrast to the approach of Johansen et al. (2000). Second, Johansen et al. (2000) model the logarithm of the moments and I consider the moments directly. Thus, measures of fit would refer to the variation of different dependent variables and cannot be compared directly (see Greene 1997, p. 256). Therefore, I compare the response surfaces separately for each dimension. Following the suggestion of Greene (1997, p. 256), the comparison is based on an analog to the usual R^2:

$$R_c^2 = 1 - \frac{\sum_{i=1}^{10} e_i^2}{\sum_{i=1}^{10} (y_i - \bar{y})^2} \tag{5.1}$$

where the y_i's represent the simulated means (variances) with respect to the 10 sample sizes. For the setup used in this paper the terms e_i are equal to the residuals from the regressions described in Step II plus the additional error made in Step III when fitting the asymptotic moments to polynomials of d. With respect to the approach of Johansen et al. (2000) the terms e_i are defined as $e_i = y_i - exp[\ln(\hat{y}_i)]$ where $\ln(\hat{y}_i)$ stands for the fitted log-moments (means or variances) from the corresponding regression. In other words, the fitted logarithms

of the mean or variance are converted back to the level. The goodness of fit measured by R_c^2 for the distribution $\text{tr}(D_{tr})$ is displayed in Table 8. Obviously, for most of the dimensions the fit for the mean and variance is better when using the setup presented in this paper. The results for the other distributions are similar.

The presented comparison does not take account of the different number of estimated parameters involved in the response surface. For the current example of $\text{tr}(D_{tr})$ I have estimated 57 (mean) and 61 (variance) coefficients according to Steps I-III but only 15 (mean) and 12 (variance) coefficients according to Johansen et al. (2000) when deleting insignificant regressors. Using an adjusted version of R_c^2 improves the relative performance of the response surfaces following Johansen et al. (2000). Nevertheless, for most of the dimensions the goodness of fit is still inferior.

Figure 1 also shows the p-values for $\text{tr}(D_{mean})$ obtained from the response surfaces based on Johansen et al. (2000) when the 0.95 percentiles regarding the response surface of Steps I-III are used. The p-values oscillate around the 5% value with increasing deviations for higher dimensions d. For smaller dimensions, however, the differences are less pronounced.

6 Summary

Following Doornik (1998) and Johansen et al. (2000) I have presented a procedure to approximate the asymptotic distributions of systems cointegration tests with a prior adjustment for deterministic terms suggested by Lütkepohl et al. (2004), Saikkonen & Lütkepohl (2000a, 2000b, 2000c) and Saikkonen & Luukkonen (1997). These tests make different assumptions on the inclusion of deterministic components like a mean term, a linear trend or a level shift.

The asymptotic distributions are approximated by the Gamma distribution. To fit a Gamma distribution one needs the mean and variance of the asymptotic test distributions. They can be obtained from response surfaces. The corresponding coefficients to compute the asymptotic moments are presented in this paper. Via the fitted Gamma distributions one can easily derive arbitrary p-values or percentiles.

Another approach to derive p-values is the use bootstrap procedures. Bootstrap methods can help to generate more accurate p-values in small samples than ordinary asymptotic approximations. Recently, Swensen (2004) has derived an asymptotically valid bootstrap

procedure for Johansen-type cointegration tests. However, Harris & Judge (1998) and van Giersbergen (1996) have shown that bootstraps regarding the Johansen tests produce a better approximation than standard critical values only in a few situations.

Appendix

In this appendix I outline how to derive the limiting distribution of $ME_{ort}(r_0)$. The proof in Saikkonen & Lütkepohl (2000a) for the limiting distribution of the corresponding trace test statistic is based on the sum of the eigenvalues. It does not involve the consideration of a joint limiting distribution of the eigenvalues. Therefore, we cannot use this framework since we need the distribution of the (r_0+1) largest estimated eigenvalue $\tilde{\lambda}_{r_0+1}$. Following the final remark in the Appendix of Saikkonen & Lütkepohl (2000c), one may refer to the approach in Johansen (1995, pp. 158-161) using the necessary convergence and distribution results proven in the Appendix of Saikkonen & Lütkepohl (2000a). That is, one derives the joint limiting distribution of the eigenvalues so that the maximum eigenvalue distribution can be easily obtained. In the following I use the notation from Johansen (1995) and Saikkonen & Lütkepohl (2000a). The reader is referred to these articles for a definition of the expressions applied here.

