SFB 649 Discussion Paper 2009-005

Panel Cointegration Testing in the Presence of a Time Trend

Bernd Droge*
Deniz Dilan Karaman Örsal*

*Humboldt-Universität zu Berlin, Germany

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

http://sfb649.wiwi.hu-berlin.de
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin

Panel Cointegration Testing in the Presence of a Time Trend *

Deniz Dilan Karaman Örsal‡ Bernd Droge§

Humboldt-Universität zu Berlin

First draft: February, 11 2008
This draft: January 20, 2009

Abstract

The purpose of this paper is to propose a new likelihood-based panel cointegration test in the presence of a linear time trend in the data generating process. This new test is an extension of the likelihood ratio (LR) test of Saikkonen & Lütkepohl (2000) for trend-adjusted data to the panel data framework, and is called the panel SL test. The idea is first to take the average of the individual LR (trace) statistics over the cross-sections and then to standardize the test statistic with the appropriate asymptotic moments. Under the null hypothesis, this standardized statistic has a limiting normal distribution as the number of time periods (T) and the number of cross-sections (N) tend to infinity sequentially. In addition to the approximation based on asymptotic moments, a second approximation approach involving the moments from a vector autoregressive process of order one is also introduced. By means of a Monte Carlo study the finite sample size and size-adjusted power properties of the test are investigated. The test presents reasonable size with the increase in T and N, and has high power in small samples.

Keywords: Panel Cointegration Test, Likelihood Ratio, Time Trend, Monte Carlo Study.

JEL classification: C33, C12, C15.

1 Introduction

Most macroeconomic variables, e.g. prices, gross domestic product, consumption etc., exhibit a trending behavior. To model this behavior in the multivariate time series literature a drift parameter is included in the vector autoregressive (VAR) model. Building on this idea, Saikkonen & Lütkepohl (2000) proposed Lagrange multiplier (LM) and likelihood ratio (LR) cointegration tests for data with a linear time trend which are different from the popular Johansen (1995) test. Saikkonen & Lütkepohl (2000) based their test on the idea of subtracting estimates of the deterministic terms from the original data and applying the cointegration test on the trend-adjusted data. The principle of subtracting estimates of the deterministic terms of the model was first suggested by Stock & Watson (1988). Saikkonen

*This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

‡Institute for Statistics and Econometrics, School of Business and Economics, Humboldt-Universität zu Berlin, Spandauer Str. 1 10099 Berlin, E-mail: karamand@staff.hu-berlin.de

§Institute for Statistics and Econometrics and CASE - Center for Applied Statistics and Economics, School of Business and Economics, Humboldt-Universität zu Berlin, Spandauer Str. 1, 10099, Berlin, E-mail: droge@wiwi.hu-berlin.de

& Lütkepohl (2000) proposed to estimate the deterministic terms under the null hypothesis using a generalized least squares (GLS) method. By construction, under the null hypothesis the limit distribution of the their tests do not depend on the deterministic terms. In a simulation study, they concluded that their tests have better properties than the test of Johansen (1995) allowing for a linear trend. Moreover, the LR type version of their tests outperforms the LM type version.

So far there are only few examples of likelihood-based panel cointegration tests which allow for a deterministic linear trend in the data generating process. Larsson et al. (2001), who extended the Johansen trace test to panel data and Breitung (2005), who based his tests on the procedure of Saikkonen (1999), showed in their studies that their panel cointegration tests can be extended to the case with deterministic terms, but they did not deliver any proof of corresponding asymptotic results. Additionally, Anderson et al. (2006) introduced a systems panel cointegration test, which allows for a linear time trend. This test is built on the method of Box & Tiao (1977) in which the number of stochastic common trends is determined by the number of certain eigenvalues close to one. Note that these eigenvalues are the squared canonical correlation coefficients between a multivariate time series and its linear projection on its own history. However, there is no likelihood-based panel cointegration test that relies on the idea of subtracting the estimated deterministic terms prior to testing for cointegration.

The goal of this paper is to close this gap. We extend the trend-adjusting procedure of Saikkonen & Lütkepohl (2000) to the panel data framework and propose an LR panel cointegration test in the presence of a linear time trend in the data generating process (DGP); recall that the LR type test was superior to the LM type version in the simulation study of Saikkonen & Lütkepohl (2000). With this new likelihood-based panel cointegration test statistic one can test for the number of cointegrating relations in the system. This is advantageous compared to the single-equation tests, which can only be used to determine whether there is a cointegrating relation or not. The proposed panel SL test statistic is a standardized version of the average of the individual LR test statistics (trace statistics) over the cross-sections. The standardization is based on the first two moments of the asymptotic trace statistic; i.e. of the limit distribution of the trace statistic. Alternatively, according to Larsson (1999) and Larsson et al. (2001) moments from an approximating VAR(1) process could be used. Under the null hypothesis, the panel SL test statistic converges in distribution to the standard normal law as the number of time periods and the number of cross-sections tend to infinity in a sequential way. Therefore standard normal quantiles may serve as critical values. To justify our approach, we show that the first two moments of the asymptotic trace statistic exists and may be obtained as limits of the moments of a statistic defined in (18), which is used to approximate the asymptotic moments by simulation. This result is an extension of a result of Karaman Örsal & Droge (2009) who corrected a related proof in Larsson et al. (2001) for the case without deterministic terms. The results of a simulation study suggest that the panel SL test has reasonable finite sample properties.

The paper is organized as follows. In Section 2 the heterogeneous panel vector error correction (VEC) model with linear time trend is introduced. Section 3 describes the estimation of the deterministic terms and Section 4 presents the new LR panel cointegration test. The size and size-adjusted power properties are examined by means of a Monte Carlo study in Section 5. Finally, Section 6 gives a summary of the main results. All proofs are deferred to the Appendix in Section 7.

2 The Model

Consider a panel data set consisting of N cross-sections (individuals) observed over T time periods and suppose that for each individual i $(i = 1, \ldots, N)$ a K-dimensional time series $y_{it} = (y_{1it}, \ldots, y_{Kit})'$, $t = 1, \ldots, T$, is observed which is generated by the following heterogenous VAR(p_i) model with linear trend:

$$\begin{aligned} y_{it} &= \mu_{0i} + \mu_{1i}t + x_{it}, \quad i = 1, \ldots, N; \quad t = 1, \ldots, T, &(1)\\ x_{it} &= A_{i1}x_{i,t-1} + \ldots + A_{i,p_i}x_{i,t-p_i} + \varepsilon_{it}. &(2)\end{aligned}$$

Here μ_{0i} and μ_{1i} are unknown K-dimensional parameter vectors, p_i is the lag order of the VAR process for the ith cross-section and $A_{i1}, ..., A_{i,p_i}$ are unknown $(K \times K)$ coefficient matrices. Moreover, we assume that the K-dimensional random errors ε_{it} are serially and cross-sectionally independent with $\varepsilon_{it} \sim N_K(0, \Omega_i)$, for some nonrandom positive definite matrix Ω_i. For simplicity the initial value condition $x_{it} = 0$, $t \leq 0$ and $i = 1, \ldots, N$, is imposed. However, the results remain valid if we assume that the initial values are drawn from a fixed probability distribution, which does not depend on the sample size.

By subtracting $x_{i,t-1}$ from both sides of (2) and rearranging terms we get the VEC form of the model x_{it}:

$$\Delta x_{it} = \Pi_i x_{i,t-1} + \sum_{j=1}^{p_i-1} \Gamma_{ij}\Delta x_{i,t-j} + \varepsilon_{it}, \quad i = 1, \ldots, N; \quad t = 1, \ldots, T, \tag{3}$$

in which $\Pi_i = -(I_K - A_{i1} - \ldots - A_{i,p_i})$ and $\Gamma_{ij} = -(A_{i,j+1} + \ldots + A_{i,p_i})$ for $j = 1, \ldots, p_i - 1$. The components of the process x_{it} are assumed to be integrated at most of order one and cointegrated with cointegrating rank r_i, $0 \leq r_i \leq K$. In other words, y_{it} is at most $I(1)$ and cointegrated at most of order r_i. Thus, the matrix Π_i can be decomposed as

$$\Pi_i = \alpha_i\beta_i', \quad i = 1, \ldots, N, \tag{4}$$

where both α_i and β_i are $(K \times r_i)$ matrices of full column rank. Note that α_i is the loading and β_i is the cointegrating matrix.

On account of (1), (2) and (3) we obtain the VEC form of y_{it}:

$$\begin{aligned} \Delta y_{it} &= \nu_i + \alpha_i\left[\beta_i' y_{i,t-1} - \tau_i(t-1)\right] + \sum_{j=1}^{p_i-1}\Gamma_{ij}\Delta y_{i,t-j} + \varepsilon_{it}, &(5)\\ &\qquad i = 1, ..., N; \quad t = p_i + 1, p_i + 2, ..., T, \end{aligned}$$

with $\nu_i = -\Pi_i\mu_{0i} + (I_K - \Gamma_{i1} - \ldots - \Gamma_{i,p_i-1})\mu_{1i}$ and $\tau_i = \beta_i'\mu_{1i}$.

