

EUI WORKING PAPERS

ECO No. 2006/29



Testing for the Cointegrating Rank of
a VAR Process with Level Shift and Trend Break

**CARSTEN TRENKLER, PENTTI SAIKKONEN
and HELMUT LÜTKEPOHL**



EUROPEAN UNIVERSITY INSTITUTE

Department of Economics

EUROPEAN UNIVERSITY INSTITUTE
DEPARTMENT OF ECONOMICS

*Testing for the Cointegrating Rank of a
VAR Process with Level Shift and Trend Break*

CARSTEN TRENKLER, PENTTI SAIKKONEN
and
HELMUT LÜTKEPOHL

This text may be downloaded for personal research purposes only. Any additional reproduction for such purposes, whether in hard copy or electronically, requires the consent of the author(s), editor(s). If cited or quoted, reference should be made to the full name of the author(s), editor(s), the title, the working paper or other series, the year, and the publisher.

The author(s)/editor(s) should inform the Economics Department of the EUI if the paper is to be published elsewhere, and should also assume responsibility for any consequent obligation(s).

ISSN 1725-6704

© 2006 Carsten Trenkler, Pentti Saikkonen and Helmut Lütkepohl

Printed in Italy
European University Institute
Badia Fiesolana
I – 50016 San Domenico di Fiesole (FI)
Italy

<http://www.iue.it/>
<http://cadmus.iue.it/dspace/index.jsp>

Testing for the Cointegrating Rank of a VAR Process with Level Shift and Trend Break

by

Carsten Trenkler
Humboldt University Berlin

Pentti Saikkonen
University of Helsinki

and

Helmut Lütkepohl
European University Institute, Florence

Address for correspondence: Carsten Trenkler, Institute of Statistics and Econometrics,
Humboldt University Berlin, Spandauer Str. 1, D-10178 Berlin, Germany;
email: trenkler@wiwi.hu-berlin.de

Abstract

A test for the cointegrating rank of a vector autoregressive (VAR) process with a possible shift and broken linear trend is proposed. The break point is assumed to be known. The setup is a VAR process for cointegrated variables. The tests are not likelihood ratio tests but the deterministic terms including the broken trends are removed first by a GLS procedure and a likelihood ratio type test is applied to the adjusted series. The asymptotic null distribution of the test is derived and it is shown by a Monte Carlo experiment that the test has better small sample properties in many cases than a corresponding Gaussian likelihood ratio test for the cointegrating rank.

Key words: Cointegration, structural break, vector autoregressive process, error correction model

JEL classification: C32

*The first author is grateful to the Deutsche Forschungsgemeinschaft, SFB 649, for financial support and the second author acknowledges financial support by the Yrjö Jahnsson Foundation and the Academy of Finland.

1 Motivation and Introduction

Structural breaks are common in economic data. Ignoring them can lead to seriously incorrect inference in particular in unit root and cointegration testing. Therefore a considerable literature has developed on testing for unit roots and cointegration when breaks occur in the deterministic term following the pioneering work of Perron (1989, 1990). He considered unit root tests for univariate time series and treated three cases of particular importance, namely the case of a shift in the level of the process, the case of a change in the trend slope and the case where both types of breaks occur. These cases have also been dealt with in the cointegration testing literature. Specifically Johansen, Mosconi and Nielsen (2000) extended the tests proposed by Johansen (1995) and considered all these cases in a Gaussian vector autoregressive (VAR) framework. They developed likelihood ratio (LR) tests for the cointegrating rank for all three situations. It was noted by Saikkonen and Lütkepohl (1999, 2000a), however, that other tests may be advantageous in terms of local power if there is just a level shift. In this paper we will extend the ideas used by the latter authors to the case of a break in the trend slope in addition to the shift in the level of the data generation process (DGP).

The general setup is a VAR process with a linear trend term which may have a level shift and a break in the trend slope at a known point in time as in the Johansen et al. (2000) paper. The deterministic term is specified in a slightly different way than in that paper. If a break is believed to have occurred in the deterministic part of the process only and does not affect the stochastic part, it seems natural to us to strictly separate the deterministic from the stochastic part in setting up the model. Therefore the deterministic part is added to a zero mean purely stochastic process in our setup.

Our test proceeds by first estimating the deterministic part of the DGP by a generalized least squares (GLS) procedure and then removing this part from the series. Thereafter an LR type test for the cointegrating rank is applied. Unlike in the case considered by Saikkonen and Lütkepohl (2000a) where a break occurs only in the level of the process, in the present setup with a possible break in the trend slope the asymptotic distribution of the test statistic under the null hypothesis depends on the break date. In this respect our test is similar to the Johansen et al. (2000) test (henceforth JMN test). The asymptotic distribution is different from that of the latter test, however. Response surface techniques will be used to provide easy to use approximations to the asymptotic distributions of the test which also allow to provide p -values of the test for any possible break date. Thereby the test will be easy to use in empirical applications without simulating new critical values for each specific case. In a Monte Carlo comparison it is shown that the new test may have considerably better small sample properties than the JMN test. In particular, the tendency of the JMN test for substantial size distortions will be seen to be reduced for our test. Moreover, in many situations the small sample power of our test will be seen to be better than that of the JMN test.

As mentioned earlier, we assume a known break date like Johansen et al. (2000). This assumption has been criticized in the related unit root and cointegration testing literature on the grounds that the break date is often not known with certainty in practice and an incorrect break point leads to a misspecified model with negative implications for the properties of the tests. In response to this critique a number of proposals were made for endogenizing the break point selection in particular in testing for unit roots. For cointegration testing when the DGP has a level shift, proposals were made and investigated by Lütkepohl, Saikkonen

and Trenkler (2004) and Saikkonen, Lütkepohl and Trenkler (2006). In the present paper we restrict the discussion to the known break date case because an extension to the unknown break date case would be a nontrivial extension for which we have not been able to work out the details so far. Moreover, we believe that the known break date case is quite relevant for applied work as well because in many situations the break point is in fact known. For example, for many German macroeconomic time series there is a break at the time of the German unification and the unification time is obviously known. An application of our tests to German data will also be given in this paper. Another case where a break has occurred at a known point in time is the European monetary union. Even though a break may have occurred at an unknown time in many economic time series, it is debatable whether the presently available models are actually suitable in such cases because the assumed breaks are still very simple. In particular, they are confined to the deterministic terms and leave the dynamic structure unchanged. If the break point is unknown it is unclear that the type of the break is known sufficiently precisely to make the tests applicable.

The structure of this study is as follows. In the next section we discuss the model setup and in Section 3 the new cointegrating rank tests are presented including asymptotic null distributions. Response surface results for approximating the asymptotic test distributions are given in Section 4 and a Monte Carlo study comparing the small sample properties of our new tests with the JMN tests are presented in Section 5. In Section 6 the tests are used to investigate the cointegration properties of a small German macroeconomic system in which the series have breaks at a known point in time. Although we restrict the discussion largely to the case of a single break point it should be noted that extensions to a multiple break situation are straightforward. To the extent necessary for our purposes, these extensions will be mentioned throughout. Some further extensions are briefly summarized in the final section together with general conclusions from our study. Response surface coefficients and mathematical derivations are provided in the Appendix.

Throughout the paper the following notation and terminology is used. The symbols Δ and L denote the differencing and lag operators, respectively. An integrated process of order d is called $I(d)$, that is, the stochastic part of the process is stationary or asymptotically stationary after differencing d times whereas it is still nonstationary after differencing $d - 1$ times only. Convergence in distribution is denoted by \xrightarrow{d} and i.i.d. abbreviates independently, identically distributed. The symbols $O_p(\cdot)$ and $o_p(\cdot)$ are used as usual for boundedness and convergence in probability, respectively. Furthermore, $\|A\|$, $\text{tr}(A)$, $\det(A)$ and $\text{rk}(A)$ denote the Euclidean norm, the trace, determinant and rank of the matrix A , respectively. If A is an $(n \times m)$ matrix of full column rank ($n > m$), an orthogonal complement is denoted by A_\perp . The zero matrix is the orthogonal complement of a nonsingular square matrix and an identity matrix of suitable dimension is the orthogonal complement of a zero matrix. The symbol I_n signifies an $(n \times n)$ identity matrix and for matrices A_1, \dots, A_s , $\text{diag}[A_1 : \dots : A_s]$ is the block-diagonal matrix with A_1, \dots, A_s on the diagonal. ML, LS, GLS, RR, LR, VAR and VECM abbreviate maximum likelihood, least squares, generalized least squares, reduced rank, likelihood ratio, vector autoregressive and vector error correction model, respectively. A summation is defined to be zero if the lower bound of the summation index exceeds the upper bound.

2 The Model Setup

The framework of Saikkonen and Lütkepohl (2000a) (henceforth S&L) is extended by allowing for a change in the trend slope. Suppose $y_t = (y_{1t}, \dots, y_{nt})'$ ($t = 1, \dots, T$) is generated by a process with constant, linear trend, level shift and change in the trend slope at known time τ ,

$$y_t = \mu_0 + \mu_1 t + \delta_0 d_t + \delta_1 b_t + x_t, \quad t = 1, 2, \dots, \quad (2.1)$$

where μ_i and δ_i ($i = 0, 1$) are unknown $(n \times 1)$ parameter vectors, and d_t and b_t are dummy variables defined by $d_t = b_t = 0$ for $t < \tau$, and $d_t = 1$ and $b_t = t - \tau + 1$ for $t \geq \tau$. The value of the break date τ is assumed to depend on the sample size such that the break occurs at a fixed fraction of the sample size. More precisely, it is assumed that

$$\tau = [T\lambda] \quad \text{with} \quad 0 < \underline{\lambda} \leq \lambda \leq \bar{\lambda} < 1, \quad (2.2)$$

where $\underline{\lambda}$ and $\bar{\lambda}$ are specified real numbers and $[\cdot]$ denotes the integer part of the argument. In other words, the break date may not be at the very beginning or at the very end of the sample. Note that $\underline{\lambda}$ and $\bar{\lambda}$ may be arbitrarily close to zero and one, respectively. Therefore our assumption regarding the break date is not very restrictive.

Of course, δ_0 or δ_1 may be zero. If $\delta_1 = 0$ is known a priori we would be back in the framework of S&L and there is nothing new. Hence, in the context of the present paper the case $\delta_1 \neq 0$ is of primary interest.

It is important to note that the deterministic part is simply added to the stochastic part x_t of the process. Our formulation differs in this respect from the setup used by Johansen et al. (2000) who introduce the deterministic terms directly in the VAR model. In our setup the process x_t is assumed to have a zero mean VAR(p) representation,

$$x_t = A_1 x_{t-1} + \dots + A_p x_{t-p} + \varepsilon_t, \quad t = 1, 2, \dots, \quad (2.3)$$

without deterministic terms. Here the A_j are $(n \times n)$ coefficient matrices. For simplicity, it is assumed that $x_t = 0$ for $t \leq 0$ and $\varepsilon_t \sim \text{i.i.d.}(0, \Omega)$, that is, the ε_t are i.i.d. vectors with zero mean and covariance matrix Ω . We also assume that all moments of ε_t of order c exist, where c is a number greater than 4. The zero initial value assumption for x_t , $t \leq 0$, could be replaced by the assumption that the initial values are from a fixed probability distribution which does not depend on the sample size.

The VECM form of the process x_t is

$$\Delta x_t = \Pi x_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta x_{t-j} + \varepsilon_t, \quad t = 1, 2, \dots, \quad (2.4)$$

where $\Pi = -(I_n - A_1 - \dots - A_p)$ and $\Gamma_j = -(A_{j+1} + \dots + A_p)$ ($j = 1, \dots, p-1$) are $(n \times n)$ matrices. The process x_t is assumed to be at most $I(1)$ and cointegrated with cointegrating rank r . Hence, the matrix Π can be written as $\Pi = \alpha\beta'$, where α and β are $(n \times r)$ matrices of full column rank. As is well-known, $\beta'x_t$ and Δx_t are then zero mean $I(0)$ processes. Defining $\Psi = I_n - \Gamma_1 - \dots - \Gamma_{p-1} = I_n + \sum_{j=1}^{p-1} jA_{j+1}$ and $C = \beta_{\perp}(\alpha'_{\perp}\Psi\beta_{\perp})^{-1}\alpha'_{\perp}$, we have the representation $x_t = C \sum_{j=1}^t \varepsilon_j + w_t$, ($t = 1, 2, \dots$), where w_t is a zero mean $I(0)$ process. Having this simple version of the Granger representation theorem makes our theoretical

derivations different from those used by Johansen et al. (2000) who use a more complicated extension of the Granger representation theorem.

In the following we use the lag polynomial

$$A(L) = I_n - A_1L - \dots - A_pL^p = I_n\Delta - \Pi L - \Gamma_1\Delta L - \dots - \Gamma_{p-1}\Delta L^{p-1}.$$

Notice that the relation between the two different parameterizations is given by $A_1 = I_n + \alpha\beta' + \Gamma_1$, $A_j = \Gamma_j - \Gamma_{j-1}$, ($j = 2, \dots, p-1$) and $A_p = -\Gamma_{p-1}$. Multiplying (2.1) by $A(L)$ yields

$$\begin{aligned} \Delta y_t &= \nu + \alpha(\beta'y_{t-1} - \phi(t-1) - \theta_1 b_{t-1}) \\ &\quad + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + \sum_{j=0}^{p-1} \gamma_j \Delta d_{t-j} + \eta_0 \Delta b_t + \varepsilon_t, \\ &\quad t = p+1, p+2, \dots, \end{aligned} \tag{2.5}$$

where $\nu = -\Pi\mu_0 + \Psi\mu_1$, $\phi = \beta'\mu_1$, $\theta_1 = \beta'\delta_1$, $\eta_0 = \Psi\delta_1 - \Pi\delta_0$ and

$$\gamma_j = \begin{cases} \delta_0 + \Pi\delta_0 + \Gamma_1\delta_1 + \dots + \Gamma_{p-1}\delta_1 & \text{for } j = 0, \\ -\Gamma_j\delta_0 + \Gamma_{j+1}\delta_1 + \dots + \Gamma_{p-1}\delta_1 & \text{for } j = 1, \dots, p-2, \\ -\Gamma_{p-1}\delta_0 & \text{for } j = p-1. \end{cases}$$

Notice that Δd_{t-j} is an impulse dummy with value one in period $t = \tau + j$ and zero elsewhere and $\Delta b_{t-j} = d_{t-j}$ is a shift dummy. There is no shift dummy in the long run relation in (2.5) and all but one of the differences Δb_{t-j} are omitted from (2.5) due to perfect collinearity. Note that (2.5) is a reparameterized form of equation (2.6) of Johansen et al. (2000) for the case of one break only.