The main ingredients of the proof in Johansen (1995, pp. 158-161) are summarized in Lemma 10.3 of Johansen (1995). Corresponding convergence results regarding the product moments in (10.13)-(10.15) of Lemma 10.3 can be obtained for the setup of Saikkonen & Lütkepohl (2000a) by using (A.4) of Lemma A.1 in Saikkonen & Lütkepohl (2000a). Note in this respect that $\Delta x_{t-j} = \Delta y_{t-j} - \mu_0$ for $j = 0, \ldots, p-1$. Further, q_t in (A.4) does not contain the term Δx_t but the necessary properties also hold for Δx_t since it is a zero-mean stationary process. Moreover, we need corresponding expressions for the distribution results in (10.16) of Lemma 10.3 and for a transformation of (10.17) given in Johansen (1995, p. 159). I will focus on the derivation of these expressions in the following.

From (10.16) and (10.17) in Johansen (1995) we obtain the following asymptotic results

$$B_T'(S_{10} - S_{11}\beta\alpha')\alpha_\perp = B_T' S_{10}\alpha_\perp \xrightarrow{d} \int_0^1 G(dW)'\alpha_\perp, \tag{A.1}$$

$$T^{-1}B_T' S_{11} B_T \xrightarrow{d} \int_0^1 GG'du. \tag{A.2}$$

with the definitions given in Johansen (1995). The crucial step in using the framework of Johansen (1995) for $ME_{ort}(r_0)$ is to replace the matrix B_T in (A.1) and (A.2) by the expression $T\beta_\perp \underline{D}_T^{-1}$ where $\underline{D}_T^{-1}$ is defined in the Appendix of Saikkonen & Lütkepohl (2000a). Thus, we have to analyse

$$T\underline{D}_T^{-1}\beta_\perp' S_{10}\alpha_\perp \quad \text{and} \tag{A.3}$$

$$T\underline{D}_T^{-1}\beta_\perp' S_{11}\beta_\perp \underline{D}_T^{-1}. \tag{A.4}$$

In order to derive the limiting distributions of (A.3) and (A.4) we have first to express S_{10} and S_{11} in terms of the quantities used in Saikkonen & Lütkepohl (2000a). To this end, we have to remember that $ME_{ort}(r_0)$ is based on the auxiliary VECM

$$\Delta y_t - \tilde{\mu}_1 = \alpha\beta'(y_{t-1} - \tilde{\mu}_0) + \sum_{j=1}^{p-1} \Gamma_j(\Delta y_{t-j} - \tilde{\mu}_1) + e_t, \quad t = p+1, p+2, \ldots. \tag{A.5}$$

where I have replaced μ_0 and μ_1 in (2.8) by some estimators $\tilde{\mu}_0$ and $\tilde{\mu}_1$ which satisfy (2.12)-(2.15) in Saikkonen & Lütkepohl (2000a). Hence, we apply a Johansen type test on this model using the the adjusted quantities $\Delta y_t - \tilde{\mu}_1$ and $y_{t-1} - \tilde{\mu}_0$. In terms of the terminology in Johansen (1995) (A.5) can be written as

$$Z_{0t} = \alpha\beta' Z_{1t} + \Psi Z_{2t} \tag{A.6}$$

applying obvious definitions for Z_{0t}, Z_{1t}, Z_{2t}, and Ψ. Note the fact that Z_{2t} is equal to $\tilde{z}_t$ used in Saikkonen & Lütkepohl (2000a).

I now prove the following Lemma.

Lemma A.1. Under the assumptions made in Saikkonen & Lütkepohl (2000a) we obtain

$$T\underline{D}_T^{-1}\beta_\perp' S_{10}\alpha_\perp \xrightarrow{d} \int_0^1 \bar{\mathbf{F}}(s)d\mathbf{B}(s)'\alpha_\perp \tag{A.7}$$

$$T\underline{D}_T^{-1}\beta_\perp' S_{11}\beta_\perp \underline{D}_T^{-1} \xrightarrow{d} \int_0^1 \mathbf{F}(s)\mathbf{F}(s)'ds. \tag{A.8}$$

where $\mathbf{B}(s)$ is a Brownian motion and $\bar{\mathbf{F}}(s)$ and $\mathbf{F}(s)$ are functions of Brownian motions defined in (2.9), (A.25), and (A.26) of Saikkonen & Lütkepohl (2000a) respectively.

Let us start with proving (A.7). First, note that $S_{10} = T^{-1}\sum_{t=p+1}^{T} R_{1t}R_{0t}'$ where R_{1t} and R_{0t} are the residuals obtained from regressing Z_{1t} and Z_{0t} on Z_{2t} respectively. In line with (6.9) of Johansen (1995) we have $R_{0t} = \alpha\beta' R_{1t} + e_t$. Thus, we can write

$$S_{10} = T^{-1}\sum_{t=p+1}^{T} R_{1t}(\alpha\beta' R_{1t} + e_t)' = T^{-1}\sum_{t=p+1}^{T} R_{1t}R_{1t}'\beta\alpha' + T^{-1}\sum_{t=p+1}^{T} R_{1t}e_t'. \tag{A.9}$$