To determine the number of cointegrating relations among the components of the process y_{it}, the rank of the matrix Π_i should be tested. The relevant null and alternative hypotheses for the cointegration tests are

$$H_0 : \text{rank}(\Pi_i) = r_i \leq r, \quad i = 1, \ldots, N \quad \text{vs.} \quad H_1 : \text{rank}(\Pi_i) = K, \quad i = 1, \ldots, N. \tag{6}$$

Under the null hypothesis all the cross-sections have at most cointegrating rank r, whereas under the alternative hypothesis the rank of Π_i, $i = 1, \ldots, N$, is K. Before testing for the cointegrating rank the data should be trend-adjusted. For the trend-adjustment, estimations of the deterministic terms μ_{0i} and μ_{1i} are required.

3 Estimation of the Deterministic Terms

To estimate the parameters μ_{0i} and μ_{1i}, the GLS method is applied. The data series is then trend-adjusted by subtracting the estimated deterministic terms from y_{it}.

For estimating the deterministic terms, we use the initial value condition $x_{it} = 0$, for $t \leq 0$. First we rewrite (1) as

$$A_i(L)y_{it} = G_{it}\mu_{0i} + H_{it}\mu_{1i} + \varepsilon_{it}, \quad i = 1, ..., N; \quad t = 1, ..., T, \tag{7}$$

with $A_i(L) = I_K - A_{i1}L - ... - A_{i,p_i}L^{p_i}$, $G_{it} = A_i(L)a_t$, $H_{it} = A_i(L)b_t$ and

$$a_t = \begin{cases} 1 & \text{for } t \geq 1 \\ 0 & \text{for } t \leq 0 \end{cases}, \qquad b_t = \begin{cases} t & \text{for } t \geq 1 \\ 0 & \text{for } t \leq 0 \end{cases}.$$

Then, (7) is premultiplied by Q_i' with

$$Q_i = \left[\Omega_i^{-1}\alpha_i(\alpha_i'\Omega_i^{-1}\alpha_i)^{-1/2} : \alpha_{i_\perp}(\alpha_{i_\perp}'\Omega_i\alpha_{i_\perp})^{-1/2}\right] \quad \text{and} \quad Q_iQ_i' = \Omega_i^{-1}, \tag{8}$$

so that the resulting error terms $Q_i'\varepsilon_{it}$ have an identity covariance matrix[1].

Replacing the unknown parameter matrices α_i, β_i, Γ_{ij} and Ω_i of the transformed model by their reduced rank (RR) estimates ($\widetilde{\alpha}_i$, $\widetilde{\beta}_i$, $\widetilde{\Gamma}_{ij}$ and $\widetilde{\Omega}_i$, respectively) from (5), the model can be written in a feasible form. Note that the unknown parameters are estimated under the null hypothesis that the cointegrating rank is r.

With the estimates of the matrices α_i, β_i, Γ_{ij} and their definitions from the previous section, the unknown $(K \times K)$ coefficient matrices A_{ij}, $i = 1, \ldots, N$ and $j = 1, \ldots, p_i$, can be estimated by

$$\begin{aligned} \widetilde{A}_{i1} &= I_K + \widetilde{\alpha}_i\widetilde{\beta}_i' + \widetilde{\Gamma}_{i1}, \\ \widetilde{A}_{ij} &= \widetilde{\Gamma}_{ij} - \widetilde{\Gamma}_{i,j-1}, \quad \text{for} \quad j = 2, ..., p_i - 1, \\ \widetilde{A}_{i,p_i} &= -\widetilde{\Gamma}_{i,p_i-1}, \end{aligned}$$

which allows to use the following:

$$\begin{aligned} \widetilde{A}_i(L) &= I_K - \widetilde{A}_{i1}L - ... - \widetilde{A}_{i,p_i}L^{p_i}, \\ \widetilde{G}_{it} &= \widetilde{A}_i(L)a_t \quad \text{and} \\ \widetilde{H}_{it} &= \widetilde{A}_i(L)b_t. \end{aligned}$$

This leads to a feasible form of the transformed model. The matrices $\widetilde{\alpha}_{i_\perp}$ and $\widetilde{\beta}_{i_\perp}$ can be obtained from the estimates $\widetilde{\alpha}_i$ and $\widetilde{\beta}_i$, respectively. To estimate Q_i, the estimates $\widetilde{\alpha}_i$, $\widetilde{\alpha}_{i_\perp}$, $\widetilde{\Omega}_i$ are inserted into (8), so that

$$\widetilde{Q}_i = \left[\widetilde{\Omega}_i^{-1}\widetilde{\alpha}_i(\widetilde{\alpha}_i'\widetilde{\Omega}_i^{-1}\widetilde{\alpha}_i)^{-1/2} : \widetilde{\alpha}_{i_\perp}(\widetilde{\alpha}_{i_\perp}'\widetilde{\Omega}_i\widetilde{\alpha}_{i_\perp})^{-1/2}\right] \quad \text{for} \quad i = 1, \ldots, N. \tag{9}$$

Finally, the estimators of μ_{0i} and μ_{1i} can be obtained by the multivariate least squares method applied to the following auxiliary regression equations, separately for each cross-section:

$$\widetilde{Q}_i'\widetilde{A}_i(L)y_{it} = \widetilde{Q}_i'\widetilde{G}_{it}\mu_{0i} + \widetilde{Q}_i'\widetilde{H}_{it}\mu_{1i} + \widetilde{Q}_i'\varepsilon_{it}, \quad i = 1, ..., N; \quad t = 1, ..., T. \tag{10}$$

As pointed out earlier, the least squares estimates of $\widetilde{\mu}_{0i}$ and $\widetilde{\mu}_{1i}$ from (10) are used to trend-adjust the data before testing for cointegration.

[1] If A is an $(n \times m)$ matrix of full column rank, its orthogonal complement is denoted by $A_\perp$, where $A_\perp$ is an $(n \times (n - m))$ matrix of full column rank such that $A'A_\perp = 0$.

4 Panel Cointegration Test

Saikkonen & Lütkepohl (2000) introduced both LM and LR cointegration test statistics. By means of a simulation study they concluded that the LR tests are preferable to LM tests. Based on this result we propose an LR panel cointegration test statistic, which is an extension of the $\mathrm{LR}^{\mathrm{GLS}}_{\mathrm{trace}}$ statistic of Saikkonen & Lütkepohl (2000) to panel data.

The new test statistic is based on the following trend-adjusted panel VEC model:

$$\Delta\widetilde{x}_{it} = \Pi_i \widetilde{x}_{i,t-1} + \sum_{j=1}^{p_i-1} \Gamma_{ij} \Delta\widetilde{x}_{i,t-j} + e_{it}, \quad i = 1, ..., N; \quad t = p_i + 1, ..., T, \tag{11}$$

with $\widetilde{x}_{it} = y_{it} - \widetilde{\mu}_{0i} - \widetilde{\mu}_{1i} t$.

The GLS-based trace statistic (LR statistic) for each cross-section is then given by

$$\mathrm{LR}^{\mathrm{GLS}}_{\mathrm{trace}_{iT}}(r) = -2 \ln Q_T\{H(r)|H(K)\} = -T \sum_{j=r+1}^{K} \ln(1 - \widehat{\lambda}_{ij}). \tag{12}$$

Here $\widehat{\lambda}_{i1} \geq \ldots \geq \widehat{\lambda}_{iK}$ are the ordered generalized eigenvalues for cross-section i which are obtained by the eigenvalue problem defined in Johansen (1995).

Under the null hypothesis it follows, as $T \to \infty$,

$$\begin{aligned} &\mathrm{LR}^{\mathrm{GLS}}_{\mathrm{trace}_{iT}}(r) \xrightarrow{w} Z_d \quad \text{with} \\ &Z_d \equiv \mathrm{tr}\left\{ \left(\int_0^1 W_*(s) \mathrm{d}W_*(s)' \right)' \left(\int_0^1 W_*(s) W_*(s)' \mathrm{d}s \right)^{-1} \left(\int_0^1 W_*(s) \mathrm{d}W_*(s)' \right) \right\}, \end{aligned} \tag{13}$$

where $W_*(s) = W(s) - sW(1)$ is a d-dimensional Brownian bridge ($d = K - r$) and $\mathrm{d}W_*(s) = \mathrm{d}W(s) - \mathrm{d}sW(1)$. The proof of this result can be found in the Appendix of Saikkonen & Lütkepohl (2000).