For given VAR order p , our formulation of the model allows to estimate the deterministic part of the DGP as in S&L. In that procedure, first stage estimators for the parameters of the error process x_t , that is, for α , β , Γ_j ($j = 1, \dots, p-1$) and Ω are based on (2.5). A conventional RR regression of Δy_t on $(y_{t-1}, t-1, b_{t-1})$ corrected for $(1, \Delta y_{t-1}, \dots, \Delta y_{t-p+1}, \Delta d_t, \dots, \Delta d_{t-p+1}, \Delta b_t)$ may be used although that procedure does not provide exact Gaussian ML estimators because there are nonlinear relations between the parameters in (2.5). In particular, the γ_j 's are functions of the other model parameters.

It may be viewed as a drawback of our model setup that the parameters μ_0 and δ_0 are not fully identified. Although we cannot estimate them consistently, it will turn out that we can estimate them sufficiently well to obtain cointegrating rank tests with desirable properties. In the following it is assumed that all deterministic parameters including unidentified ones are estimated in a first step. The observations may then be adjusted for deterministic terms and cointegration tests are based on the adjusted series. These tests will be discussed in the next section.

3 Cointegrating Rank Tests

We wish to test the null hypothesis

$$H_0(r_0) : \text{rk}(\Pi) = r_0 \quad \text{vs.} \quad H_1(r_0) : \text{rk}(\Pi) > r_0. \tag{3.1}$$

For a given break date, S&L propose to estimate the parameters of the deterministic part first. We use their approach also here. Thus, define

$$a_{0t} = \begin{cases} 1 & \text{for } t \geq 1 \\ 0 & \text{for } t < 0 \end{cases} \quad \text{and} \quad a_{1t} = \begin{cases} t & \text{for } t \geq 1 \\ 0 & \text{for } t < 0 \end{cases}.$$

Multiplying (2.1) from the left by $A(L)$ gives

$$A(L)y_t = H_{0t}\mu_0 + H_{1t}\mu_1 + H_{2t}\delta_0 + H_{3t}\delta_1 + \varepsilon_t, \quad t = 1, 2, \dots, \quad (3.2)$$

where $y_t = 0$ for $t \leq 0$, $H_{it} = A(L)a_{it}$ ($i = 0, 1$), $H_{2t} = A(L)d_t$, and $H_{3t} = A(L)b_t$. We also introduce the matrix

$$Q = [\Omega^{-1}\alpha(\alpha'\Omega^{-1}\alpha)^{-1/2} : \alpha_{\perp}(\alpha'_{\perp}\Omega\alpha_{\perp})^{-1/2}]$$

such that $QQ' = \Omega^{-1}$.

As in S&L, we first estimate the parameters α , β , Γ_j ($j = 1, \dots, p-1$), and Ω by applying RR regression to (2.5). The resulting estimators are denoted by $\tilde{\alpha}$, $\tilde{\beta}$, $\tilde{\Gamma}_j$, and $\tilde{\Omega}$, and substituting them for the corresponding theoretical parameters gives the estimators $\tilde{A}_1 = I_n + \tilde{\alpha}\tilde{\beta}' + \tilde{\Gamma}_1$, $\tilde{A}_j = \tilde{\Gamma}_j - \tilde{\Gamma}_{j-1}$, ($j = 2, \dots, p-1$) and $\tilde{A}_p = -\tilde{\Gamma}_{p-1}$ used to define $\tilde{A}(L) = I_n - \tilde{A}_1L - \dots - \tilde{A}_pL^p$ and $\tilde{\Psi} = I_n - \tilde{\Gamma}_1 - \dots - \tilde{\Gamma}_{p-1}$. These estimators are further used to obtain $\tilde{H}_{it} = \tilde{A}(L)a_{it}$ ($i = 0, 1$), $\tilde{H}_{2t} = \tilde{A}(L)d_t$ and $\tilde{H}_{3t} = \tilde{A}(L)b_t$ as well as \tilde{Q} with obvious notation. Using these estimators we transform the feasible version of equation (3.2) to get the multivariate auxiliary regression model

$$\tilde{Q}'\tilde{A}(L)y_t = \tilde{Q}'\tilde{H}_{0t}\mu_0 + \tilde{Q}'\tilde{H}_{1t}\mu_1 + \tilde{Q}'\tilde{H}_{2t}\delta_0 + \tilde{Q}'\tilde{H}_{3t}\delta_1 + \varsigma_t, \quad t = 1, 2, \dots, T. \quad (3.3)$$

Our estimators of μ_0 , μ_1 , δ_0 and δ_1 , denoted by $\hat{\mu}_0$, $\hat{\mu}_1$, $\hat{\delta}_0$ and $\hat{\delta}_1$, respectively, are obtained from this multivariate regression model by LS. Thereby we effectively obtain a feasible GLS estimator of the parameters of the deterministic term in (2.1). Asymptotic properties of these estimators are given in the following lemma which is an analog of Theorem 2.1 of S&L. The proof is given in the Appendix.

Lemma 3.1. Under the null hypothesis $H_0(r_0)$ and the assumptions spelled out in Section 2,

- (i) $\beta'(\hat{\mu}_0 - \mu_0) = O_p(T^{-1/2})$;
- (ii) $\beta'_{\perp}(\hat{\mu}_0 - \mu_0) = O_p(1)$;
- (iii) $\beta'(\hat{\delta}_0 - \delta_0) = O_p(T^{-1/2})$;
- (iv) $\beta'_{\perp}(\hat{\delta}_0 - \delta_0) = O_p(1)$;
- (v) $\beta'(\hat{\mu}_1 - \mu_1) = O_p(T^{-3/2})$;
- (vi) $\beta'(\hat{\delta}_1 - \delta_1) = O_p(T^{-3/2})$;
- (vii) $[T^{1/2}\beta'_{\perp}(\hat{\mu}_1 - \mu_1) : T^{1/2}\beta'_{\perp}(\hat{\delta}_1 - \delta_1)] \xrightarrow{d} \beta'_{\perp}C[\zeta_1 : \zeta_2]$,

where

$$[\zeta_1 : \zeta_2] = [B(1) : B(1) - B(\lambda)] \begin{bmatrix} 1 & 1 - \lambda \\ 1 - \lambda & 1 - \lambda \end{bmatrix}^{-1}$$

and $B(s)$ is an n -dimensional Brownian motion with covariance matrix Ω and $C = \beta_{\perp}(\alpha'_{\perp}\Psi\beta_{\perp})^{-1}\alpha'_{\perp}$ as before. Moreover, all quantities converge jointly in distribution upon appropriate standardizations. \square

As in S&L, the parameters μ_0 and δ_0 are not estimated consistently in the direction of β_\perp . It suffices for our purposes, however, that the estimators are bounded in probability. Since there is no trend break in the model used by S&L, the results for δ_1 are new. The joint asymptotic distribution of $\hat{\mu}_1$ and $\hat{\delta}_1$ in the direction of β_\perp depends on the fraction or relative sample length λ where the break occurs. This fact will also be reflected in the asymptotic distribution of the test statistic for our hypothesis of interest.

Our test of the null hypothesis $H_0(r_0)$ is based on a sample analog of the series x_t obtained as

$$\hat{x}_t = y_t - \hat{\mu}_0 - \hat{\mu}_1 t - \hat{\delta}_0 d_t - \hat{\delta}_1 b_t.$$

The series \hat{x}_t can be used to compute LR type test statistics in the same way as the usual LR test statistic based on the VECM (2.4). More precisely, the test statistic can be determined from

$$\Delta \hat{x}_t = \Pi \hat{x}_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta \hat{x}_{t-j} + e_t, \quad t = p+1, \dots, T, \quad (3.4)$$

by solving the generalized eigenvalue problem $\det(\hat{\Pi} \hat{K}_T \hat{\Pi}' - \lambda \hat{\Omega}) = 0$, where $\hat{\Pi}$ is the LS estimator of Π obtained from (3.4), $\hat{\Omega}$ is the corresponding residual covariance matrix and

$$\hat{K}_T = \sum_{t=p+1}^T \hat{x}_{t-1} \hat{x}'_{t-1} - \sum_{t=p+1}^T \hat{x}_{t-1} \Delta \hat{X}'_{t-1} \left(\sum_{t=p+1}^T \Delta \hat{X}_{t-1} \Delta \hat{X}'_{t-1} \right)^{-1} \sum_{t=p+1}^T \Delta \hat{X}_{t-1} \hat{x}'_{t-1}$$

with $\Delta \hat{X}_{t-1} = [\Delta \hat{x}'_{t-1} : \dots : \Delta \hat{x}'_{t-p+1}]'$. Denoting the resulting ordered eigenvalues by $\hat{\lambda}_1 \geq \dots \geq \hat{\lambda}_n$, the LR type statistic for the pair of hypotheses in (3.1) can be shown to be

$$LR(r_0) = \sum_{j=r_0+1}^n \log(1 + \hat{\lambda}_j). \quad (3.5)$$

To distinguish it from the JMN-LR test, we will refer to the test statistic in (3.5) and the related test as GLS statistic and GLS test, respectively, in the following because the deterministic term is estimated by a feasible GLS procedure rather than ML. The limiting distribution of this test statistic under the null hypothesis is given in the following theorem, where $W(s)$ is an $(n - r_0)$ -dimensional standard Brownian motion and $\mathbf{1}(\cdot)$ denotes the indicator function. The proof is also given in the Appendix.

Theorem 3.1. Under the null hypothesis $H_0(r_0)$ and the assumptions spelled out in Section 2,

$$LR(r_0) \xrightarrow{d} \text{tr} \left\{ \left(\int_0^1 W_*(s) dW_*(s)' \right)' \left(\int_0^1 W_*(s) W_*(s)' ds \right)^{-1} \left(\int_0^1 W_*(s) dW_*(s)' \right) \right\},$$

where

$$W_*(s) = W(s) - [\xi_1 : \xi_2] \begin{bmatrix} s \\ (s - \lambda) \mathbf{1}(s > \lambda) \end{bmatrix}$$

with

$$[\xi_1 : \xi_2] = [W(1) : W(1) - W(\lambda)] \begin{bmatrix} 1 & 1 - \lambda \\ 1 - \lambda & 1 - \lambda \end{bmatrix}^{-1}$$

and $dW_*(s) = dW(s) - ds\xi_1 - \mathbf{1}(s > \lambda)ds\xi_2$ so that $\int_0^1 W_*(s)dW_*(s)'$ abbreviates $\int_0^1 W_*(s)dW(s)' - \int_0^1 W_*(s)ds\xi_1' - \int_0^1 \mathbf{1}(s > \lambda)W_*(s)ds\xi_2'$. \square

Obviously, the asymptotic distribution of the test statistic depends on the sample fraction λ of the break point. In contrast, the asymptotic distribution of the test statistic was found to be independent of the break date in S&L's setup where only a level shift is present. Thus, the dependence on the break date is introduced here exclusively due to the break in the trend slope. If that is excluded from the model, we will be back in the case considered by S&L and the asymptotic distribution of the test statistic will be independent of λ . In contrast, in the setup of Johansen et al. (2000) the asymptotic distribution of their LR statistic for the cointegrating rank will also depend on λ if only a level shift and no change in the trend slope is present. In any case, the asymptotic distribution in our Theorem 3.1 is different from the asymptotic distribution of the corresponding JMN test.

For notational convenience we have stated the asymptotic distribution for the case of one break point only. In practice one may want to allow for more than one break. Therefore it is worth noting that the result can be extended to that situation as well. To generalize the result in Theorem 3.1 it is helpful to rewrite the limiting distribution such that we can consider independent Brownian bridges confined to each of the sub-regimes. Using the indicator functions $\mathbf{1}(s \leq \lambda)$ and $\mathbf{1}(s > \lambda)$ we have for $W_*(s)$ of Theorem 3.1,

$$W_*(s) = \{W(s) - \lambda^{-1}W(\lambda)s\} \mathbf{1}(s \leq \lambda) + \left\{ W(s) - \frac{W(1) - W(\lambda)}{1 - \lambda}s + \frac{\lambda W(1) - W(\lambda)}{1 - \lambda} \right\} \mathbf{1}(s > \lambda). \quad (3.6)$$

The terms multiplied with s represent the stochastic trend slopes in the respective regimes and the term $[\lambda W(1) - W(\lambda)]/(1 - \lambda)$ ensures that the correction of $W(s)$ has the appropriate level in the second regime. To be precise, the latter term corrects for the differences between the stochastic trend slopes $W(1)$ and $[W(1) - W(\lambda)]/(1 - \lambda)$. If we plug in the regime boundaries for s , it is seen that the terms in (3.6) represent independent Brownian bridges confined to (the length of) each of the two regimes.