Hence, we get for $S_{10}\alpha_\perp$

$$\begin{aligned} S_{10}\alpha_\perp &= T^{-1}\sum_{t=p+1}^{T} R_{1t}e_t'\alpha_\perp \\ &= T^{-1}\left(\sum_{t=p+1}^{T} Z_{1t}e_t' - \sum_{t=p+1}^{T} Z_{1t}Z_{2t}'\left(\sum_{t=p+1}^{T} Z_{2t}Z_{2t}'\right)^{-1}\sum_{t=p+1}^{T} Z_{2t}e_t'\right)\alpha_\perp \\ &= T^{-1}\sum_{t=p+1}^{T} Z_{1t}e_t'\alpha_\perp - T^{-1}\sum_{t=p+1}^{T} Z_{1t}Z_{2t}'\left(\sum_{t=p+1}^{T} Z_{2t}Z_{2t}'\right)^{-1}\sum_{t=p+1}^{T} Z_{2t}e_t'\alpha_\perp \\ &= A_{1T} - A_{2T}. \end{aligned} \tag{A.10}$$

where A_{1T} and A_{2T} are described below. I first show $T\underline{D}_T^{-1}\beta_\perp' A_{2T} = o_p(1)$ and then $T\underline{D}_T^{-1}\beta_\perp' A_{1T} \xrightarrow{d} \int_0^1 \bar{\mathbf{F}}(s)d\mathbf{B}(s)'\alpha_\perp$.

We can write

$$\begin{aligned} T\underline{D}_T^{-1}\beta_\perp' A_{2T} &= T\underline{D}_T^{-1}\beta_\perp' T^{-1}\sum_{t=p+1}^{T} Z_{1t}Z_{2t}'\left(\sum_{t=p+1}^{T} Z_{2t}Z_{2t}'\right)^{-1}\sum_{t=p+1}^{T} Z_{2t}e_t'\alpha_\perp \\ &= T^{-1/2}\underline{D}_T^{-1}\beta_\perp'\sum_{t=p+1}^{T} Z_{1t}Z_{2t}'\left(T^{-1}\sum_{t=p+1}^{T} Z_{2t}Z_{2t}'\right)^{-1}T^{-1/2}\sum_{t=p+1}^{T} Z_{2t}e_t'\alpha_\perp \\ &= T^{-1/2}\underline{D}_T^{-1}\beta_\perp'\sum_{t=p+1}^{T} Z_{1t}Z_{2t}'\cdot O_p(1)\cdot O_p(1) \\ &= T^{-1/2}\underline{D}_T^{-1}\sum_{t=p+1}^{T}\bar{v}_{t-1}Z_{2t}'\cdot O_p(1)\cdot O_p(1) \\ &= o_p(1)\cdot O_p(1) \\ &= o_p(1) \end{aligned} \tag{A.11}$$

where $\bar{v}_{t-1} = \beta_\perp' Z_{1t} = \beta_\perp'(y_{t-1} - \tilde{\mu}_0)$. Note, the third line of (A.11) is obtained from (A.4) and (A.21) in Saikkonen & Lütkepohl (2000a) replacing $\tilde{\alpha}_\perp$ by $\alpha_\perp$ and given the definition of $\tilde{q}_t$ used there. Moreover, the last but one line is justified owing to (A.16) of Saikkonen & Lütkepohl (2000a) by replacing $\tilde{\beta}_\perp$ with $\beta_\perp$.

Next, we have

$$\begin{aligned} T\underline{D}_T^{-1}\beta_\perp' A_{1T} &= T\underline{D}_T^{-1}\beta_\perp' T^{-1}\sum_{t=p+1}^{T} Z_{1t}e_t'\alpha_\perp \\ &= \underline{D}_T^{-1}\sum_{t=p+1}^{T}\bar{v}_{t-1}e_t'\alpha_\perp \quad \xrightarrow{d} \int_0^1 \bar{\mathbf{F}}(s)d\mathbf{B}(s)'\alpha_\perp \end{aligned} \tag{A.12}$$

where the last line follows from (A.26) of Saikkonen & Lütkepohl (2000a). Thus, (A.12) and (A.11) prove (A.7).