Next, following Larsson et al. (2001), the average of the N individual trace statistic,

$$\overline{\mathrm{LR}}^{\mathrm{GLS}}_{\mathrm{trace}_{NT}}(r) = \frac{1}{N} \sum_{i=1}^{N} \mathrm{LR}^{\mathrm{GLS}}_{\mathrm{trace}_{iT}}(r), \tag{14}$$

is called the $\mathrm{LR}^{\mathrm{GLS}}_{\mathrm{trace}}$-bar statistic. After subtracting the mean and dividing by the standard deviation of the asymptotic trace statistic Z_d, the standardized $\mathrm{LR}^{\mathrm{GLS}}_{\mathrm{trace}}$-bar test (henceforth panel SL test) statistic is given by

$$\gamma_{\overline{\mathrm{LR}}^{\mathrm{GLS}}_{\mathrm{trace}}} = \frac{\sqrt{N}[\overline{\mathrm{LR}}^{\mathrm{GLS}}_{\mathrm{trace}_{NT}}(r) - \mathbb{E}(Z_d)]}{\sqrt{\mathrm{Var}(Z_d)}}, \tag{15}$$

in which $\mathbb{E}(Z_d)$ and $\mathrm{Var}(Z_d)$ are the mean and variance, respectively, of the individual asymptotic trace statistic in (13).

As usual, the mean and variance of Z_d can be approximated by simulation for different values of $d = K - r$ (see Lütkepohl & Saikkonen, 2000). To accomplish this, one generates, for example, $T = 1000$ independent d-dimensional standard normal variates

$\varepsilon_t \sim N(0, I_d)$. Next,

$$A_T = \frac{1}{T^2} \sum_{t=1}^{T} \left[\sum_{m=1}^{t-1} (\varepsilon_m - \bar{\varepsilon}) \right] \left[\sum_{m=1}^{t-1} (\varepsilon_m - \bar{\varepsilon}) \right]', \tag{16}$$

$$B_T = \frac{1}{T} \sum_{t=1}^{T} \left[\sum_{m=1}^{t-1} (\varepsilon_m - \bar{\varepsilon}) \right] (\varepsilon_t - \bar{\varepsilon})', \tag{17}$$

are computed with $\bar{\varepsilon} = T^{-1} \sum_{t=1}^{T} \varepsilon_t$. Because of $A_T \xrightarrow{\omega} \int_0^1 W_*(s) W_*(s)' \mathrm{d}s$ and $B_T \xrightarrow{\omega} \int_0^1 W_*(s) \mathrm{d}W_*(s)'$, it follows

$$Z_{T,d} := \mathrm{tr}\{B_T' A_T^{-1} B_T\} \xrightarrow{w} Z_d. \tag{18}$$

By replicating the experiment 20000 times, approximations of the first two moments of the asymptotic $\mathrm{LR}_{\mathrm{trace}}^{\mathrm{GLS}}$ statistic are computed as sample moments of $Z_{T,d}$ for different values of d. The resulting approximations of the mean and variance of Z_d are presented in Table 1.

Table 1: Simulated first two moments of Z_d.

$d = K - r$	$\mathbb{E}(Z_d)$	$\mathrm{Var}(Z_d)$	$d = K - r$	$\mathbb{E}(Z_d)$	$\mathrm{Var}(Z_d)$
1	2.69	4.38	7	97.91	143.68
2	8.86	13.37	8	127.55	187.28
3	18.85	28.23	9	161.20	238.00
4	32.78	47.94	10	198.43	300.91
5	50.58	73.74	11	239.70	357.05
6	72.44	105.33	12	284.87	424.86

The proposed test statistic is only justified if the first two moments of the asymptotic trace statistic Z_d exist and may be obtained as limits of the corresponding moments of the statistic $Z_{T,d}$. Therefore we prove in Section 7 the following result.

Theorem 1. *It holds* $\mathbb{E}(Z_d^2) < \infty$ *and* $\lim_{T\to\infty} \mathbb{E}(Z_{T,d}^r) \to \mathbb{E}(Z_d^r)$ *for* $r = 1, 2$.

The following theorem is an immediate consequence of the above result together with the central limit theorem and motivates that quantiles of the standard normal law may serve as critical values for the test procedure.

Theorem 2. *Under the null hypothesis,* $H_0 : \mathrm{rank}(\Pi) = r_i \leq r$ *for all* $i = 1, \ldots, N$, *the panel cointegration statistic* $\gamma_{\overline{\mathrm{LR}}_{\mathrm{trace}}^{\mathrm{GLS}}}$ *is asymptotically* $N(0, 1)$ *distributed as* $T \to \infty$, *followed by* $N \to \infty$.

Under certain conditions[2] the asymptotic distribution of the panel cointegration statistic $\gamma_{\overline{\mathrm{LR}}_{\mathrm{trace}}^{\mathrm{GLS}}}$ can also be established when T and N tend jointly to infinity.

It is obvious from (6) that the panel cointegration test is one-sided, and a test at an asymptotic significance level α rejects H_0 defined in (6) if

$$\gamma_{\overline{\mathrm{LR}}_{\mathrm{trace}}^{\mathrm{GLS}}}(r) > z_{1-\alpha},$$

[2]see Phillips & Moon (1999), for conditions under which the sequential convergence implies joint convergence.

where $z_{1-\alpha}$ is the $(1-\alpha)$ quantile of the standard normal distribution.

The sequential testing procedure of Johansen (1988) may be applied to determine the cointegrating rank of the process. First, $H_0 : \text{rank}(\Pi_i) = r_i \leq 0$ is tested. If this null hypothesis is rejected, then $H_0 : \text{rank}(\Pi_i) = r_i \leq 1$ is tested. This procedure continues until the null hypothesis cannot be rejected or $H_0 : \text{rank}(\Pi_i) = r_i \leq K-1$ is rejected. If $H_0 : \text{rank}(\Pi_i) = r_i \leq K-1$ is rejected, then (1) is stable[3].

Following the theory in Larsson (1999) and Larsson et al. (2001) we suggest a second approximation of the moments for the standardization of the panel SL statistic. Larsson (1999) and Larsson et al. (2001) proposed to use the moments from an approximating VAR(1) process, even if the true DGP is a VAR process of higher order. This is motivated by the fact that the moments of the log-likelihood for a VAR(s) process can be approximated sufficiently well by the moments from the log-likelihood for a VAR(1) process, in which s denotes the maximum lag order of the VAR process. In particular, they showed

Theorem 3. *For all positive integers n,*

$$\mathbb{E}[(-2\ln Q_T^{(s)})^n] = \mathbb{E}[(-2\ln Q_T^{(1)})^n] + O(T^{-1}).$$

Here, $-2\ln Q_T^{(s)}$ is the maximum log-likelihood for a VAR(s) process and $-2\ln Q_T^{(1)}$ is the maximum log-likelihood for a VAR(1) process, which can be formulated as

$$\begin{aligned} -2\ln Q_T^{(1)} &= tr\left\{\left[\sum_{t=1}^{T}\left(\sum_{m=1}^{t-1}(\varepsilon_m-\bar{\varepsilon})\right)\left(\sum_{m=1}^{t-1}(\varepsilon_m-\bar{\varepsilon})\right)'\right]^{-1}\right. \\ &\quad \left[\sum_{t=1}^{T}\left(\sum_{m=1}^{t-1}(\varepsilon_m-\bar{\varepsilon})\right)(\varepsilon_t-\bar{\varepsilon})'\right]\left[T^{-1}\sum_{t=1}^{T}(\varepsilon_t-\bar{\varepsilon})(\varepsilon_t-\bar{\varepsilon})'\right]^{-1} \\ &\quad \left.\left[\sum_{t=1}^{T}(\varepsilon_t-\bar{\varepsilon})\left(\sum_{m=1}^{t-1}(\varepsilon_m-\bar{\varepsilon})\right)'\right]\right\} + O_p(T^{-1}), \end{aligned}$$

with $\varepsilon_t \sim N(0, I_d)$ and $\bar{\varepsilon} = T^{-1}\sum_{t=1}^{T}\varepsilon_t$.

Using 50000 replications for different time spans T and values $d = K - r$ the VAR(1) mean and variance are computed by means of a simulation. The results are tabulated in Table 2.

Table 2: Mean and variance values of the VAR(1) approximation.

$d = K - r$	1		2		3		4	
$T - 1$	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
10	2.11	1.75	6.60	3.50	13.21	4.69	21.65	5.27
25	2.42	2.95	7.77	7.42	16.01	12.63	26.98	17.82
50	2.53	3.54	8.28	9.90	17.34	18.31	29.61	28.41
100	2.61	3.90	8.59	11.44	18.15	22.70	31.27	37.21
200	2.66	4.21	8.76	12.49	18.56	25.27	32.10	42.87
500	2.67	4.21	8.86	13.25	18.85	27.17	32.57	45.76
1000	2.67	4.37	8.86	13.41	18.87	27.73	32.80	46.78

[3]Remark: A VAR(p_i) process is stable if $\det(A_i(z)) \neq 0$ for $|z| \leq 1$ with $A_i(z) = I_K - A_{i1}z - ... - A_{i,p_i}z^{p_i}$ (see Lütkepohl, 2005).