Using this setup, the asymptotic analysis can be generalized easily to the case of $q - 1$ breaks inducing q sub-samples. In this case, W_* can be written as the sum of q independent sub-sample Brownian Bridges. Thus, we can rewrite the limiting distribution of Theorem 3.1 similar to Theorem 3.2 of Johansen et al. (2000) and obtain for the general situation of $q - 1$ breaks

$$\text{tr} \left\{ \left(\sum_{j=1}^q D_j l_j \right)' \left(\sum_{j=1}^q P_j l_j^2 \right)^{-1} \left(\sum_{j=1}^q D_j l_j \right) \right\}, \quad (3.7)$$

where l_j ($j = 1, \dots, q$) are the relative sample lengths,

$$D_j = \int_0^1 W_*^{(j)}(s)dW_*^{(j)}(s)', \quad P_j = \int_0^1 W_*^{(j)}(s)W_*^{(j)}(s)'ds, \quad \text{and} \quad W_*^{(j)}(s) = W^{(j)}(s) - sW^{(j)}(1)$$

are independent $(n - r_0)$ -dimensional standard Brownian bridges. Hence, it follows that the limiting distribution only depends on the relative lengths of the sub-sample periods and not on their ordering. Thus, in the one-break case the relative break points λ and $1 - \lambda$ produce the same critical values because of the symmetry.

In contrast to Johansen et al. (2000) we do not need to add additional χ^2 distributed terms to the expression (3.7). In the framework of Johansen et al. (2000) these terms result from preserving the dimension of the error correction term in (2.5) while one of the sample lengths l_j tends to zero. Our test is based on the VECM (3.4) which contains the broken deterministic terms only indirectly through \hat{x}_t and not directly. Hence, the dimension of the error correction term in (3.4) does not depend on the number of breaks. Accordingly, we can work with the representation (3.7). In the next section the computation of critical values and p -values for our test is discussed based on this distribution.

4 Response Surface

Because the limiting distribution in Theorem 3.1 depends on the relative break point it is convenient to follow the response surface approach of Johansen et al. (2000) in order to derive percentage points of the distribution or p -values for the tests. The idea underlying the response surface analysis is to approximate the distribution given in (3.7) by a Gamma distribution with two parameters. It has been demonstrated by Doornik (1998) that the shape of the Johansen rank test distributions can be approximated well by Gamma distributions. This result carries over to the distributions of cointegration tests with prior adjustment of deterministic terms, as shown by Trenkler (2004). The parameters of the Gamma distribution can be related to the mean and variance of the distribution of interest. Therefore, the aim of the response surface is to provide accurate estimations of the asymptotic means and variances of the distributions of the $LR(r_0)$ statistics as a function of the number of stochastic trends under the null hypothesis, $k = n - r_0$, and the relative sample lengths. The estimated moments are used to fit approximating Gamma distributions, from which p -values or any desired quantiles can be computed.

The simulation design follows Johansen et al. (2000). We allow for up to two breaks, i.e. three sub-sample periods. For $q = 3$ sub-sample periods we have three ordered relative sample lengths $l_1 \leq l_2 \leq l_3 = 1 - l_1 - l_2$. Hence, l_3 follows from l_1 and l_2 and, therefore, l_3 need not be explicitly considered. For the case $q = 2$ we have two relative sample lengths $l_2 \leq l_3 = 1 - l_2$ and set $l_1 = 0$. Finally, if there is no break and, hence, $q = 1$ we are left with $l_3 = 1$ and set $l_1 = l_2 = 0$. We have simulated the limiting distribution of $LR(r_0)$ for different values of k , l_1 , l_2 , and sample sizes T . In line with (3.7) we generated three random walks with T steps, computed D_j and P_j and scaled the latter according to the relative sample lengths l_j ($j = 1, 2, 3$). For the simulations we used the parameter values given in Table 1. This results in 1600 different cases which were simulated $N = 100\,000$ times. Based on the N repetitions we computed the means and variances of the asymptotic distributions for the 1600 cases. The computations for the simulation study were performed by using code written in GAUSS V6 for Windows. The Monster-KISS random number generator with a fixed seed has been used to generate standard normally distributed random numbers.

As described by Johansen et al. (2000), the logarithm of the moments can be very accu-

Table 1: Simulation Details for Response Surfaces

Replications N :	100 000
Dimensions k :	1,2,...,8
Relative sample lengths (l_1, l_2) : (20 pairs)	(0, 0), (0, 0.05), (0, 0.1), (0, 0.15), (0, 0.2), (0, 0.25) (0, 0.3), (0, 0.35), (0, 0.4), (0, 0.45), (0, 0.5), (0.1, 0.1) (0.1, 0.2), (0.1, 0.3), (0.1, 0.4), (0.2, 0.2), (0.2, 0.3) (0.2, 0.4), (0.3, 0.3), (0.33, 0.33)
Sample Size T :	500/ t for $t = 1, \dots, 10$

rately approximated by a third-order polynomial in k , l_1 , l_2 , and T^{-1} given by

$$\begin{aligned} \log(\text{moment}) &\approx f_{\text{moment}}(k, l_1, l_2, T) \\ &= \sum_{h=0}^2 \left(\iota_h + \sum_{i=1}^4 \kappa_{ih} v_i + \sum_{i=1}^4 \sum_{j \geq i} \rho_{ijh} v_i v_j + \sum_{i=1}^4 \sum_{j \geq i} \sum_{k \geq j} \varphi_{ijkh} v_i v_j v_k \right) s_h, \end{aligned} \quad (4.1)$$

where $v_1 = k$, $v_2 = l_1$, $v_3 = l_2$, $v_4 = T^{-1}$, $s_h = k^{-h}$. Moreover, ι_h , κ_{ih} , ρ_{ijh} and φ_{ijkh} are parameters to be determined. Note that we have $v_1 s_1 = s_0 = 1$ and $v_1 s_2 = s_1$ such that some of the parameters of (4.1) are not identified. Therefore, they were set to zero and we are left with 75 parameters which were estimated by ordinary least squares. Using these parameter estimates the log asymptotic moments are then approximated by letting $T \rightarrow \infty$. Applying the exponential function we obtained the estimators

$$\begin{aligned} \widehat{\text{mean}} &= m = \exp\{f_{\text{mean}}(k, l_1, l_2, \infty)\} \\ \widehat{\text{variance}} &= v = \exp\{f_{\text{variance}}(k, l_1, l_2, \infty)\}. \end{aligned} \quad (4.2)$$

As mentioned earlier, the asymptotic distributions were approximated by the Gamma distribution

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

where the parameters a and b are related to the mean (μ_Γ) and variance (σ_Γ^2) of the Gamma distribution by $a = \mu_\Gamma^2 / \sigma_\Gamma^2$ and $b = \mu_\Gamma / \sigma_\Gamma$. To fit the Gamma distribution we just replaced μ_Γ and σ_Γ^2 by the estimated asymptotic moments m and v of our limiting distributions. One can also use a χ^2 distribution with non-integer degrees of freedom for the practical implementation of the approximation. The relationship between both distributions is given by $2bY \sim \chi^2(2a)$, where Y has a Gamma distribution.

As can be seen in Table 2, the fit of (4.1) for our test distribution is comparable to the approximations for the JMN test. Under the heading *Restricted Model* we present results obtained by sequentially deleting the insignificant parameters in (4.1) on the basis of t -tests. We started from the least significant parameter and applied a 5% significance level. Note that we could only eliminate a smaller number of parameters from (4.1) than in Johansen et al. (2000). In any case, the estimated standard errors are very small and can be neglected when computing critical or p -values as illustrated in Johansen et al. (2000). The response surface coefficients of the restricted models for the mean and variance are collected in Table 10 in the Appendix.

Table 2: Goodness of Fit Measures for Response Surfaces

Test	Unrestricted Model			Restricted Model		
	# Par.	R^2	$10^3\hat{\sigma}$	# Par.	R^2	$10^3\hat{\sigma}$
GLS, log mean	75	0.999993	2.84	54	0.999993	2.86
GLS, log variance	75	0.999924	8.93	63	0.999923	8.97
JMN, log mean	75	0.999998	1.15	31	0.999996	1.77
JMN, log variance	75	0.999940	6.76	24	0.999894	8.86

5 Monte Carlo Simulations

A Monte Carlo experiment was performed to analyse the finite sample properties of our new test proposal. Moreover, we compare our test with the JMN test. In order to assess the small sample effects of introducing breaks in the deterministic terms we also consider the respective standard procedures which only make allowance for a linear trend and a level term. The simulations are based on the following x_t process from Toda (1994) which was also used by a number of other authors for investigating the properties of cointegrating rank tests (see, e.g., Hubrich, Lütkepohl and Saikkonen (2001)):

$$x_t = A_1 x_{t-1} + \varepsilon_t = \begin{bmatrix} \boldsymbol{\psi} & 0 \\ 0 & I_{n-r} \end{bmatrix} x_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d. } N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} I_r & \Theta \\ \Theta' & I_{n-r} \end{bmatrix} \right), \quad (5.1)$$

where $\boldsymbol{\psi} = \text{diag}(\psi_1, \dots, \psi_r)$ and Θ are $(r \times r)$ and $(r \times (n - r))$ matrices, respectively. As shown by Toda, this type of process is useful for investigating the properties of LR tests for the cointegrating rank because other cointegrated VAR(1) processes of interest can be obtained from (5.1) by linear transformations which leave such tests invariant. Obviously, if $|\psi_i| < 1$ ($i = 1, \dots, r$) we have r stationary series and, thus, the cointegrating rank is equal to r . Hence, Θ describes the contemporaneous error term correlation between the stationary and nonstationary components. We have used two- and four-dimensional processes for simulations and report some of the results in more detail here. For given VAR order p and break date τ , the test results are invariant to the parameter values of the deterministic terms. Therefore, we use $\mu_i = 0$ and $\delta_i = 0$ ($i = 0, 1$) as parameter values throughout without loss of generality. In other words, the deterministic terms including the breaks are actually zero and, hence, $y_t = x_t$, although we take deterministic terms into account in computing the test statistics. Thereby we pretend that the analyst does not know that the deterministic terms are zero. An advantage of this specification is that we can easily compare our results with those for the standard cointegration tests within the simulation experiment. The application of the standard procedures to our process can be interpreted as taking into account a priori knowledge that there is no structural break. Thus, a comparison of the respective small sample properties should enable us to quantify the effects associated with modelling breaks in the deterministic components.

Samples are simulated by starting with initial values of zero. We have considered sample sizes of $T = 50, 100$ and 200 . Furthermore, three different relative break points given by $\lambda = 0.25, \lambda = 0.50$, and $\lambda = 0.75$ are studied. The number of replications is 5000. Thus, the standard error of an estimator of a true rejection probability P is $s_P = \sqrt{P(1 - P)/5000}$, e.g., $s_{0.05} = 0.0031$. Again, GAUSS V6 for Windows has been used for the simulations.

Table 3: Empirical Sizes of Cointegrating Rank Tests for DGP (5.1), $n = 2$, $p = 1$, $\Theta = 0$

	Panel A: $H_0(0) : r = 0$, true rank $r = 0$		Panel B: $H_0(1) : r = 1$, true rank $r = 1$, ($\psi_1 = 0.9$)		Panel C: $H_0(1) : r = 1$, true rank $r = 1$, ($\psi_1 = 0.7$)	
	$T = 50$	$T = 100$	$T = 50$	$T = 100$	$T = 50$	$T = 100$
GLS	0.0472	0.0474	0.0138	0.0204	0.0378	0.0524
GLS_{TR25}	0.0500	0.0502	0.0108	0.0170	0.0320	0.0484
GLS_{TR50}	0.0556	0.0448	0.0156	0.0170	0.0390	0.0496
GLS_{TR75}	0.0532	0.0486	0.0126	0.0172	0.0360	0.0486
JOH	0.0628	0.0552	0.0036	0.0096	0.0144	0.0466
$JMNT_{R25}$	0.0632	0.0550	0.0020	0.0044	0.0096	0.0326
$JMNT_{R50}$	0.0650	0.0616	0.0046	0.0062	0.0096	0.0316
$JMNT_{R75}$	0.0718	0.0638	0.0042	0.0064	0.0108	0.0368

Note: JOH and GLS refer to the standard Johansen and GLS test procedures, respectively, with an unrestricted linear trend and no break as proposed by Johansen (1995) and Saikkonen and Lütkepohl (2000b). GLS_{TRxx} and $JMNT_{Rxx}$ denote the GLS tests proposed in Section 3 and the corresponding JMN tests, respectively, with level shift and trend break at sample fraction $\lambda = 0.xx$.

In Tables 3 and 5 (Panel C), we present rejection frequencies for a correct null hypothesis $H_0(r_0) : r = r_0$. Hence, the rejection frequencies should give an indication of the tests' sizes in small samples. Therefore we use the term *size* in the following when we refer to this case. We always apply a significance level of 5% and the tests are not applied sequentially. Thus, the results for testing $H_0(1) : r = 1$ are not conditioned on the outcome of the test of $H_0(0) : r = 0$ etc.. Finally, we do not report size adjusted rejection frequencies related to the power of the tests because such an adjustment is not possible in applied work.