Now, we turn to prove (A.8). Using $S_{11} = T^{-1}\sum_{t=p+1}^{T} R_{1t}R_{1t}'$ we have

$$\begin{aligned} & T\underline{D}_T^{-1}\beta_\perp' S_{11}\beta_\perp \underline{D}_T^{-1} \\ &= \underline{D}_T^{-1}\beta_\perp' \sum_{t=p+1}^{T} R_{1t}R_{1t}'\beta_\perp \underline{D}_T^{-1} \\ &= \underline{D}_T^{-1}\beta_\perp' \left(\sum_{t=p+1}^{T} Z_{1t}Z_{1t}' - \sum_{t=p+1}^{T} Z_{1t}Z_{2t}' \left(\sum_{t=p+1}^{T} Z_{2t}Z_{2t}' \right)^{-1} \sum_{t=p+1}^{T} Z_{2t}Z_{1t}' \right) \beta_\perp \underline{D}_T^{-1} \\ &= \underline{D}_T^{-1}\beta_\perp' \sum_{t=p+1}^{T} Z_{1t}Z_{1t}'\beta_\perp \underline{D}_T^{-1} - \underline{D}_T^{-1}\beta_\perp' \sum_{t=p+1}^{T} Z_{1t}Z_{2t}' \left(\sum_{t=p+1}^{T} Z_{2t}Z_{2t}' \right)^{-1} \sum_{t=p+1}^{T} Z_{2t}Z_{1t}'\beta_\perp \underline{D}_T^{-1} \\ &= \underline{D}_T^{-1}\beta_\perp' \sum_{t=p+1}^{T} Z_{1t}Z_{1t}'\beta_\perp \underline{D}_T^{-1} - o_p(1) \\ &= \underline{D}_T^{-1} \sum_{t=p+1}^{T} \bar{v}_{t-1}\bar{v}_{t-1}'\underline{D}_T^{-1} \quad \xrightarrow{d} \int_0^1 \mathbf{F}(s)\mathbf{F}(s)'ds \end{aligned} \tag{A.13}$$

where the fourth equality follows from arguments used for (A.11) and the last line is obtained by applying (A.25) similar to (A.27) of Saikkonen & Lütkepohl (2000a). This completes the proof of Lemma A.1. ■

Using Lemma A.1 we can derive the joint limiting distribution of the eigenvalues in line with the approach of Johansen (1995, pp. 158-161). Then, following the argumentation in Hansen & Johansen (1998, p. 124) we obtain the limiting distribution of $ME_{ort}(r_0)$ as stated in Table 2.

References

Doornik, J. A. (1998). Approximations to the asymptotic distributions of cointegration tests, *Journal of Economic Surveys* **12**: 573–593.

Ford, M. P. & Ford, D. J. (2001). Investigation of GAUSS' random number generators, *http://www.aptech.com/random/rndu36.pdf*.

Greene, W. H. (1997). *Econometric Analysis*, Prentice Hall, New Jersey.

Hansen, P. R. & Johansen, S. (1998). *Workbook on Cointegration*, Oxford University Press, Oxford.

Harris, R. I. D. & Judge, G. G. (1998). Small sample testing for cointegration using the bootstrap approach, *Economics Letters* **58**: 31–37.

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 Vector Autoregressive Models*, Oxford University Press, Oxford.

Johansen, S., Mosconi, R. & Nielsen, B. (2000). Cointegration analysis in the presence of structural breaks in the deterministic trend, *Econometrics Journal* **3**: 216–249.

Lütkepohl, H., Saikkonen, P. & Trenkler, C. (2001). Maximum eigenvalue versus trace tests for the cointegrating rank of a VAR process, *Econometrics Journal* **4**: 287–310.

Lütkepohl, H., Saikkonen, P. & Trenkler, C. (2004). Testing for the cointegrating rank of a VAR process with a structural shift at unknown time, *Econometrica,* **72**: 647–662.

MacKinnon, J. G., Haug, A. A. & Michelis, L. (1999). Numerical distribution functions of likelihood ration tests for cointegration, *Applied Econometrics* **14**: 563–577.

Marsaglia, G. (2000). The monster, a random number generator with period over 10^{2587} as long as the previously touted longest period one, *mimeo*, Department of Statistics, Florida State University.

Saikkonen, P. & Lütkepohl, H. (2000a). Testing for the cointegrating rank of a VAR process with an intercept, *Econometric Theory* **16**: 373–406.

Saikkonen, P. & Lütkepohl, H. (2000b). Testing for the cointegration rank of a VAR process with structural shifts, *Journal of Business & Economic Statistics,* **18**: 451–464.

Saikkonen, P. & Lütkepohl, H. (2000c). Trend adjustment prior to testing for the cointegration rank of a VAR process, *Journal of Time Series Analysis* **21**: 435–456.

Saikkonen, P. & Luukkonen, R. (1997). Testing cointegration in infinite vector autoregressive processes, *Journal of Econometrics* **81**: 93–126.

Swensen, A. R. (2004). Bootstrap algorithms for testing and determining the cointegration rank in VAR models, *Working Paper*, Department of Mathematics, University of Oslo.

Trenkler, C. (2003). A new set of critical values for systems cointegration tests with a prior adjustment for deterministic terms, *Economics Bulletin* **3**: 1–9.

van Giersbergen, N. P. A. (1996). Bootstrapping the trace statistic in VAR models: Monte Carlo results and applications, *Oxford Bulletin of Economics and Statistics* **58**: 391–408.