5 Monte Carlo Study

Three different DGPs are considered to investigate the finite sample properties of the panel SL test. Particular interest is in checking how the test reacts to the changes in the crucial parameters of the three DGPs.

5.1 DGP A

Since Saikkonen & Lütkepohl (2000) based their simulation study on the Toda (1994, 1995) process, we consider a modified version of this process for panel data.

For $i = 1, ..., N$ and $t = 1, \ldots, T$, the general form of the bivariate Toda process in the presence of a linear trend in the data is

$$y_{it} = \begin{pmatrix} 0 \\ \delta_i \end{pmatrix} + \begin{pmatrix} \psi_a & 0 \\ 0 & \psi_b \end{pmatrix} y_{i,t-1} + \varepsilon_{it}, \quad \varepsilon_{it} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \theta \\ \theta & 1 \end{bmatrix} \right) i.i.d. \tag{19}$$

Throughout the simulation study the initial values y_{i0} are set to zero. The parameter θ represents the correlation between the innovations to the stationary and nonstationary components of the relevant cross-section. If $\theta \neq 0$, then there is instantaneous correlation between the innovations to the stationary and nonstationary components of the process y_{it}. The Toda process is frequently used in the literature because from its canonical form other processes can be obtained by regular linear transformations of y_{it}, and the tests under consideration are invariant to these transformations.

If $\psi_a = \psi_b = 1$, the true cointegrating rank is zero, and there is no cointegrating relation between the components of the process. Then, (19) becomes

$$y_{it} = \delta_i e_2 + I_2 y_{i,t-1} + \varepsilon_{it}, \quad \varepsilon_{it} \sim N(0, I_2), \tag{20}$$

with $e_2 = \begin{pmatrix} 0 & 1 \end{pmatrix}'$. Thus, the process consists of two nonstationary processes. If $\delta_i \neq 0$, a heterogeneous linear trend parameter is present in the second nonstationary process because in a nonstationary unit root processes a drift parameter generates a linear trend. Moreover, there is no instantaneous correlation between the innovations of the two nonstationary components[4], i.e. $\theta = 0$.

If $|\psi_a| < 1$ and $\psi_b = 1$, the true cointegrating rank of the process is one, and (19) can be written as

$$y_{it} = \begin{pmatrix} 0 \\ \delta_i \end{pmatrix} + \begin{pmatrix} \psi_a & 0 \\ 0 & 1 \end{pmatrix} y_{i,t-1} + \varepsilon_{it}, \quad \varepsilon_{it} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \theta \\ \theta & 1 \end{bmatrix} \right) i.i.d. \tag{21}$$

Hence, the process consists of a stationary and a nonstationary component. Instantaneous correlation is present if $\theta \neq 0$, and in the nonstationary component there is a linear trend for $\delta_i \neq 0$.

If $|\psi_a|, |\psi_b| < 1$, then the true cointegrating rank of the process is two, and the VAR process y_{it} is stable. This can be formulated as

$$y_{it} = \begin{pmatrix} \psi_a & 0 \\ 0 & \psi_b \end{pmatrix} y_{i,t-1} + \varepsilon_{it}, \quad \varepsilon_{it} \sim N(0, I_2), \tag{22}$$

[4]Since θ denotes correlation between the innovations to the stationary and nonstationary components of the process, $\theta = 0$.

in which the process consists of two stationary components and $\theta = 0$. The parameter δ_i is excluded from the model as a drift parameter will not create a linear time trend for stationary processes. Besides this we obtain the same simulation results even when we include a drift parameter.

Throughout the simulation study, we consider the same values for the parameters θ, ψ_a and ψ_b as in Saikkonen & Lütkepohl (2000): $\theta \in \{0, 0.8\}$, $\psi_a, \psi_b \in \{0.5, 0.7, 0.8, 0.9, 0.95, 1\}$. The time and cross-section dimensions are the values, which are also taken by Larsson et al. (2001): $N \in \{1, 5, 10, 25, 50\}$ and $T - p \in \{10, 25, 50, 100, 200, 500, 1000\}$, where p denotes the VAR order of the underlying DGP[5]. The drift parameter is independently generated from a uniform distribution $\delta_i \sim U(0, 2)$. In addition to this, we also consider that the drift parameter is homogeneous, i.e. $\delta_i = 1$ for all i. However, this has no effect on the properties of the test. Indeed, the same results are achieved for both heterogeneous and homogeneous cases (cp. Saikkonen & Lütkepohl, 2000; Trenkler, 2002).

5.2 DGP B

The second DGP is a VAR(2) process, which allows for a better examination of the properties of the test based on the VAR(1) approximation of the moments. In particular we see how the test behaves when the underlying VAR process has a higher order than one.

If the true cointegrating rank is zero, the DGP has the form

$$y_{it} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0.8 & 0 \\ 0 & 0.4 \end{pmatrix} y_{i,t-1} + \begin{pmatrix} 0.2 & 0 \\ 0 & 0.6 \end{pmatrix} y_{i,t-2} + \varepsilon_{it}, \quad \varepsilon_{it} \sim N(0, I_2), \tag{23}$$

with

$$\Pi_i = \Pi = -\left(I_2 - \begin{pmatrix} 0.8 & 0 \\ 0 & 0.4 \end{pmatrix} - \begin{pmatrix} 0.2 & 0 \\ 0 & 0.6 \end{pmatrix} \right) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad i = 1, \ldots, N.$$

If the true cointegrating rank is one, the DGP is

$$y_{it} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} \psi & 0 \\ 0 & 0.4 \end{pmatrix} y_{i,t-1} + \begin{pmatrix} 0.2 & 0 \\ 0 & 0.6 \end{pmatrix} y_{i,t-2} + \varepsilon_{it}, \quad \varepsilon_{it} \sim N(0, I_2), \tag{24}$$

with

$$\Pi_i = \Pi = -\left(I_2 - \begin{pmatrix} \psi & 0 \\ 0 & 0.4 \end{pmatrix} - \begin{pmatrix} 0.2 & 0 \\ 0 & 0.6 \end{pmatrix} \right) = \begin{pmatrix} \psi - 0.8 & 0 \\ 0 & 0 \end{pmatrix}, \quad i = 1, \ldots, N.$$

If $\psi < 0.8$, then the DGP consists of a stationary and a nonstationary component. To generate the same Π_i matrices as in DGP A, the ψ parameter takes the values $\psi \in \{0.5, 0.6, 0.7, 0.75\}$. The drift parameter takes the value 1 for all i because a cross-section varying trend term does not affect the results of the simulation study.

A VAR(2) process with a true cointegrating rank of two can be generated as follows:

$$y_{it} = \begin{pmatrix} \psi & 0 \\ 0 & 0.3 \end{pmatrix} y_{i,t-1} + \begin{pmatrix} 0.2 & 0 \\ 0 & 0.2 \end{pmatrix} y_{i,t-2} + \varepsilon_{it}, \quad \varepsilon_{it} \sim N(0, I_2). \tag{25}$$

If we assume again that $\psi \in \{0.5, 0.6, 0.7, 0.75\}$, the DGP is composed of two stationary processes. The drift parameter is not included in the expression as this will not generate a linear trend.

[5]In our study, we consider additionally $T - p \in \{500, 1000\}$ to find out the properties of the tests when T is large.

5.3 DGP C

The third DGP considered in this simulation study is that of Breitung (2005). DGP C differs from the other two DGPs in so far as both the drift parameter and the parameters of the coefficient matrix are heterogeneous over the cross-sections. This is quite suitable for the heterogeneous structure of the model introduced in (1) and (2). The DGP is based on the following VAR(1) model.

$$y_{it} = \mu_i \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 - a_{1i} & -a_{1i}b \\ -a_{2i} & 1 - a_{2i}b \end{pmatrix} y_{i,t-1} + \varepsilon_{it}, \tag{26}$$

in which $\varepsilon_{it} = u_{it} + \vartheta_i u_{i,t-1}$, $u_{it} \sim N(0, I_2)$ $i.i.d$ and $y_{i0} = 0$, $i = 1, \ldots, N$. If $\vartheta_i \neq 0$, then there is correlation between the components of the process y_{it}. Furthermore, the cross-sectional varying parameters are generated from uniform distributions: $\mu_i \sim U(0, 1)$, $\vartheta_i \sim U(0, 0.5)$, $a_{Ki} \sim U(0.1, 0.5)$ for $K = 1, 2$ and $b = 1$.

5.4 Simulation Results

In this section the simulation results based on the three different data generating processes explained above are presented. Throughout the simulation study the test statistics are computed with two different approximations, i.e approximations based on asymptotic moments and on VAR(1) moments. Similar to the Monte Carlo study of Breitung (2005), we compare our panel SL test with the panel test of Larsson et al. (2001) allowing a linear time trend (henceforth LLL test), which is an extension of the test of Johansen (1995) with deterministic terms. The LLL panel test statistic is computed using the asymptotic moments presented in Breitung (2005). To include the results for the LLL test based on VAR(1) moments, the moments are calculated analogous to the procedure described in Section 4.