Table 3 contains the empirical sizes of the different tests for various bivariate versions of the process (5.1) with zero error term correlation ($\Theta = 0$). In general, the GLS tests have preferable size properties, especially for DGPs with one cointegration relationship and in samples of lengths $T = 50$ and $T = 100$. The consideration of breaks increases the sizes if $r = 0$ (compare Panel A). As a result, the JMN test displays some overrejection for $T = 50$ and 100. By contrast, the rejection frequencies fall for DGPs with $r = 1$ such that the problem of underrejection is worsened for smaller sample sizes (compare Panels B and C). Again, the JMN test is affected more strongly than the GLS test. Although the rejection frequencies of the latter tests are still a bit away from the nominal 5% level, they are in most cases quite substantially closer to that value than those of the JMN tests. The problem is also apparent in the original test versions proposed by Johansen (1995) and Saikkonen and Lütkepohl (2000b) for DGPs with linear trends but without a break. These test versions are denoted by *JOH* and *GLS* in Table 3. The relative break point has some impact but does not substantially alter the assessment of the tests' performances. Note that the rejection frequencies for $\lambda = 0.25$ and $\lambda = 0.75$ are not identical although they will be asymptotically. The change of the autoregressive parameter ψ_1 from 0.9 to 0.7 for processes with $r = 1$ increases the sizes. The latter reduces the size distortions. As expected, increasing the sample size clearly improves the tests' size properties.

Without showing the detailed results, we mention that the introduction of error term correlation ($\Theta \neq 0$) to bivariate processes with $r = 1$ usually produces higher empirical size values for the Johansen tests but smaller sizes for the GLS tests. Accordingly, the JMN tests may reject too often in some situations whereas the GLS tests often fall below the nominal level.

Table 4 presents the empirical powers of the cointegration tests for two-dimensional versions of the Toda process. In addition, we show power curves for $T = 100$ and a break point $\tau = 50$ in Figure 1. The parameter ψ_1 varies from 0.5 to 1 in 0.05-steps in the figure. Accordingly, the true cointegrating rank of the process is zero if $\psi_1 = 1$, otherwise, it is one. Note, that all these results are for the case where the null hypothesis $H_0(0) : r = 0$ is tested. Hence, the parameter ψ_1 can be regarded as a measure of the distance of the DGP from the null hypothesis: the smaller ψ_1 , the more we deviate from the null hypothesis. The steepness of the power curves in Figure 1 allows to evaluate the power of the tests in small samples in case of varying empirical sizes. As mentioned earlier, the latter is obtained for $\psi_1 = 1$.

Clearly, considering a trend break and a level shift reduces the finite sample power of the tests. The power loss can amount to up to one-third of the original power of the standard tests for the DGPs considered here. The GLS and the JMN tests are affected in a similar fashion.

As seen in Table 4, the GLS test outperforms the JMN test for small sample sizes if there is no error term correlation ($\Theta = 0$). This result is at least to some extent due to the better size properties of the GLS test. Moreover, if $\Theta = 0.4$, the GLS test is also more powerful in many cases. Figure 1 demonstrates that the power curves of the GLS tests are

Table 4: Empirical Powers of Cointegrating Rank Tests for DGP (5.1), $n = 2$, $p = 1$, $r = 1$, $H_0(0) : r = 0$

	Panel A: $\psi_1 = 0.9$, $\Theta = 0$			Panel B: $\psi_1 = 0.9$, $\Theta = 0.4$		
	$T = 50$	$T = 100$	$T = 200$	$T = 50$	$T = 100$	$T = 200$
<i>GLS</i>	0.0690	0.1286	0.4030	0.0790	0.1618	0.5168
<i>GLS</i> _{<i>TR</i>25}	0.0660	0.0998	0.2776	0.0692	0.1170	0.3452
<i>GLS</i> _{<i>TR</i>50}	0.0644	0.0970	0.2744	0.0656	0.1134	0.3378
<i>GLS</i> _{<i>TR</i>75}	0.0666	0.1146	0.3018	0.0730	0.1300	0.3876
<i>JOH</i>	0.0670	0.1041	0.2848	0.0710	0.1324	0.3894
<i>JMN</i> _{<i>TR</i>25}	0.0698	0.0888	0.1892	0.0682	0.1030	0.2490
<i>JMN</i> _{<i>TR</i>50}	0.0692	0.0828	0.1760	0.0750	0.0092	0.2354
<i>JMN</i> _{<i>TR</i>75}	0.0760	0.0924	0.1966	0.0756	0.1044	0.2590

	Panel C: $\psi_1 = 0.7$, $\Theta = 0$			Panel D: $\psi_1 = 0.7$, $\Theta = 0.4$		
	$T = 50$	$T = 100$	$T = 200$	$T = 50$	$T = 100$	$T = 200$
<i>GLS</i>	0.2356	0.6844	0.9754	0.2944	0.7760	0.9862
<i>GLS</i> _{<i>TR</i>25}	0.1726	0.5220	0.9374	0.2062	0.6242	0.9586
<i>GLS</i> _{<i>TR</i>50}	0.1692	0.5170	0.9374	0.2018	0.6126	0.9620
<i>GLS</i> _{<i>TR</i>75}	0.1810	0.5474	0.9520	0.2236	0.6372	0.9716
<i>JOH</i>	0.1892	0.6316	0.9988	0.2498	0.7724	1.0000
<i>JMN</i> _{<i>TR</i>25}	0.1490	0.4392	0.9724	0.1796	0.5630	0.9960
<i>JMN</i> _{<i>TR</i>50}	0.1306	0.3964	0.9690	0.1632	0.5300	0.9928
<i>JMN</i> _{<i>TR</i>75}	0.1550	0.4426	0.9756	0.1894	0.5710	0.9942

Note: *JOH* and *GLS* refer to the standard Johansen and GLS test procedures, respectively, with an unrestricted linear trend and no break as proposed by Johansen (1995) and Saikkonen and Lütkepohl (2000b). *GLS*_{*TR**xx*} and *JMN*_{*TR**xx*} denote the GLS tests proposed in Section 3 and the corresponding JMN tests, respectively, with level shift and trend break at sample fraction $\lambda = 0.xx$.

somewhat steeper than those of the JMN procedures for values of ψ_1 close to one, i.e. for alternatives close to the null hypothesis. However, the power curves of the Johansen tests become steeper if ψ_1 is small. As a result, the JMN test has higher small sample power for $\psi_1 < 0.55$ for no error term correlation (Figure 1, Panel A) and $\psi_1 < 0.60$ if $\Theta = 0.4$ (Figure 1, Panel B). Of course, having good power close to the null hypothesis is of particular importance because it is the region of the parameter space where it is typically most difficult to discriminate between the null and alternative hypotheses. Hence, applying the GLS test is a good strategy at least for the processes considered in our Monte Carlo simulations.

The general power level of the tests increases with the sample size, the magnitude of the error term correlation, and with falling values of ψ_1 . The latter should be no surprise, since ψ_1 measures the distance of the DGP from the null hypothesis. However, this also means that the power of the tests can be very low for values of ψ_1 close to one, especially in case of very small samples and zero or weak error term correlation. As in case of the tests' sizes, the location of the break has some effect on the small sample power of the tests. The power tends to be higher for breaks towards the end of the sample than for those closer to the

Table 5: Rejection Frequencies of Cointegrating Rank Tests for DGP (5.1), $n = 4$, $p = 1$, $\Theta = [(0.4 \ 0.4)']$, $r = 2$ ($\psi_1 = 0.7$, $\psi_2 = 0.7$)

	Panel A: $H_0(0) : r = 0$				Panel B: $H_0(1) : r = 1$				Panel C: $H_0(2) : r = 2$			
	$T = 50$		$T = 200$		$T = 50$		$T = 200$		$T = 50$		$T = 200$	
	$T = 100$	$T = 200$	$T = 100$	$T = 200$	$T = 100$	$T = 200$	$T = 100$	$T = 200$	$T = 100$	$T = 200$	$T = 100$	$T = 200$
<i>GLS</i>	0.5626	0.9808	1.0000	1.0000	0.1306	0.4598	0.9328	0.9328	0.0192	0.0356	0.0482	0.0482
<i>GLS_{TR25}</i>	0.4050	0.9264	0.9996	0.9996	0.0900	0.3264	0.8384	0.8384	0.0124	0.0368	0.0394	0.0394
<i>GLS_{TR50}</i>	0.3998	0.9210	0.9998	0.9998	0.0918	0.3060	0.8250	0.8250	0.0168	0.0274	0.0402	0.0402
<i>GLS_{TR75}</i>	0.4260	0.9334	1.0000	1.0000	0.0982	0.3184	0.8592	0.8592	0.0156	0.0258	0.0396	0.0396
<i>GLS_{SH25}</i>	0.5234	0.9688	1.0000	1.0000	0.1368	0.4368	0.9166	0.9166	0.0214	0.0352	0.0476	0.0476
<i>GLS_{SH50}</i>	0.5236	0.9718	1.0000	1.0000	0.1356	0.4372	0.9240	0.9240	0.0206	0.0394	0.0502	0.0502
<i>GLS_{SH75}</i>	0.5324	0.9690	1.0000	1.0000	0.1352	0.4324	0.9226	0.9226	0.0224	0.0416	0.0462	0.0462
<i>JOH</i>	0.5882	0.9922	1.0000	1.0000	0.1128	0.4420	0.9682	0.9682	0.0134	0.0426	0.0608	0.0608
<i>JMNT_{TR25}</i>	0.4694	0.9520	1.0000	1.0000	0.0842	0.3142	0.8770	0.8770	0.0082	0.0302	0.0564	0.0564
<i>JMNT_{TR50}</i>	0.4644	0.9482	1.0000	1.0000	0.0814	0.2986	0.8668	0.8668	0.0098	0.0268	0.0616	0.0616
<i>JMNT_{TR75}</i>	0.4880	0.9572	1.0000	1.0000	0.0898	0.3098	0.8768	0.8768	0.0106	0.0314	0.0608	0.0608

Note: *JOH* and *GLS* refer to the standard Johansen and GLS test procedures, respectively, with an unrestricted linear trend and no break as proposed by Johansen (1995) and Saikkonen and Lütkepohl (2000b). *GLS_{TRxx}* and *JMNT_{TRxx}* denote the GLS tests proposed in Section 3 and the corresponding JMN tests, respectively, with level shift and trend break at sample fraction $\lambda = 0.xx$. *GLS_{SHxx}* denotes the corresponding test version proposed by S&L which accounts for a level shift only.

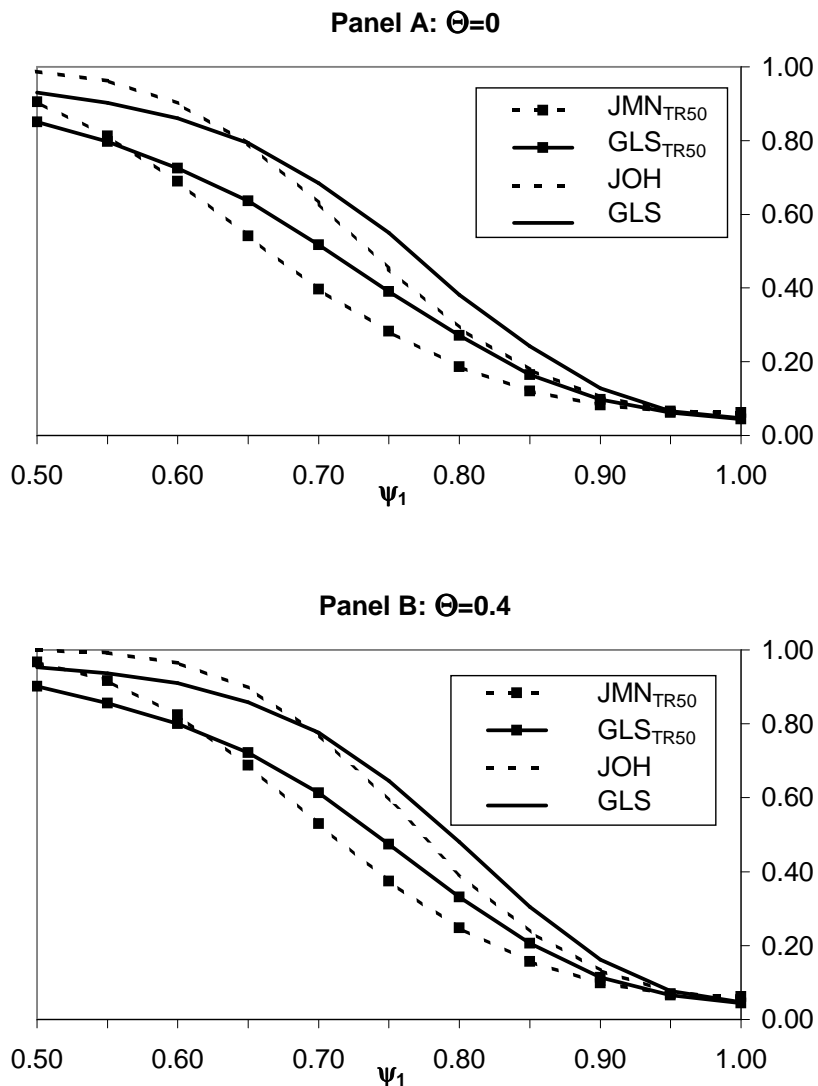


Figure 1: Relative rejection frequencies of null hypothesis $H_0(0) : r = 0$ for bivariate DGPs with $r = 0$ ($\psi_1 = 1$) or $r = 1$ ($\psi_1 < 1$), sample size $T = 100$, true break point $\tau = 50$, and nominal significance level 0.05.

beginning of the sample (compare the TR75 and TR25 results).

We have also studied four-dimensional DGPs with two cointegrating vectors for different autoregressive parameters ψ_1 and ψ_2 and covariance matrices Θ . Some results are shown in Table 5. They are generally in line with those for the bivariate processes. In particular, none of the tests is uniformly superior for all situations considered. For comparison purposes we also present results for the case considered by S&L where just a level shift is accounted for. These tests are denoted by GLS_{SHxx} in Table 5. They have clearly substantially better power than the corresponding GLS and JMN tests which account for a break in the trend slope. This result shows that it is not a good idea to include trend breaks in the test when just a level shift has occurred.

Finally, we have analyzed the finite sample performance of the tests in large-dimensional systems with many parameters to be estimated. To this end, we have again simulated four-dimensional VAR(1) processes but we have fitted higher lag orders when performing the tests.

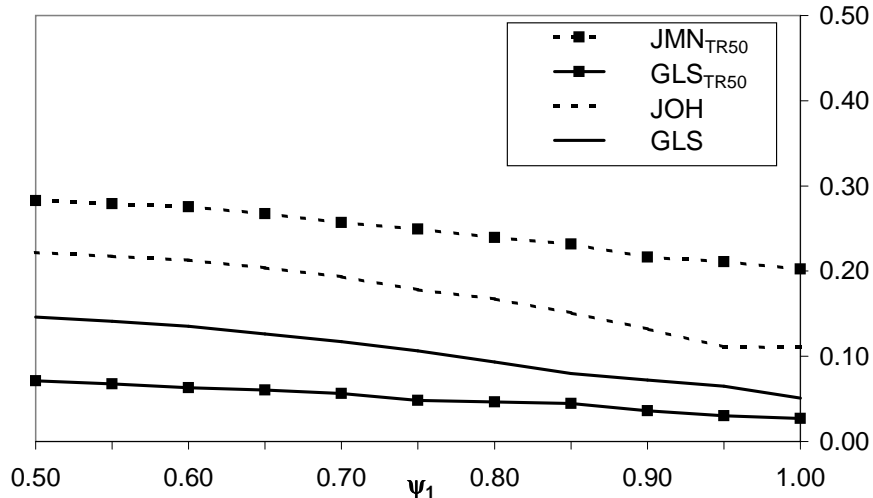


Figure 2: Relative rejection frequencies of null hypothesis $H_0(0) : r = 0$ for four-dimensional DGPs with $r = 0$ ($\psi_1 = 1$) or $r = 1$ ($\psi_1 < 1$), sample size $T = 100$, true break point $\tau = 50$, $\Theta = [(0.4 \ 0.4)' : (0.4 \ 0.4)']$, and nominal significance level 0.05. The fitted lag order is five.

Thereby, we are able to check the reaction of the tests to increased estimation uncertainty due to higher lag orders. As an example, Figure 2 displays the tests' rejection frequencies for a specific four-dimensional DGP with one autoregressive lag and $r = 0$ ($\psi_1 = 1$) or $r = 1$ ($\psi_1 < 1$). The fitted lag order for all tests is five. Obviously, the Johansen tests, especially the JMN version with a trend break, display excessive size distortions. This kind of size distortion in large systems has already been pointed out by Gonzalo and Pitarakis (1999) even for the case of no breaks and only one fitted autoregressive lag. Our simulation results indicate that this dimensionality effect gets much worse for the Johansen procedure if a break in the linear trend and if higher lag orders are considered. In contrast, the GLS test without break has roughly a correct empirical size of about 5%, whereas the test version with a break is somewhat conservative. Nevertheless, the slopes of the power curves are approximately equal for the tests with a break and without breaks, respectively. The power curves are very flat, however, which makes it difficult to draw proper conclusions regarding the cointegrating rank in the present situation. Although this is a slightly negative prospect for empirical work, it should perhaps be taken as encouragement to use both tests simultaneously in the difficult situation where breaks have occurred. Accounting for size distortions neither the GLS nor the JMN test has small sample power advantages for the current situation. An application of the tests is discussed in the next section.

6 Empirical Application: The Great Ratios

Using the GLS and JMN procedures we test in the following whether the so-called “great ratios” of consumption and investment to output are stationary for the German economy. This kind of research work was pioneered by King, Plosser, Stock and Watson (1991). They refer to a standard RBC model which suggests that the economy converges to a balanced growth path. The convergence behaviour implies that the logarithms of per capita output, consumption and investment are driven by a common stochastic trend such that consump-

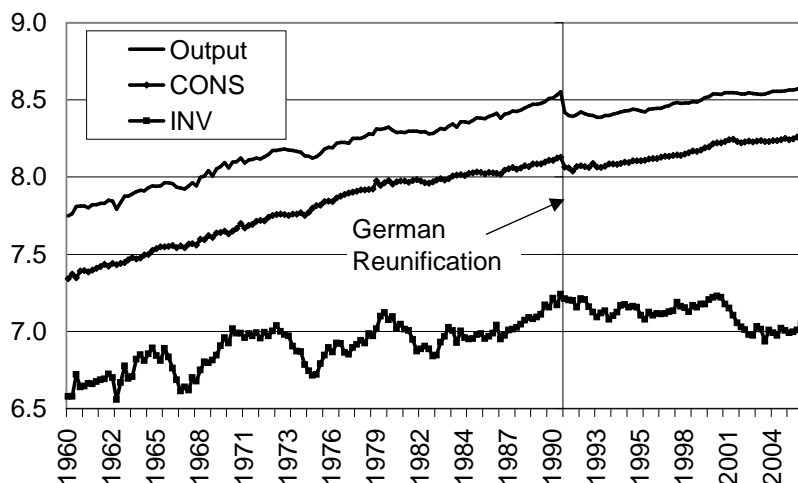


Figure 3: Logarithms of German real per capita private output, consumption, and investment between 1960:1-2005:4. Data are seasonally adjusted.

tion and investment form one-to-one cointegration relationships with output resulting in stationary ratios.

While King et al. (1991) found clear evidence for stationary great ratios in the U.S. between 1949 and 1988 the results for other periods and countries are mixed (for references to other studies see Attfield and Temple (2003)). It has been argued, for instance, by Attfield and Temple (2003) and Clemente, Montañés and Ponz (1999) that the great ratios may have been affected by structural breaks. Then the distortions emerging from ignoring these breaks may have been the reasons for the rather weak evidence for stationary great ratios. In fact, allowing for a level shift Attfield and Temple (2003) found two cointegrating relationships with JMN tests for the UK and U.S. during the period from 1955 to 2002. However, they rejected the unit coefficient restriction at the 5% level. Clemente et al. (1999) obtained stationarity of the great ratios for a larger number of OECD countries between 1959 and 1995 once they allowed for level shifts within the ADF unit root testing framework. Unfortunately, there were also countries for which stationarity of the ratios could not be concluded. Specifically, the consumption-output ratio for Germany may not be stationary. Clearly, the German reunification, that is, the aggregation of the East and West German data has led to a structural break. In the following we will analyze the possibility of stationary great ratios for Germany taking into account the reunification break.

We use quarterly data for the logarithms of German real per capita private output, consumption and investment spanning the period 1960:1 to 2005:4. That is, we have 180 observations. The series are computed as follows. Nominal private output is obtained by subtracting government consumption expenditures from GDP as in King et al. (1991). Nominal consumption is defined as household consumption expenditures and nominal investment as the sum of gross fixed capital formation and changes in inventories. Then, the implicit GDP deflator was used to get real series since no long-term series of the real aggregate series were directly available. In contrast to Attfield and Temple (2003) we applied the GDP deflator instead of the consumption-based price deflator because the former is a plausible measure for the overall price level of the three series. In order to obtain per capita quantities we divided the deflated series by quarterly population figures which were obtained by a log-

linear interpolation of the yearly population series. Finally, the logarithmic transformation was applied. The data are seasonally adjusted and refer to West Germany until 1990:4 and afterwards to reunified Germany. Graphs are presented in Figure 3. All data series are from the International Financial Statistics and were provided by the Financial and Economic Data Center of the Collaborative Research Center “Economic Risk” at the Humboldt-Universität zu Berlin. The computations were done with the software program JMulTi (Lütkepohl and Krätzig (2004)) and own programs written in GAUSS.

There are obviously strong reasons why the German reunification may have caused a structural break in the analyzed time series (see also Figure 3). Hence, when testing for cointegration we allow for a break in the level and the linear trend component of the series in the first quarter of 1991. Compared to Attfield and Temple (2003) and Clemente et al. (1999) we are more general by modelling a trend break in addition to simple level shifts. The reason for doing so is that the reunification has resulted in changes in the growth rates after the reunification in the per capita series because of the grossly different economic conditions in East and West Germany.

We started our empirical analysis by determining the integration order of the logarithms of German real per capita private output, consumption and investment. The generalized ADF test suggested by Perron (1989), which allows for a level shift and a trend break in 1991:1, indicated that all three series can be regarded as $I(1)$. We do not show the detailed results but work under the assumption of $I(1)$ series in the following.

Table 6: ADF Unit Root Test Results for Consumption-Output and Investment-Output Ratios 1960:1-2005:4

Consumption-Output Ratio				Investment-Output Ratio			
Lags	C	C, T	Breaks, C, T	Lags	C	C, T	Breaks, C, T
2	-2.033 [0.273] HQ	-2.664 [0.253] HQ	-1.510 HQ	0	—	-3.033 [0.126] HQ	-4.074* HQ
4	—	-3.098 [0.110] AIC	-2.107 AIC	2	-1.779 [0.390] AIC, HQ	—	—
5	-1.772 [0.393] AIC	—	—	5	—	-3.462** [0.047] AIC	-4.690** AIC

Note: C and T stand for constant and trend, respectively. Breaks refers to break in 1991:1 in level and trend, i.e., Perron (1989) test is conducted. The critical values for relative break point $\lambda = 0.7$ are -4.75 (1%), -4.18 (5%), and -3.86 (10%) Perron (1989, Table VI.B). ** and * indicate significance at the 5% and 10% level, respectively. For the ADF tests, p -values are given in brackets. Lag lengths suggested by AIC and HQ information criteria for respective deterministic terms.

We first checked the stationarity properties of the univariate consumption-output and investment-output ratios (log consumption - log output and log investment - log output, respectively). The results of unit root tests are presented in Table 6. Augmented Dickey-Fuller (ADF) tests with and without trend break are given in the table. The lag orders were determined by the Hannan-Quinn (HQ) and Akaike (AIC) model selection criteria.

Table 7: Results of Cointegration Tests (p -values) for Consumption-Output Ratio 1960:1-2005:4

$H_0(r_0)$	VAR(3)		VAR(5)		VAR(3)		VAR(5)	
	<i>JOH</i>	<i>GLS</i>	<i>JOH</i>	<i>GLS</i>	$JMN_{1991:1}$	$GLS_{1991:1}$	$JMN_{1991:1}$	$GLS_{1991:1}$
$r_0 = 0$	0.176	0.462	0.188	0.227	0.789	0.493	0.259	0.188
$r_0 = 1$	0.277	0.919	0.148	0.964	0.783	0.972	0.907	0.986

Note: *JOH* and *GLS* refer to the standard Johansen and GLS tests, respectively, with an unrestricted linear trend. No breaks in the deterministic terms are allowed for. The p -values for the Johansen and GLS tests are computed according to the response surfaces of Doornik (1998) and Trenkler (2004), respectively.

$JMN_{1991:1}$ and $GLS_{1991:1}$ refer to the JMN and GLS tests, respectively, with an unrestricted linear trend and a level shift and trend break in 1991:1. The p -values for the JMN and GLS tests are computed according to the response surfaces of Johansen et al. (2000) and Section 4, respectively. The break date 1991:1 translates to relative sample lengths of 0.679 and 0.321. The latter is chosen for the p -value determination since it is the shorter one.

In addition to tests with constant and trend and a break in the trend we also present results of tests with just a constant as deterministic term because tests with over specified deterministic terms may lack power. In our case the results are quite clear. A unit root in the consumption-output ratio cannot be rejected by any of the tests. In contrast, a unit root is rejected for the investment-output ratio at least if a trend break is allowed for and even in one case when just a constant (C) and a trend (T) are included without allowing for a break. Our analysis could stop here and we could conclude that we did not find evidence for a stable consumption-output ratio in Germany during the sample period. It is possible, however, that consumption and output are still driven by a common stochastic trend and, hence, that there is a cointegration relation between the two series although the difference between the two series is not $I(0)$.

To investigate this possibility we have performed the systems cointegration tests which were discussed in the previous sections. Some results for the bivariate consumption-output system are presented in Table 7. Again results for tests with and without allowance for trend breaks are given. The lag orders are those suggested by AIC and HQ as in the univariate tests. Also the multivariate tests provide no evidence of a possible cointegration relation. More precisely, a cointegrating rank of zero cannot be rejected by any of the cointegration tests at a 10% level of significance. Of course, this finding could be a consequence of the low power of the tests especially when the lag order is not very small.

To get a better feeling for the properties of the multivariate tests and to illustrate their performance in practice we also applied them to the bivariate investment-output system. Some p -values are shown in Table 8. In this case we hope, of course, to reject $H_0(0) : r = 0$ because a cointegration relation was already found by the previous unit root tests. For this system the rank zero null hypothesis can indeed be rejected at the 5% level by both the JMN and the GLS tests if a trend break in 1991:1 is allowed for at least if the HQ lag order one is considered. On the other hand, the tests which do not allow for the trend break cannot confirm the cointegration relation. These results illustrate the importance of taking into account a break in the deterministic term if such a break exists and they are also well

Table 8: Results of Cointegration Tests (p -values) for Investment-Output Ratio 1960:1-2005:4

$H_0(r_0)$	VAR(1)		VAR(6)		VAR(1)		VAR(5)	
	<i>JOH</i>	<i>GLS</i>	<i>JOH</i>	<i>GLS</i>	$JMN_{1991:1}$	$GLS_{1991:1}$	$JMN_{1991:1}$	$GLS_{1991:1}$
$r_0 = 0$	0.297	0.352	0.118	0.200	0.007	0.022	0.141	0.584
$r_0 = 1$	0.459	0.817	0.642	0.948	0.250	0.402	0.696	0.517

Note: see note to Table 7.