Note that the total number of replications is 1000. While generating the random error terms, seeded values are used and the first 50 observations are deleted, so that the starting values are not zero anymore. All the programming is done in GAUSS 6.0.

5.4.1 Simulation Results for DGP A

With the approximation based on asymptotic moments the size[6] of the panel SL test for the true cointegrating rank of zero (see Table 3) varies between 0.053 (for $T = 25$, $N = 25$) and 0.118 (for $T = 10$, $N = 50$). If the test statistic is approximated with VAR(1) moments, the empirical size of the test is around the 5% level for $T = 500, 1000$ and otherwise it is severely oversized. Even worse, the LLL test is severely oversized for short time periods under both approximations, and the distortion increases with the increase in N. Moreover, its size comes close to the nominal 5% significance level for longer time periods; it reaches 0.055 for $T = 1000$, $N = 10$, when the VAR(1) moments are used. Overall, based on the asymptotic moments the panel SL test shows the best size properties if the true cointegrating rank of the process is zero. Note that with an increase in T, the size results with different approximations converge to each other.

[6] In the tables presenting the empirical size results, the columns denoted by "asymp" refer to the results of the tests based on the moments of the asymptotic trace statistic, whereas the columns denoted by "VAR(1)" present the results of the tests based on VAR(1) moments.

Table 3: Empirical size results of the tests for DGP A and true cointegrating rank of zero.

		panel SL		LLL	
T-1	N	asymp	VAR(1)	asymp	VAR(1)
10	1	0.072	0.372	0.188	0.765
	5	0.089	0.720	0.394	0.996
	10	0.088	0.890	0.604	1.000
	25	0.104	0.999	0.906	1.000
	50	0.118	1.000	0.990	1.000
25	1	0.067	0.160	0.083	0.286
	5	0.073	0.286	0.114	0.611
	10	0.058	0.395	0.129	0.774
	25	0.053	0.636	0.266	0.983
	50	0.075	0.836	0.379	1.000
50	1	0.081	0.128	0.079	0.165
	5	0.062	0.154	0.100	0.253
	10	0.067	0.181	0.100	0.376
	25	0.061	0.281	0.142	0.604
	50	0.057	0.426	0.178	0.819
100	1	0.064	0.076	0.064	0.092
	5	0.056	0.096	0.058	0.112
	10	0.060	0.114	0.076	0.168
	25	0.077	0.160	0.119	0.284
	50	0.075	0.220	0.147	0.387
200	1	0.076	0.084	0.070	0.082
	5	0.061	0.071	0.070	0.080
	10	0.056	0.071	0.074	0.088
	25	0.077	0.109	0.110	0.139
	50	0.074	0.105	0.124	0.155
500	1	0.069	0.069	0.064	0.062
	5	0.076	0.077	0.082	0.077
	10	0.074	0.074	0.079	0.072
	25	0.068	0.069	0.088	0.078
	50	0.074	0.076	0.115	0.094
1000	1	0.061	0.061	0.080	0.076
	5	0.070	0.070	0.068	0.058
	10	0.066	0.066	0.066	0.055
	25	0.068	0.068	0.081	0.059
	50	0.069	0.069	0.118	0.073

Table 4: Empirical size results of the tests for DGP A and true cointegrating rank of one with $\theta = 0$.

		$\psi_a = 0.7$				$\psi_a = 0.95$			
		panel SL		LLL		panel SL		LLL	
T-1	N	asymp	VAR(1)	asymp	VAR(1)	asymp	VAR(1)	asymp	VAR(1)
10	1	0.022	0.087	0.025	0.187	0.018	0.083	0.021	0.162
	5	0.008	0.095	0.002	0.324	0.003	0.086	0.004	0.311
	10	0.001	0.122	0.002	0.411	0.001	0.102	0.001	0.406
	25	0.000	0.174	0.001	0.654	0.000	0.110	0.000	0.667
	50	0.000	0.238	0.000	0.858	0.000	0.102	0.000	0.860
25	1	0.039	0.068	0.012	0.049	0.015	0.021	0.007	0.036
	5	0.016	0.054	0.003	0.025	0.003	0.010	0.001	0.011
	10	0.006	0.042	0.001	0.020	0.000	0.002	0.000	0.006
	25	0.004	0.049	0.000	0.024	0.000	0.000	0.000	0.002
	50	0.001	0.054	0.000	0.010	0.000	0.000	0.000	0.000
50	1	0.062	0.080	0.030	0.063	0.015	0.024	0.000	0.019
	5	0.058	0.085	0.016	0.046	0.002	0.005	0.000	0.000
	10	0.038	0.089	0.009	0.046	0.000	0.001	0.000	0.002
	25	0.038	0.106	0.002	0.042	0.000	0.000	0.000	0.000
	50	0.031	0.122	0.002	0.028	0.000	0.000	0.000	0.000
100	1	0.060	0.071	0.053	0.064	0.013	0.018	0.013	0.013
	5	0.069	0.092	0.036	0.063	0.001	0.003	0.000	0.000
	10	0.064	0.088	0.059	0.078	0.000	0.000	0.002	0.003
	25	0.063	0.111	0.058	0.099	0.000	0.000	0.000	0.000
	50	0.079	0.149	0.048	0.119	0.000	0.000	0.000	0.000
200	1	0.063	0.062	0.068	0.077	0.022	0.020	0.014	0.016
	5	0.067	0.069	0.071	0.082	0.012	0.011	0.003	0.003
	10	0.059	0.068	0.069	0.081	0.003	0.003	0.003	0.004
	25	0.060	0.069	0.073	0.088	0.000	0.000	0.001	0.001
	50	0.068	0.084	0.082	0.109	0.000	0.000	0.000	0.000
500	1	0.055	0.057	0.063	0.064	0.041	0.044	0.049	0.049
	5	0.066	0.073	0.076	0.076	0.041	0.046	0.046	0.046
	10	0.077	0.084	0.070	0.070	0.027	0.029	0.037	0.037
	25	0.064	0.072	0.068	0.067	0.012	0.014	0.025	0.025
	50	0.083	0.095	0.074	0.073	0.013	0.016	0.020	0.020
1000	1	0.065	0.065	0.066	0.066	0.050	0.050	0.067	0.067
	5	0.066	0.067	0.073	0.071	0.053	0.053	0.073	0.072
	10	0.064	0.066	0.051	0.050	0.050	0.051	0.058	0.058
	25	0.069	0.072	0.071	0.068	0.034	0.035	0.073	0.071
	50	0.056	0.060	0.075	0.071	0.020	0.024	0.086	0.078

Table 5: Empirical size results of the tests for DGP A and true cointegrating rank of one with $\theta = 0.8$.

		$\psi_a = 0.7$				$\psi_a = 0.95$			
		panel SL		LLL		panel SL		LLL	
T-1	N	asymp	VAR(1)	asymp	VAR(1)	asymp	VAR(1)	asymp	VAR(1)
10	1	0.029	0.087	0.026	0.236	0.019	0.083	0.022	0.162
	5	0.016	0.095	0.016	0.482	0.005	0.086	0.005	0.323
	10	0.012	0.122	0.015	0.688	0.001	0.102	0.001	0.416
	25	0.003	0.174	0.008	0.911	0.000	0.110	0.000	0.679
	50	0.001	0.238	0.003	0.991	0.000	0.102	0.000	0.874
25	1	0.045	0.068	0.043	0.122	0.015	0.021	0.008	0.037
	5	0.019	0.054	0.034	0.190	0.003	0.010	0.000	0.013
	10	0.011	0.042	0.040	0.253	0.000	0.002	0.000	0.008
	25	0.008	0.049	0.045	0.479	0.000	0.000	0.000	0.003
	50	0.003	0.054	0.059	0.710	0.000	0.000	0.000	0.000
50	1	0.059	0.080	0.079	0.118	0.014	0.024	0.011	0.019
	5	0.031	0.085	0.080	0.158	0.004	0.005	0.002	0.003
	10	0.019	0.089	0.086	0.215	0.000	0.001	0.001	0.005
	25	0.007	0.106	0.137	0.380	0.000	0.000	0.000	0.000
	50	0.004	0.122	0.209	0.561	0.000	0.000	0.000	0.000
100	1	0.041	0.071	0.074	0.094	0.016	0.018	0.023	0.034
	5	0.033	0.092	0.093	0.127	0.000	0.003	0.010	0.017
	10	0.015	0.088	0.079	0.125	0.000	0.000	0.010	0.013
	25	0.008	0.111	0.116	0.198	0.000	0.000	0.004	0.011
	50	0.005	0.149	0.176	0.287	0.000	0.000	0.000	0.003
200	1	0.037	0.062	0.079	0.088	0.013	0.020	0.049	0.054
	5	0.039	0.069	0.077	0.094	0.003	0.011	0.056	0.069
	10	0.027	0.068	0.067	0.082	0.000	0.003	0.047	0.058
	25	0.017	0.069	0.091	0.122	0.000	0.000	0.078	0.101
	50	0.008	0.084	0.120	0.161	0.000	0.000	0.099	0.150
500	1	0.071	0.057	0.068	0.069	0.017	0.044	0.080	0.084
	5	0.054	0.073	0.068	0.069	0.003	0.046	0.092	0.093
	10	0.039	0.084	0.083	0.084	0.001	0.029	0.109	0.109
	25	0.032	0.072	0.082	0.081	0.000	0.014	0.137	0.137
	50	0.019	0.095	0.096	0.092	0.000	0.016	0.178	0.175
1000	1	0.060	0.065	0.071	0.071	0.019	0.050	0.080	0.080
	5	0.049	0.067	0.062	0.061	0.002	0.053	0.070	0.069
	10	0.045	0.066	0.065	0.062	0.000	0.051	0.086	0.084
	25	0.044	0.072	0.061	0.059	0.000	0.035	0.081	0.077
	50	0.036	0.060	0.082	0.079	0.000	0.024	0.148	0.137