Table 9: Results of Cointegration Tests (p -values) for German Macro Data 1960:1-2005:4

$H_0(r_0)$	VAR(4)		VAR(5)		VAR(4)		VAR(5)	
	<i>JOH</i>	<i>GLS</i>	<i>JOH</i>	<i>GLS</i>	$JMN_{1991:1}$	$GLS_{1991:1}$	$JMN_{1991:1}$	$GLS_{1991:1}$
$r_0 = 0$	0.341	0.450	0.102	0.109	0.307	0.047	0.091	0.065
$r_0 = 1$	0.338	0.947	0.281	0.965	0.391	0.834	0.123	0.705
$r_0 = 2$	0.281	1.000	0.212	0.958	0.255	0.947	0.269	0.961

Note: see note to Table 7.

in line with our simulation results which indicated that increasing the lag order may reduce the power of the cointegration tests dramatically. For the lag order 5 which was suggested by AIC for the models with break, our tests cannot reject the rank zero hypothesis at the 10% level.

As a further check we have applied our cointegrating rank tests to the three-dimensional system consisting of all three variables. Some results are given in Table 9. Again tests with and without allowance for a trend break are presented and again the tests which do not account for the break cannot reject the rank zero hypothesis at the 10% level. In contrast the JMN and our GLS tests with trend breaks can reject cointegrating rank zero at the 10% level at least for one of the two lag orders shown in Table 9. In fact, our new GLS test is the only one which can reject rank zero for both lag orders 4 and 5. This finding is in line with the simulation result that the tests can have quite different power for particular DGPs and underscores the previous conclusion that applying both tests may be beneficial in practice.

In Table 9 it is also apparent that the tests do not find more than one cointegration relation between the three series. A cointegrating rank of $r = 1$ is of course fully consistent with the previous unit root and bivariate cointegration testing results which indicated a cointegration relation between investment and output and no such relation between consumption and output. Had we found a second cointegration relation between the three series there would necessarily also be a cointegration relation between all pairs of two series because in that case all three series are driven by a common stochastic trend. Although such a result was not expected on the basis of the previous tests, it is not totally impossible that taking the information in all three series into account enables the tests to reject cointegrating rank one even if univariate and bivariate analyses do not suggest such an outcome. Unfortunately, in our example we do not find a second cointegration relation at least at common significance levels and, hence, there is not much evidence that the great ratios were both stable for Germany during the sampling period from 1960 to 2005.

Our cointegration test results are in line with findings of Clemente et al. (1999) regarding

the failure of a stationary consumption-output ratio and of D’Adda and Scorcu (2003) with respect to a nonstable capital-output ratio for Germany. Thus, serious doubts have to be raised regarding the validity of the real business cycle theory for Germany.

7 Conclusions and Extensions

In this paper we have considered cointegrating rank tests for VAR processes with a break in the deterministic trend component at a known time point. In contrast to other tests which accommodate trend breaks in the present framework, we propose to estimate the deterministic terms first and adjust the series for deterministic terms including the break. A Johansen type cointegrating rank test is then applied to the adjusted series. We have derived the asymptotic distribution of the test statistic under the null hypothesis. Because the null distribution depends on the relative fraction of the sample where the break occurs, we have provided response surfaces to obtain critical values for the test and to approximate p -values. It is shown by a Monte Carlo study that our test has better small sample properties than existing other tests in many situations. In particular, the size distortion is often smaller than for the corresponding JMN test and the power is comparable or even higher especially close to the null hypothesis. Because there are also situations where the JMN test has better power and in some situations of practical relevance none of the tests has very attractive power properties, we propose to use both tests simultaneously and base a decision on the number of cointegration relations in a system of interest on both of them. An investigation of the stability of the “great ratios” for Germany illustrates the appeal of the new test for applied work.

There are a number of possible extensions which may be desirable for applied work. First of all, there may be more than one break in the deterministic trend function during the sample period. Although we have focussed the theoretical derivations on the case of a single break point to avoid more complicated expressions, we have also presented the necessary modifications if there is more than one break and our response surface results refer to the more general case. In principle, the theory for this case can be handled in an analogous way as the case of one break. The notation will become more complicated, however. Also, in addition to the deterministic terms considered in the present paper, further dummy variables may be included. More precisely, impulse dummies and seasonal dummies can be included in the model without affecting the asymptotic results. The necessary modifications for these extensions are straightforward.

There is at least one potentially interesting extension of our results which is nontrivial, namely the case of an unknown break point. Although we have argued that the case of a known break point is at least equally relevant and although we have provided an example where the break point is in fact known, there may be situations where estimating the break point by some statistical procedure may be desirable. We leave this case for future research.

Appendix

A.1 Response Surface Coefficients

The response surface coefficients for computing critical values and p -values for our tests are given in Table 10.

Table 10: Estimated Response Surface Coefficients for Mean and Variance

	mean	variance
constant	2.4402237	2.2377192
k	0.56642166	0.67248661
l_1	1.6881464	-1.8645617
l_2	-0.16741988	1.5842396
k^2	-0.036711384	-0.043986793
$k \cdot l_1$	-0.12654483	-
$k \cdot l_2$	0.028632527	-0.24851423
l_1^2	-7.2612954	12.095382
$l_1 \cdot l_2$	-1.9837337	5.0821793
l_2^2	-1.6794244	-1.5583336
k^3	0.0011810636	0.0012910484
$k^2 \cdot l_1$	0.0043692769	0.010518609
$k^2 \cdot l_2$	-0.0013398893	0.013510933
$k \cdot l_1^2$	0.18296009	-0.47646731
$k \cdot l_1 \cdot l_2$	0.029314412	-0.24048797
$k \cdot l_2^2$	0.030349768	0.089839081
l_1^3	11.803034	-22.104882
$l_1^2 \cdot l_2$	-2.4870918	7.7658803
$l_1 \cdot l_2^2$	4.0200467	-8.7651217
l_2^3	2.1430130	-0.33556879
$1/k$	-3.0135200	-1.6752679
l_1/k	1.1124296	11.709656
l_2/k	5.1272149	-1.8671894
l_1^2/k	4.3452158	-60.229949
$(l_1 \cdot l_2)/k$	3.5022236	-10.142186
l_2^2/k	-8.6822664	4.5029279
l_1^3/k	-16.767237	129.75575
$l_1^2 \cdot l_2/k$	5.9727547	-58.276995
$l_1 \cdot l_2^2/k$	-7.0978257	32.313807
l_2^3/k	5.7110493	-
$1/k^2$	1.0331268	0.29558742
l_1/k^2	-0.64788931	-4.9775552
l_2/k^2	-2.9655130	4.3265064
l_1^2/k^2	-	30.965573
l_2^2/k^2	7.6083137	-14.418641
l_1^3/k^2	5.7695930	-82.599414
$(l_1^2 \cdot l_2)/k^2$	-6.5947593	48.316674
$(l_1 \cdot l_2^2)/k^2$	-	-15.333499
l_2^3/k^2	-6.9391802	10.881697

A.2 Proofs

Proof of Lemma 3.1

The proof of Lemma 3.1 is similar to the proof of Theorem 2.1 of S&L. So several details will be skipped. First note that the results given in Lemma 2.1 of S&L also hold in the present context. In other words, upon appropriate normalization, the RR estimator $\tilde{\beta}$ is consistent of order $O_p(T^{-1})$ whereas $\tilde{\alpha}$, $\tilde{\Gamma}_j$ ($j = 1, \dots, p-1$), and $\tilde{\Omega}$ are consistent of order $O_p(T^{-1/2})$. This can be proved in the same way as the aforementioned previous result or Lemma A.3 of Johansen et al. (2000). In subsequent derivations all relevant quantities will be invariant to normalizations of $\tilde{\alpha}$ and $\tilde{\beta}$ so that we can assume that some kind of normalization has been imposed.

As in the proof of Theorem 2.1 of S&L it follows from definitions that

$$\tilde{H}_{0t} = \begin{cases} I_n, & t = 1, \\ I_n - \sum_{j=1}^{t-1} \tilde{A}_j, & t = 2, \dots, p, \\ -\tilde{\alpha}\tilde{\beta}', & t = p+1, \dots, T, \end{cases}$$

$$\tilde{H}_{1t} = \begin{cases} I_n, & t = 1, \\ tI_n - \sum_{j=1}^{t-1} (t-j) \tilde{A}_j, & t = 2, \dots, p, \\ \tilde{\Psi} - (t-1)\tilde{\alpha}\tilde{\beta}', & t = p+1, \dots, T, \end{cases}$$

$$\tilde{H}_{2t} = \begin{cases} 0, & t < \tau, \\ I_n, & t = \tau, \\ I_n - \sum_{j=1}^{t-\tau} \tilde{A}_j, & t = \tau+1, \dots, \tau+p-1, \\ -\tilde{\alpha}\tilde{\beta}', & t = \tau+p, \dots, T, \end{cases}$$

and

$$\tilde{H}_{3t} = \begin{cases} 0, & t < \tau, \\ I_n, & t = \tau, \\ I_n - \sum_{j=1}^{t-\tau} (t-\tau+1-j) \tilde{A}_j, & t = \tau+1, \dots, \tau+p-1, \\ \tilde{\Psi} - (t-\tau)\tilde{\alpha}\tilde{\beta}', & t = \tau+p, \dots, T. \end{cases}$$

Define

$$\underline{\gamma}_1 = \begin{bmatrix} \underline{\gamma}_{11} \\ \underline{\gamma}_{12} \end{bmatrix} = \begin{bmatrix} \tilde{\beta}'_{\perp} \mu_0 \\ \tilde{\beta}'_{\perp} \delta_0 \end{bmatrix},$$

$$\underline{\gamma}_2 = \begin{bmatrix} \underline{\gamma}_{21} \\ \underline{\gamma}_{22} \\ \underline{\gamma}_{23} \\ \underline{\gamma}_{24} \end{bmatrix} = \begin{bmatrix} \tilde{\beta}' \mu_0 \\ \tilde{\beta}' \delta_0 \\ \tilde{\beta}' \mu_1 \\ \tilde{\beta}' \delta_1 \end{bmatrix},$$

and

$$\underline{\gamma}_3 = \begin{bmatrix} \underline{\gamma}_{31} \\ \underline{\gamma}_{32} \end{bmatrix} = \begin{bmatrix} \tilde{\beta}'_{\perp} \mu_1 \\ \tilde{\beta}'_{\perp} \delta_1 \end{bmatrix}.$$

As in S&L, the idea is to first obtain asymptotic properties of LS estimators of these three “parameter vectors”. To express (3.3) in terms of $\underline{\gamma}_1$, $\underline{\gamma}_2$ and $\underline{\gamma}_3$ we transform the matrices \tilde{H}_{it} ($i = 0, 1, 2, 3$) accordingly. Thus, define

$$\tilde{G}_{1t} = \tilde{Q}' \left[\tilde{H}_{0t} \tilde{\beta}_{\perp} : \tilde{H}_{2t} \tilde{\beta}_{\perp} \right],$$

$$\tilde{G}_{2t} = \tilde{Q}' \left[\tilde{H}_{0t} \tilde{\beta} : \tilde{H}_{2t} \tilde{\beta} : \tilde{H}_{1t} \tilde{\beta} : \tilde{H}_{3t} \tilde{\beta} \right]$$

and

$$\tilde{G}_{3t} = \tilde{Q}' \left[\tilde{H}_{1t} \tilde{\beta}_\perp : \tilde{H}_{3t} \tilde{\beta}_\perp \right],$$

where $\tilde{\beta} = \tilde{\beta}(\tilde{\beta}'\tilde{\beta})^{-1}$ and similarly for $\tilde{\beta}_\perp$ as well as the corresponding parameters to be used below. With this notation (3.3) can be written as

$$\tilde{Q}' \tilde{A}(L) y_t = \tilde{G}_{1t} \underline{\gamma}_1 + \tilde{G}_{2t} \underline{\gamma}_2 + \tilde{G}_{3t} \underline{\gamma}_3 + \varsigma_t, \quad t = 1, 2, \dots, T. \quad (A.1)$$

The LS estimators of $\underline{\gamma}_1$, $\underline{\gamma}_2$ and $\underline{\gamma}_3$ obtained from this equation are denoted by $\hat{\underline{\gamma}}_1$, $\hat{\underline{\gamma}}_2$ and $\hat{\underline{\gamma}}_3$, respectively. Arguments needed to obtain their asymptotic properties are very similar to those used in the case of equation (A.12) of S&L. To see this, note that the deterministic part in the present model (2.1) differs from its counterpart, equation (1.1) of S&L, only in two respects. First, the impulse dummy used in S&L has been dropped and, second, the variable b_t has been included. Dropping the impulse dummy is clearly immaterial and, as far as rates of convergence are concerned, the variable b_t behaves in the same way as the trend term t (assuming condition (2.2)). Thus, observing that \tilde{G}_{1t} takes nonzero values only for a fixed number of time indices t we can proceed in the same way as in the proof of Theorem 2.1 of S&L and conclude that the appropriately standardized moment matrix in the aforementioned LS estimation is asymptotically block diagonal between \tilde{G}_{1t} and $[\tilde{G}_{2t} : \tilde{G}_{3t}]$ and that

$$\hat{\underline{\gamma}}_1 = \underline{\gamma}_1 + O_p(1), \quad \Upsilon_T(\hat{\underline{\gamma}}_2 - \underline{\gamma}_2) = O_p(1) \quad \text{and} \quad T^{1/2}(\hat{\underline{\gamma}}_3 - \underline{\gamma}_3) = O_p(1), \quad (A.2)$$

where $\Upsilon_T = \text{diag}[T^{1/2}I_{2r} : T^{3/2}I_{2r}]$. As in the case of Theorem 2.1 of S&L we can conclude from the first two results in (A.2) that (i) - (vi) of Lemma 3.1 hold. Thus, we are left with (vii).

The aforementioned asymptotic block diagonality of the appropriately standardized moment matrix in the LS estimation of (A.1) implies that we can drop the first term on the right hand side of (A.1) when studying asymptotic properties of the LS estimators of $\underline{\gamma}_2$ and $\underline{\gamma}_3$. Further, deriving an explicit expression for the error term ς_t (cf. (A.13) of S&L) it can be shown that the asymptotic distributions of the LS estimators $\hat{\underline{\gamma}}_2$ and $\hat{\underline{\gamma}}_3$ can be obtained by ignoring errors due to using the estimators $\tilde{\alpha}$, $\tilde{\beta}$, $\tilde{\Gamma}_j$ and $\tilde{\Omega}$ instead of their true values. This means that we can proceed by replacing ς_t by $Q'\varepsilon_t$ and obtain

$$\begin{aligned} & \begin{bmatrix} \Upsilon_T(\hat{\underline{\gamma}}_2 - \underline{\gamma}_2) \\ T^{1/2}(\hat{\underline{\gamma}}_3 - \underline{\gamma}_3) \end{bmatrix} \\ &= \begin{bmatrix} \Upsilon_T^{-1} \sum_{t=1}^T \tilde{G}'_{2t} \tilde{G}_{2t} \Upsilon_T^{-1} & T^{-1/2} \Upsilon_T^{-1} \sum_{t=1}^T \tilde{G}'_{2t} \tilde{G}_{3t} \\ T^{-1/2} \sum_{t=1}^T \tilde{G}'_{3t} \tilde{G}_{2t} \Upsilon_T^{-1} & T^{-1} \sum_{t=1}^T \tilde{G}'_{3t} \tilde{G}_{3t} \end{bmatrix}^{-1} \begin{bmatrix} \Upsilon_T^{-1} \sum_{t=1}^T \tilde{G}'_{2t} Q' \varepsilon_t \\ T^{-1/2} \sum_{t=1}^T \tilde{G}'_{3t} Q' \varepsilon_t \end{bmatrix} + o_p(1) \\ & \stackrel{def}{=} \tilde{M}_T^{-1} \tilde{m}_T + o_p(1). \end{aligned} \quad (A.3)$$

We partition $\tilde{M}_T = [\tilde{M}_{ijT}]_{i,j=1,2}$, $\tilde{M}_T^{-1} = [\tilde{M}_T^{ij}]_{i,j=1,2}$ and $\tilde{m}_T = [\tilde{m}'_{1T}, \tilde{m}'_{2T}]'$ conformably with the other partitions in (A.3). Using the inversion formula for partitioned matrices one then obtains

$$T^{1/2}(\hat{\gamma}_3 - \gamma_3) = \tilde{M}_T^{22}(\tilde{m}_{2T} - \tilde{M}_{21T}\tilde{M}_{11T}^{-1}\tilde{m}_{1T}) \quad (\text{A.4})$$

and

$$\tilde{M}_T^{22} = (\tilde{M}_{22T} - \tilde{M}_{21T}\tilde{M}_{11T}^{-1}\tilde{M}_{12T})^{-1}.$$

From the definitions of \tilde{G}_{2t} and \tilde{G}_{3t} and the consistency of the involved estimators it is seen that, as far as subsequent asymptotic derivations are concerned, we can use the approximations

$$\tilde{G}_{2t} \approx -\tilde{Q}'[\tilde{\alpha} : d_t\tilde{\alpha} : t\tilde{\alpha} : b_t\tilde{\alpha}] \approx -k'_t \otimes Q'\alpha$$

and

$$\tilde{G}_{3t} \approx -\tilde{Q}'[\tilde{\Psi}\tilde{\beta}_\perp : d_t\tilde{\Psi}\tilde{\beta}_\perp] \approx -k'_{1t} \otimes Q'\Psi\tilde{\beta}_\perp,$$

where $k_t = [k'_{1t}, k'_{2t}]' = [1, d_t, t, b_t]'$ and $k_{1t} = [1, d_t]'$. It is straightforward to establish the existence and nonsingularity of the limit

$$\lim_{T \rightarrow \infty} \Upsilon_T^{-1} \sum_{t=1}^T k_t k'_t \Upsilon_T^{-1} = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{22} \end{bmatrix} = \Lambda,$$

where

$$\Lambda_{11} = \begin{bmatrix} 1 & 1 - \lambda \\ 1 - \lambda & 1 - \lambda \end{bmatrix}$$

and explicit expressions of $\Lambda_{12} = \Lambda'_{21}$ and Λ_{22} will not be needed.

Using the preceding approximations of \tilde{G}_{2t} and \tilde{G}_{3t} we get

$$\begin{aligned} \tilde{M}_{22T} &= T^{-1} \sum_{t=1}^T k_{1t} k'_{1t} \otimes \tilde{\beta}'_\perp \Psi' \Omega^{-1} \Psi \tilde{\beta}_\perp + o_p(1) \\ &= \Lambda_{11} \otimes \tilde{\beta}'_\perp \Psi' \Omega^{-1} \Psi \tilde{\beta}_\perp + o_p(1), \end{aligned}$$

$$\begin{aligned} \tilde{M}_{21T} = \tilde{M}'_{12T} &= T^{-1/2} \sum_{t=1}^T k_{1t} k'_t \Upsilon_T^{-1} \otimes \tilde{\beta}'_\perp \Psi' \Omega^{-1} \alpha + o_p(1) \\ &= [\Lambda_{11} : \Lambda_{12}] \otimes \tilde{\beta}'_\perp \Psi' \Omega^{-1} \alpha + o_p(1), \end{aligned}$$

and

$$\begin{aligned} \tilde{M}_{11T} &= \Upsilon_T^{-1} \sum_{t=1}^T k_t k'_t \Upsilon_T^{-1} \otimes \alpha' \Omega^{-1} \alpha + o_p(1) \\ &= \Lambda \otimes \alpha' \Omega^{-1} \alpha + o_p(1). \end{aligned}$$

Thus, by straightforward calculation,

$$\begin{aligned}\tilde{M}_{21T}\tilde{M}_{11T}^{-1}\tilde{M}_{12T} &= [\Lambda_{11} : \Lambda_{12}] \Lambda^{-1} \begin{bmatrix} \Lambda_{11} \\ \Lambda_{21} \end{bmatrix} \otimes \bar{\beta}'_{\perp} \Psi' \Omega^{-1} \alpha (\alpha' \Omega^{-1} \alpha)^{-1} \alpha' \Omega^{-1} \Psi \bar{\beta}_{\perp} + o_p(1) \\ &= \Lambda_{11} \otimes \bar{\beta}'_{\perp} \Psi' \Omega^{-1} \alpha (\alpha' \Omega^{-1} \alpha)^{-1} \alpha' \Omega^{-1} \Psi \bar{\beta}_{\perp} + o_p(1)\end{aligned}$$

and, furthermore,

$$\begin{aligned}\tilde{M}_T^{22} &= \left(\Lambda_{11} \otimes \bar{\beta}'_{\perp} \Psi' \Omega^{-1} \Psi \bar{\beta}_{\perp} - \Lambda_{11} \otimes \bar{\beta}'_{\perp} \Psi' \Omega^{-1} \alpha (\alpha' \Omega^{-1} \alpha)^{-1} \alpha' \Omega^{-1} \Psi \bar{\beta}_{\perp} \right)^{-1} + o_p(1) \\ &= \left(\Lambda_{11} \otimes \bar{\beta}'_{\perp} \Psi' [\Omega^{-1} - \Omega^{-1} \alpha (\alpha' \Omega^{-1} \alpha)^{-1} \alpha' \Omega^{-1}] \Psi \bar{\beta}_{\perp} \right)^{-1} + o_p(1) \\ &= \Lambda_{11}^{-1} \otimes \left(\bar{\beta}'_{\perp} \Psi' \alpha_{\perp} (\alpha'_{\perp} \Omega^{-1} \alpha_{\perp})^{-1} \alpha'_{\perp} \Psi \bar{\beta}_{\perp} \right)^{-1} + o_p(1) \\ &= \Lambda_{11}^{-1} \otimes (\alpha'_{\perp} \Psi \bar{\beta}_{\perp})^{-1} \alpha'_{\perp} \Omega^{-1} \alpha_{\perp} \left(\bar{\beta}'_{\perp} \Psi' \alpha_{\perp} \right)^{-1} + o_p(1).\end{aligned}\tag{A.5}$$

Here the third equality makes use of Lemma 10.1, Eq. (10.6), of Johansen (1995).

We also have

$$\tilde{m}_{1T} = -\Upsilon_T \sum_{t=1}^T k_t \otimes \alpha' \Omega^{-1} \varepsilon_t + o_p(1)$$

and

$$\tilde{m}_{2T} = -T^{-1/2} \sum_{t=1}^T k_{1t} \otimes \bar{\beta}'_{\perp} \Psi' \Omega^{-1} \varepsilon_t + o_p(1).$$

Now,

$$\begin{aligned}\tilde{m}_{2T} - \tilde{M}_{21T}\tilde{M}_{11T}^{-1}\tilde{m}_{1T} &= -T^{-1/2} \sum_{t=1}^T k_{1t} \otimes \bar{\beta}'_{\perp} \Psi' \Omega^{-1} \varepsilon_t \\ &\quad + \left([\Lambda_{11} : \Lambda_{12}] \otimes \bar{\beta}'_{\perp} \Psi' \Omega^{-1} \alpha \right) (\Lambda^{-1} \otimes (\alpha' \Omega^{-1} \alpha)^{-1}) \\ &\quad \times \Upsilon_T \sum_{t=1}^T k_t \otimes \alpha' \Omega^{-1} \varepsilon_t + o_p(1),\end{aligned}$$

where

$$\left([\Lambda_{11} : \Lambda_{12}] \otimes \bar{\beta}'_{\perp} \Psi' \Omega^{-1} \alpha \right) (\Lambda^{-1} \otimes (\alpha' \Omega^{-1} \alpha)^{-1}) = [I_2 : 0] \otimes \bar{\beta}'_{\perp} \Psi' \Omega^{-1} \alpha (\alpha' \Omega^{-1} \alpha)^{-1}.$$

Thus, it follows that

$$\begin{aligned}\tilde{m}_{2T} - \tilde{M}_{21T}\tilde{M}_{11T}^{-1}\tilde{m}_{1T} &= -T^{-1/2} \sum_{t=1}^T k_{1t} \otimes \bar{\beta}'_{\perp} \Psi' \Omega^{-1} \varepsilon_t \\ &\quad + T^{-1/2} \sum_{t=1}^T k_{1t} \otimes \bar{\beta}'_{\perp} \Psi' \Omega^{-1} \alpha (\alpha' \Omega^{-1} \alpha)^{-1} \alpha' \Omega^{-1} \varepsilon_t + o_p(1) \\ &= -T^{-1/2} \sum_{t=1}^T k_{1t} \otimes \bar{\beta}'_{\perp} \Psi' \left(\Omega^{-1} - \Omega^{-1} \alpha (\alpha' \Omega^{-1} \alpha)^{-1} \alpha' \Omega^{-1} \right) \varepsilon_t + o_p(1) \\ &= -T^{-1/2} \sum_{t=1}^T k_{1t} \otimes \bar{\beta}'_{\perp} \Psi' \alpha_{\perp} (\alpha'_{\perp} \Omega^{-1} \alpha_{\perp})^{-1} \alpha'_{\perp} \varepsilon_t + o_p(1),\end{aligned}\tag{A.6}$$

where we have again used Lemma 10.1, Eq. (10.6), of Johansen (1995).

Combining (A.4), (A.5) and (A.6) shows that

$$\begin{aligned} T^{1/2}(\hat{\underline{\gamma}}_3 - \underline{\gamma}_3) &= -\left(\Lambda_{11}^{-1} \otimes (\alpha'_\perp \Psi \bar{\beta}_\perp)^{-1} \alpha'_\perp \Omega^{-1} \alpha_\perp (\bar{\beta}'_\perp \Psi' \alpha_\perp)^{-1}\right) \\ &\quad \times T^{-1/2} \sum_{t=1}^T k_{1t} \otimes \bar{\beta}'_\perp \Psi' \alpha_\perp (\alpha'_\perp \Omega^{-1} \alpha_\perp)^{-1} \alpha'_\perp \varepsilon_t + o_p(1) \\ &= -T^{-1/2} \sum_{t=1}^T \Lambda_{11}^{-1} k_{1t} \otimes \beta'_\perp C \varepsilon_t + o_p(1) \end{aligned}$$

or that

$$[T^{1/2}(\hat{\underline{\gamma}}_{31} - \underline{\gamma}_{31}) : T^{1/2}(\hat{\underline{\gamma}}_{32} - \underline{\gamma}_{32})] = -T^{-1/2} \sum_{t=1}^T \beta'_\perp C \varepsilon_t k'_{1t} \Lambda_{11}^{-1} + o_p(1).$$

The definition of k_{1t} and a standard application of the functional central limit theorem imply that

$$T^{-1/2} \sum_{t=1}^T \varepsilon_t k'_{1t} \xrightarrow{d} [B(1) : (B(1) - B(\lambda))].$$

As in S&L the proof can now be completed by using this weak convergence and the preceding equality.