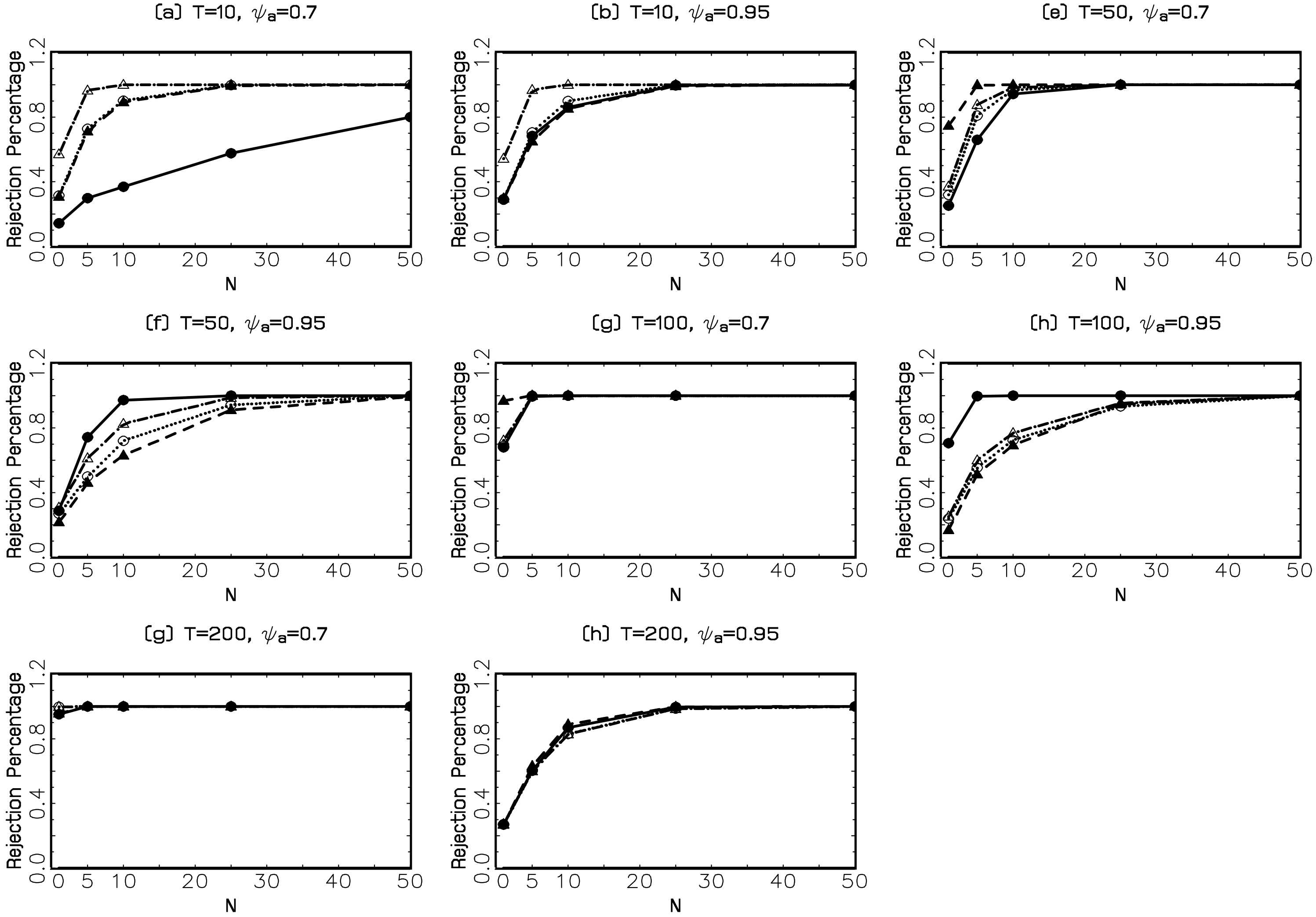

Figure 1: Empirical size-adjusted power results of the tests for DGP A and true cointegrating rank of one with $\theta = 0$ when the hypothesized rank is zero. ●——— panel SL-asymp, ▲ – – – – panel SL-VAR(1), ○........ LLL-asymp, △ – · – · –· LLL-VAR(1).

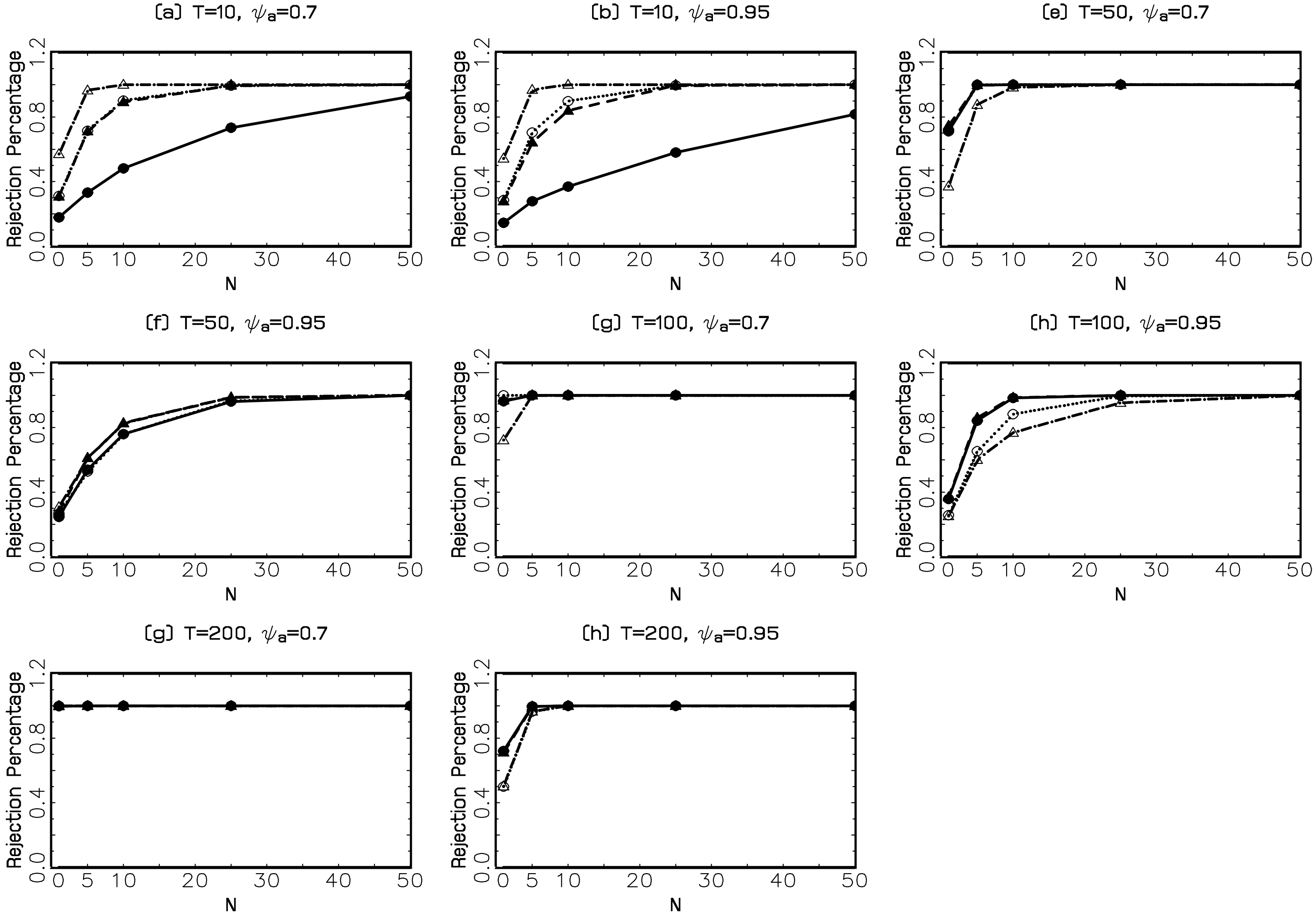

Figure 2: Empirical size-adjusted power results of the tests for DGP A and true cointegrating rank of one with $\theta = 0.8$ when the hypothesized rank is zero. •——— panel SL-asymp, ▲ – – – – panel SL-VAR(1), ○........ LLL-asymp, △ – · – · –· LLL-VAR(1).