Proof of Theorem 3.1

First note that

$$\hat{x}_t = x_t - (\hat{\mu}_0 - \mu_0) - (\hat{\mu}_1 - \mu_1)t - (\hat{\delta}_0 - \delta_0)d_t - (\hat{\delta}_1 - \delta_1)b_t \quad (\text{A.7})$$

which in conjunction with Lemma 3.1 and a standard functional central limit theorem gives

$$\begin{aligned} T^{-1/2} \beta'_\perp \hat{x}_{[Ts]} &= T^{-1/2} \beta'_\perp x_{[Ts]} - T^{1/2} \beta'_\perp (\hat{\mu}_1 - \mu_1) ([Ts] / T) \\ &\quad - T^{1/2} \beta'_\perp (\hat{\delta}_1 - \delta_1) (b_{[Ts]} / T) + o_p(1) \\ &\xrightarrow{d} \beta'_\perp C \left(B(s) - [\zeta_1 : \zeta_2] \left[(s - \lambda) \mathbf{1}(s > \lambda) \right] \right) \\ &\stackrel{def}{=} \beta'_\perp C B_*(s). \end{aligned} \quad (\text{A.8})$$

The error term e_t in (3.4) has the structure

$$e_t = \varepsilon_t - \alpha \beta' (\hat{x}_{t-1} - x_{t-1}) + \Delta \hat{x}_t - \Delta x_t - \sum_{j=1}^{p-1} \Gamma_j (\Delta \hat{x}_{t-j} - \Delta x_{t-j}). \quad (\text{A.9})$$

As in previous similar proofs in Johansen (1995, Theorem 11.1) and S&L (Theorem 3.1) it can be shown that the limiting distribution of $LR(r_0)$ depends on the weak limits of

$$T^{-2} \sum_{t=1}^T \beta'_\perp \hat{x}_{t-1} \hat{x}'_{t-1} \beta_\perp \quad \text{and} \quad T^{-1} \sum_{t=1}^T \beta'_\perp \hat{x}_{t-1} e'_t \alpha_\perp.$$

From (A.8) and a standard application of the continuous mapping theorem we first find that

$$T^{-2} \sum_{t=1}^T \beta'_{\perp} \hat{x}_{t-1} \hat{x}'_{t-1} \beta_{\perp} \xrightarrow{d} \beta'_{\perp} C \int_0^1 B_*(s) B_*(s)' ds C' \beta_{\perp}. \quad (\text{A.10})$$

Next, using (A.7), (A.9), Lemma 3.1, and arguments similar to those in (A.21) of S&L it can be shown that¹

$$\begin{aligned} T^{-1} \sum_{t=1}^T \beta'_{\perp} \hat{x}_{t-1} e'_t \alpha_{\perp} &= T^{-1} \sum_{t=1}^T [\beta'_{\perp} x_{t-1} - \beta'_{\perp} (\hat{\mu}_1 - \mu_1)(t-1) - \beta'_{\perp} (\hat{\delta}_1 - \delta_1) b_{t-1}] \\ &\quad \times [\varepsilon'_t C' \beta_{\perp} - (\hat{\mu}_1 - \mu_1)' \beta_{\perp} - (\hat{\delta}_1 - \delta_1)' \beta_{\perp} \Delta b_t] \bar{\beta}'_{\perp} \Psi' \alpha_{\perp} + o_p(1), \end{aligned}$$

where use has also been made of the fact that $\Delta b_t = d_t$ implying that the term $\Delta b_t - \sum_{j=1}^{p-1} \Gamma_j \Delta b_t$ can be approximated by $\Psi \Delta b_t$. For later purposes we also notice the identity $C' \beta_{\perp} \bar{\beta}'_{\perp} \Psi' \alpha_{\perp} = \alpha_{\perp}$ obtained from the definitions. To simplify notation, set $\varrho_T = [\varrho_{1T} : \varrho_{2T}] = T^{1/2} [\beta'_{\perp} (\hat{\mu}_1 - \mu_1) : \beta'_{\perp} (\hat{\delta}_1 - \delta_1)]$ and $k_{2t} = [t, b_t]'$ as before. Then the preceding equation can be written as

$$\begin{aligned} &T^{-1} \sum_{t=1}^T \beta'_{\perp} \hat{x}_{t-1} e'_t \alpha_{\perp} \\ &= T^{-1} \sum_{t=1}^T (\beta'_{\perp} x_{t-1} - T^{-1/2} \varrho_T k_{2,t-1}) (\varepsilon'_t C' \beta_{\perp} - T^{-1/2} \Delta k'_{2,t} \varrho'_T) \bar{\beta}'_{\perp} \Psi' \alpha_{\perp} + o_p(1) \\ &= T^{-1} \sum_{t=1}^T \beta'_{\perp} x_{t-1} e'_t \alpha_{\perp} - \varrho_T T^{-3/2} \sum_{t=1}^T k_{2,t-1} \varepsilon'_t \alpha_{\perp} \\ &\quad - T^{-3/2} \sum_{t=1}^T \beta'_{\perp} x_{t-1} \Delta k'_{2,t} \varrho'_T \bar{\beta}'_{\perp} \Psi' \alpha_{\perp} \\ &\quad + \varrho_T T^{-2} \sum_{t=1}^T k_{2,t-1} \Delta k'_{2,t} \varrho'_T \bar{\beta}'_{\perp} \Psi' \alpha_{\perp} + o_p(1) \\ &\stackrel{def}{=} A_{1T} + A_{2T} + A_{3T} + A_{4T}. \end{aligned}$$

From a well-known weak convergence result to a stochastic integral one obtains

$$A_{1T} \xrightarrow{d} \beta'_{\perp} C \int_0^1 B(s) dB(s)' \alpha_{\perp}$$

whereas a central limit theorem and Lemma 3.1(vii) give

$$A_{2T} \xrightarrow{d} -\beta'_{\perp} C [\zeta_1 : \zeta_2] \int_0^1 \left[(s - \lambda) \mathbf{1}(s > \lambda) \right]^s dB(s)' \alpha_{\perp}$$

and

$$\varrho'_T \bar{\beta}'_{\perp} \Psi' \alpha_{\perp} \xrightarrow{d} \begin{bmatrix} \zeta'_1 \\ \zeta'_2 \end{bmatrix} C' \beta_{\perp} \bar{\beta}'_{\perp} \Psi' \alpha_{\perp} = \begin{bmatrix} \zeta'_1 \\ \zeta'_2 \end{bmatrix} \alpha_{\perp}.$$

¹Note that there is a typo in (A.21) of S&L. α_{\perp} after ε'_t should be deleted.

From the last result, the definition of k_{2t} , and standard weak convergence arguments we find that

$$A_{3T} \xrightarrow{d} -\beta'_\perp C \int_0^1 B(s)[\mathbf{1}, \mathbf{1}(s > \lambda)] ds \begin{bmatrix} \zeta'_1 \\ \zeta'_2 \end{bmatrix} \alpha_\perp$$

and

$$A_{4T} \xrightarrow{d} \beta'_\perp C [\zeta_1 : \zeta_2] \int_0^1 \begin{bmatrix} s & s\mathbf{1}(s > \lambda) \\ (s - \lambda)\mathbf{1}(s > \lambda) & (s - \lambda)\mathbf{1}(s > \lambda) \end{bmatrix} ds \begin{bmatrix} \zeta'_1 \\ \zeta'_2 \end{bmatrix} \alpha_\perp.$$

Combining the above results gives

$$\begin{aligned} A_{1T} + A_{2T} &\xrightarrow{d} \beta'_\perp C \int_0^1 \left(B(s) - [\zeta_1 : \zeta_2] \begin{bmatrix} s \\ (s - \lambda)\mathbf{1}(s > \lambda) \end{bmatrix} \right) dB(s)' \alpha_\perp \\ &= \beta'_\perp C \int_0^1 B_*(s) dB(s)' \alpha_\perp \end{aligned}$$

and

$$\begin{aligned} A_{3T} + A_{4T} &\xrightarrow{d} -\beta'_\perp C \int_0^1 \left(B(s) - [\zeta_1 : \zeta_2] \begin{bmatrix} s \\ (s - \lambda)\mathbf{1}(s > \lambda) \end{bmatrix} \right) [ds\zeta'_1 + \mathbf{1}(s > \lambda)ds\zeta'_2] \alpha_\perp \\ &= -\beta'_\perp C \int_0^1 B_*(s) [ds\zeta'_1 + \mathbf{1}(s > \lambda)ds\zeta'_2] \alpha_\perp, \end{aligned}$$

where the latter result can be checked by straightforward calculation. Thus,

$$\begin{aligned} T^{-1} \sum_{t=1}^T \beta'_\perp \hat{x}_{t-1} e'_t \alpha_\perp &\xrightarrow{d} \beta'_\perp C \int_0^1 B_*(s) dB(s)' \alpha_\perp - \beta'_\perp C \int_0^1 B_*(s) [ds\zeta'_1 + \mathbf{1}(s > \lambda)ds\zeta'_2] \alpha_\perp \\ &= \beta'_\perp C \int_0^1 B_*(s) dB_*(s)' \alpha_\perp, \end{aligned} \tag{A.11}$$

where we have used the notation $dB_*(s) = dB(s) - ds\zeta_1 - \mathbf{1}(s > \lambda)ds\zeta_2$ with a similar interpretation as in Theorem 3.1.

As in the proofs in Johansen (1995, Theorem 11.1) and S&L (Theorem 3.1) we can now use (A.10) and (A.11) to obtain

$$\begin{aligned} LR(r_0) &\xrightarrow{d} \text{tr} \left\{ (\alpha'_\perp \Omega^{-1} \alpha_\perp)^{-1} \alpha'_\perp \left(\beta'_\perp C \int_0^1 B_*(s) dB_*(s)' \alpha_\perp \right)' \right. \\ &\quad \left. \times \left(\beta'_\perp C \int_0^1 B_*(s) B_*(s)' ds C' \beta_\perp \right)^{-1} \left(\beta'_\perp C \int_0^1 B_*(s) dB_*(s)' \alpha_\perp \right) \right\} \\ &= \text{tr} \left\{ (\alpha'_\perp \Omega^{-1} \alpha_\perp)^{-1} \alpha'_\perp \left(\alpha'_\perp \int_0^1 B_*(s) dB_*(s)' \alpha_\perp \right)' \right. \\ &\quad \left. \times \left(\alpha'_\perp \int_0^1 B_*(s) B_*(s)' ds \alpha_\perp \right)^{-1} \left(\alpha'_\perp \int_0^1 B_*(s) dB_*(s)' \alpha_\perp \right) \right\}, \end{aligned}$$

where the latter equality follows from the definition of the matrix C . Define now the $(n - r_0)$ -dimensional standard Brownian motion $W(s) = (\alpha'_\perp \Omega^{-1} \alpha_\perp)^{-1/2} \alpha'_\perp B(s)$ and note that, by the definitions of $B_*(s)$, $[\zeta_1 : \zeta_2]$, and $[\xi_1 : \xi_2]$ we can write $(\alpha'_\perp \Omega^{-1} \alpha_\perp)^{-1/2} \alpha'_\perp B_*(s) = W_*(s)$. Thus, it follows that the last trace above becomes as stated in the theorem. This completes the proof.

References

- Attfield, C. L. F. and Temple, J. R. W. (2003). Measuring trend output: How useful are the great ratios?, *Discussion Paper No 03/555*, University of Bristol.
- Clemente, J., Montañés, A. and Ponz, M. (1999). Interest rate arbitrage in currency baskets: Forecasting weights and measuring risk, *Applied Economics Letters* **6**: 687–691.
- D’Adda, C. and Scorcu, A. E. (2003). On the time stability of the output-capital ratio, *Economic Modelling* **20**: 1175–1189.
- Doornik, J. A. (1998). Approximations to the asymptotic distributions of cointegration tests, *Journal of Economic Surveys* **12**: 573–593.
- Gonzalo, J. and Pitarakis, J.-Y. (1999). Dimensionality effect in cointegration analysis, in R. Engle and H. White (eds), *Cointegration, Causality, and Forecasting. A Festschrift in Honour of Clive W.J. Granger*, Oxford University Press, Oxford, pp. 212–229.
- Hubrich, K., Lütkepohl, H. and Saikkonen, P. (2001). A review of systems cointegration tests, *Econometric Reviews* **20**: 247–318.
- Johansen, S. (1995). *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press, Oxford.
- Johansen, S., Mosconi, R. and Nielsen, B. (2000). Cointegration analysis in the presence of structural breaks in the deterministic trend, *Econometrics Journal* **3**: 216–249.
- King, R. G., Plosser, C. I., Stock, J. H. and Watson, M. W. (1991). Stochastic trends and economic fluctuations, *American Economic Review* **81**: 819–840.
- Lütkepohl, H. and Krätzig, M. (eds) (2004). *Applied Time Series Econometrics*, Cambridge University Press, Cambridge.
- Lütkepohl, H., Saikkonen, P. and Trenkler, C. (2004). Testing for the cointegrating rank of a VAR process with level shift at unknown time, *Econometrica* **72**: 647–662.
- Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis, *Econometrica* **57**: 1361–1401, Erratum, (1993) **61**: 248–249.
- Perron, P. (1990). Testing for a unit root in a time series with a changing mean, *Journal of Business & Economic Statistics* **8**: 153–162.
- Saikkonen, P. and Lütkepohl, H. (1999). Local power of likelihood ratio tests for the cointegrating rank of a VAR process, *Econometric Theory* **15**: 50–78.
- Saikkonen, P. and Lütkepohl, H. (2000a). Testing for the cointegrating rank of a VAR process with structural shifts, *Journal of Business & Economic Statistics* **18**: 451–464.
- Saikkonen, P. and Lütkepohl, H. (2000b). Trend adjustment prior to testing for the cointegrating rank of a vector autoregressive process, *Journal of Time Series Analysis* **21**: 435–456.

- Saikkonen, P., Lütkepohl, H. and Trenkler, C. (2006). Break date estimation for VAR processes with level shift with an application to cointegration testing, *Econometric Theory* **22**: 15–68.
- Toda, H. Y. (1994). Finite sample properties of likelihood ratio tests for cointegrating ranks when linear trends are present, *Review of Economics and Statistics* **76**: 66–79.
- Trenkler, C. (2004). Determining p -values for systems cointegration tests with a prior adjustment for deterministic terms, *CASE Discussion Paper 37*, Humboldt-Universität zu Berlin.