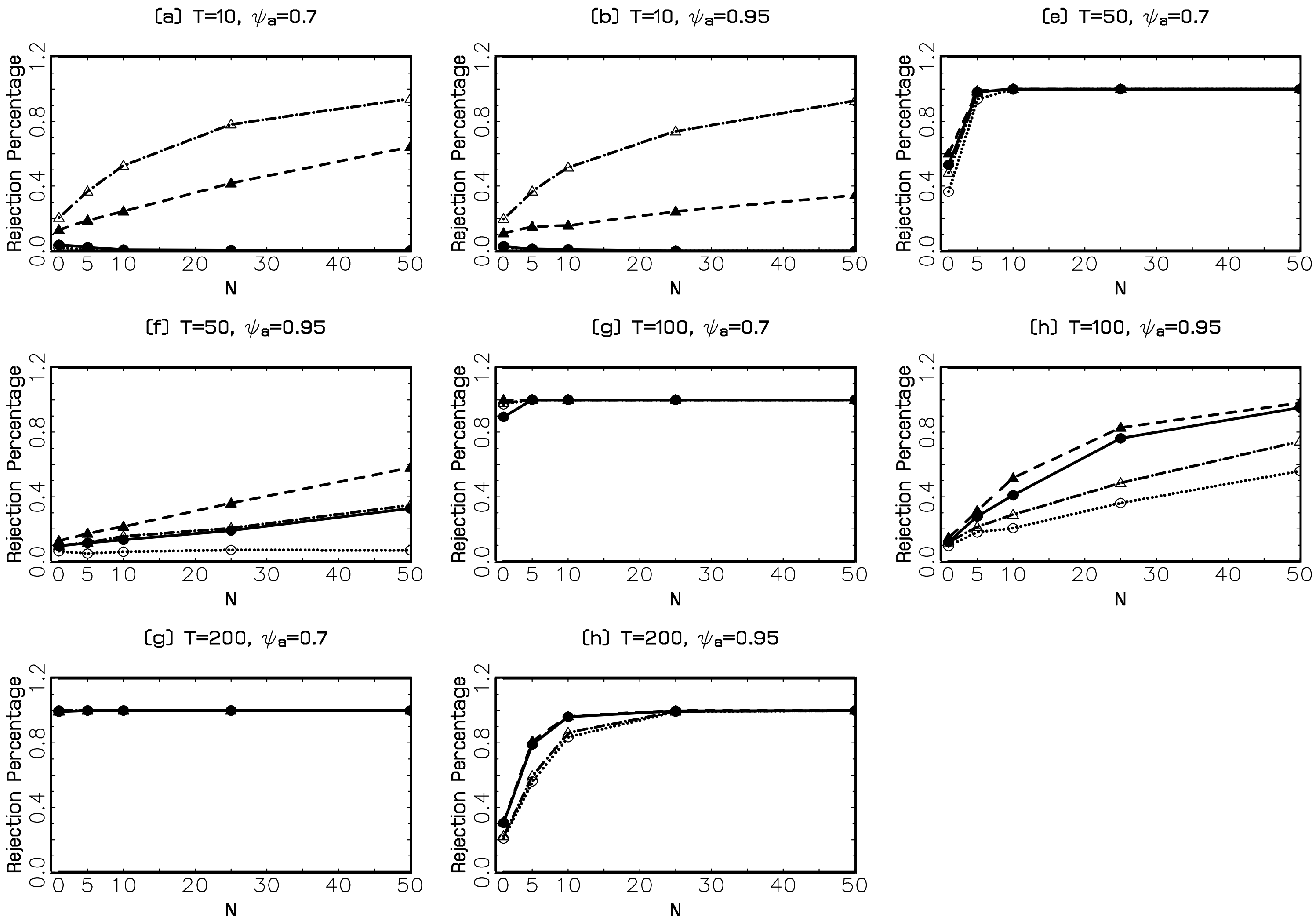

Figure 3: Empirical power results of the tests for DGP A and true cointegrating rank of two when hypothesized rank is one. •——— panel SL-asymp, ▲ – – – – panel SL-VAR(1), ○........ LLL-asymp, △ – · – · –· LLL-VAR(1).

To save space, just the extreme cases, i.e $\psi_a = 0.7$, $\psi_a = 0.95$, are shown for the true cointegrating rank of one. When the asymptotic moments are used to approximate the panel statistics, the true hypothesis of $r = 1$ for $\psi_a = 0.7$ cannot be rejected if $T = 10, 25$ and $N \geq 10$ (see Table 4). With the increase in T the size of the panel SL test rises and is around the 5% level for $T \geq 100$, and it varies between 0.056 (for $T = 1000$, $N = 50$) and 0.083 (for $T = 500$, $N = 50$). If VAR(1) moments are used, the size of the panel SL test comes close to the 5% level for $T = 25$. Moreover, based on the VAR(1) moments the LLL test shows poor size properties for small T. However, if $T = 1000$, the size of the LLL test under both approximations is around the 5% level. For $\psi_a = 0.95$, the panel SL test is undersized for almost all T and N combinations, except for $T = 1000$, $N \leq 10$ (see Table 4). In the latter case the size is exactly 5% with both approximations. The LLL test is also undersized for almost all cases, but the most important difference between the properties of the two tests is that if VAR(1) moments are used, the LLL test is oversized for $T = 10$. With an increase in T, the size of the LLL test moves close to the 5% nominal level. However, the panel SL test has once more better size properties than the LLL test when T increases.

As it is apparent from Table 5 when $\psi_a = 0.7$ and $\theta = 0.8$ the panel SL test has reasonable size either for $N \leq 10$ or $T = 1000$. Hence, for almost all T and N combinations the size of the panel SL test is zero if the true cointegrating rank is one, $\psi_a = 0.95$ and $N \geq 10$. (see Table 5). If the panel SL test statistic is approximated with VAR(1) moments, the test has just the correct size for $T = 25$, $N = 10, 25$ and $T = 1000$, $N = 5$ as $\psi_a = 0.7$. Otherwise the test is size distorted for both ψ_a being either 0.7 or 0.95. However, with the approximation based on asymptotic moments the LLL test is undersized for small T. With an increase in T the size approaches the nominal level, and the test becomes oversized with further increases in T and N. The LLL test approximated with VAR(1) moments is again severely oversized for short time periods, and the size moves around the 5% level, but does not approach it even for large T. In general, none of the tests have nice size properties for $\psi_a = 0.95$.

In line with Banerjee et al. (2004) we observe nonmonotonicities in the results on the size properties of the tests. The sizes of the tests do not increase or decrease monotonically with the increase in T and/or N.

Figures 1-3 present the power results for DGP A[7]. For the true cointegrating rank of one with $\theta = 0$, it is obvious from Figure 1 that the size-adjusted power of the LLL test is slightly better than the size-adjusted power of the panel SL test when $T = 10$. As expected, for small T the approximation based on VAR(1) moments lead to higher power than the approximation based on asymptotic moments. Moreover, the powers of the tests approach unity even for small T if N increases and their powers are almost always unity if T and N are sufficiently large. The same conclusions are also visible in Figure 2, in which the true cointegrating rank is one and $\theta = 0.8$.

From Figure 3 it can be concluded that if both test statistics are approximated with asymptotic moments, the false hypothesis of one cointegrating relation cannot be rejected for $T = 10$. On the contrary, if the test statistics are approximated with VAR(1) moments the powers of the tests increase with an increase in N, and the power of the LLL test is higher. If ψ_a parameter[8] increases, larger T and N are necessary so that the powers of the tests tend

[7]The size-adjusted power results for the true cointegrating rank of zero are not illustrated as the power of the tests for the false hypothesis of one cointegrating relation is around zero.

[8]For DGP A to achieve the true cointegrating rank of two, ψ_b parameter is held constant at 0.5.

Table 6: Empirical size results of the tests for DGP B and true cointegrating rank of zero.

		panel SL		LLL	
T-2	N	asymp	VAR(1)	asymp	VAR(1)
10	1	0.161	0.497	0.446	0.925
	5	0.277	0.898	0.905	1.000
	10	0.430	0.985	0.997	1.000
	25	0.694	1.000	1.000	1.000
	50	0.902	1.000	1.000	1.000
25	1	0.095	0.203	0.129	0.356
	5	0.105	0.386	0.238	0.746
	10	0.092	0.480	0.306	0.928
	25	0.147	0.782	0.596	0.998
	50	0.216	0.964	0.858	1.000
50	1	0.082	0.129	0.110	0.196
	5	0.074	0.170	0.121	0.325
	10	0.094	0.251	0.148	0.462
	25	0.095	0.349	0.240	0.756
	50	0.128	0.564	0.401	0.933
100	1	0.076	0.093	0.075	0.098
	5	0.073	0.110	0.086	0.158
	10	0.066	0.123	0.090	0.187
	25	0.067	0.164	0.141	0.316
	50	0.093	0.241	0.226	0.511
200	1	0.063	0.071	0.077	0.084
	5	0.057	0.070	0.068	0.087
	10	0.066	0.081	0.085	0.102
	25	0.079	0.110	0.132	0.162
	50	0.074	0.123	0.155	0.208
500	1	0.055	0.055	0.063	0.062
	5	0.062	0.063	0.071	0.070
	10	0.057	0.059	0.092	0.083
	25	0.063	0.064	0.090	0.082
	50	0.074	0.075	0.135	0.107
1000	1	0.063	0.063	0.061	0.056
	5	0.064	0.064	0.083	0.070
	10	0.065	0.065	0.073	0.068
	25	0.062	0.062	0.086	0.071
	50	0.078	0.078	0.111	0.077

Table 7: Empirical size results of the tests for DGP B and true cointegrating rank of one.

		$\psi = 0.5$				$\psi = 0.75$			
		panel SL		LLL		panel SL		LLL	
T-2	N	asymp	VAR(1)	asymp	VAR(1)	asymp	VAR(1)	asymp	VAR(1)
10	1	0.037	0.096	0.074	0.326	0.030	0.094	0.067	0.352
	5	0.013	0.106	0.074	0.650	0.014	0.109	0.069	0.698
	10	0.004	0.097	0.059	0.871	0.007	0.093	0.098	0.891
	25	0.001	0.090	0.084	0.992	0.000	0.078	0.117	0.998
	50	0.000	0.074	0.095	1.000	0.000	0.065	0.189	1.000
25	1	0.033	0.064	0.009	0.045	0.027	0.040	0.009	0.047
	5	0.012	0.045	0.000	0.044	0.006	0.020	0.005	0.034
	10	0.007	0.035	0.003	0.027	0.000	0.011	0.001	0.016
	25	0.002	0.037	0.001	0.028	0.000	0.002	0.000	0.015
	50	0.001	0.044	0.000	0.019	0.000	0.001	0.000	0.005
50	1	0.047	0.074	0.017	0.036	0.018	0.026	0.012	0.018
	5	0.034	0.059	0.008	0.031	0.004	0.007	0.000	0.004
	10	0.030	0.064	0.004	0.020	0.000	0.002	0.000	0.000
	25	0.014	0.072	0.000	0.012	0.000	0.000	0.000	0.000
	50	0.028	0.104	0.001	0.010	0.000	0.000	0.000	0.000
100	1	0.078	0.092	0.050	0.072	0.016	0.019	0.005	0.012
	5	0.050	0.073	0.031	0.055	0.005	0.006	0.002	0.005
	10	0.058	0.094	0.023	0.043	0.000	0.000	0.000	0.000
	25	0.059	0.110	0.027	0.058	0.000	0.000	0.000	0.000
	50	0.087	0.137	0.027	0.067	0.000	0.000	0.000	0.000
200	1	0.074	0.077	0.073	0.084	0.031	0.033	0.012	0.014
	5	0.045	0.050	0.056	0.069	0.004	0.004	0.001	0.002
	10	0.075	0.082	0.075	0.090	0.003	0.004	0.000	0.001
	25	0.069	0.081	0.076	0.093	0.000	0.001	0.000	0.000
	50	0.074	0.090	0.074	0.106	0.000	0.000	0.000	0.000
500	1	0.051	0.064	0.060	0.061	0.048	0.043	0.031	0.033
	5	0.075	0.069	0.078	0.081	0.035	0.037	0.034	0.034
	10	0.074	0.088	0.069	0.069	0.030	0.025	0.025	0.025
	25	0.064	0.063	0.077	0.077	0.009	0.013	0.010	0.010
	50	0.068	0.084	0.075	0.073	0.004	0.008	0.013	0.013
1000	1	0.057	0.058	0.058	0.058	0.058	0.059	0.054	0.054
	5	0.065	0.065	0.076	0.075	0.046	0.046	0.073	0.073
	10	0.075	0.076	0.068	0.067	0.052	0.053	0.070	0.068
	25	0.072	0.075	0.064	0.061	0.038	0.040	0.055	0.052
	50	0.057	0.063	0.089	0.085	0.027	0.029	0.087	0.078

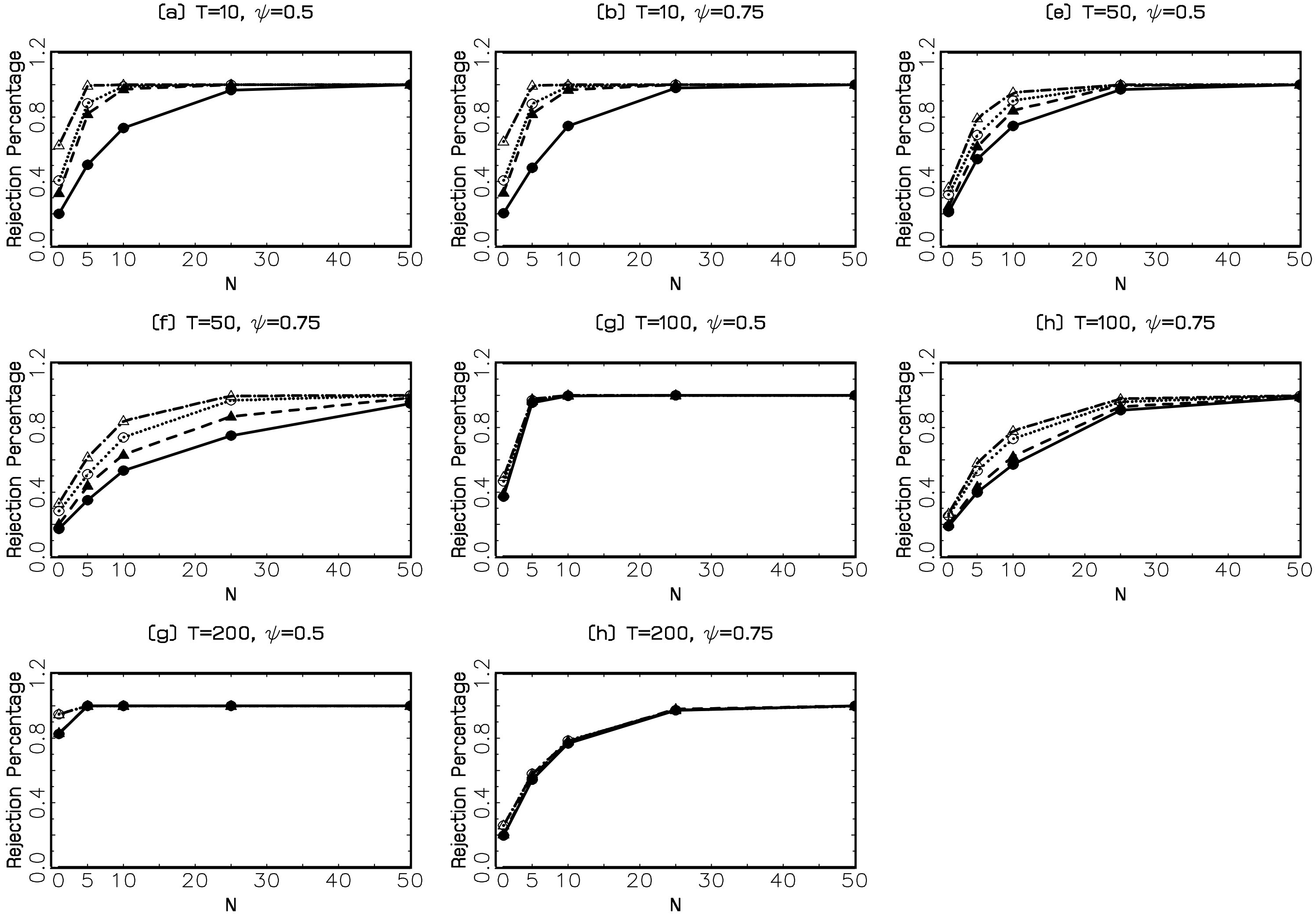

Figure 4: Empirical size-adjusted power results of the tests for DGP B and true cointegrating rank of one when the hypothesized rank is zero. •——— panel SL-asymp, ▲ − − − − panel SL-VAR(1), ○........ LLL-asymp, △ − · − · −· LLL-VAR(1).

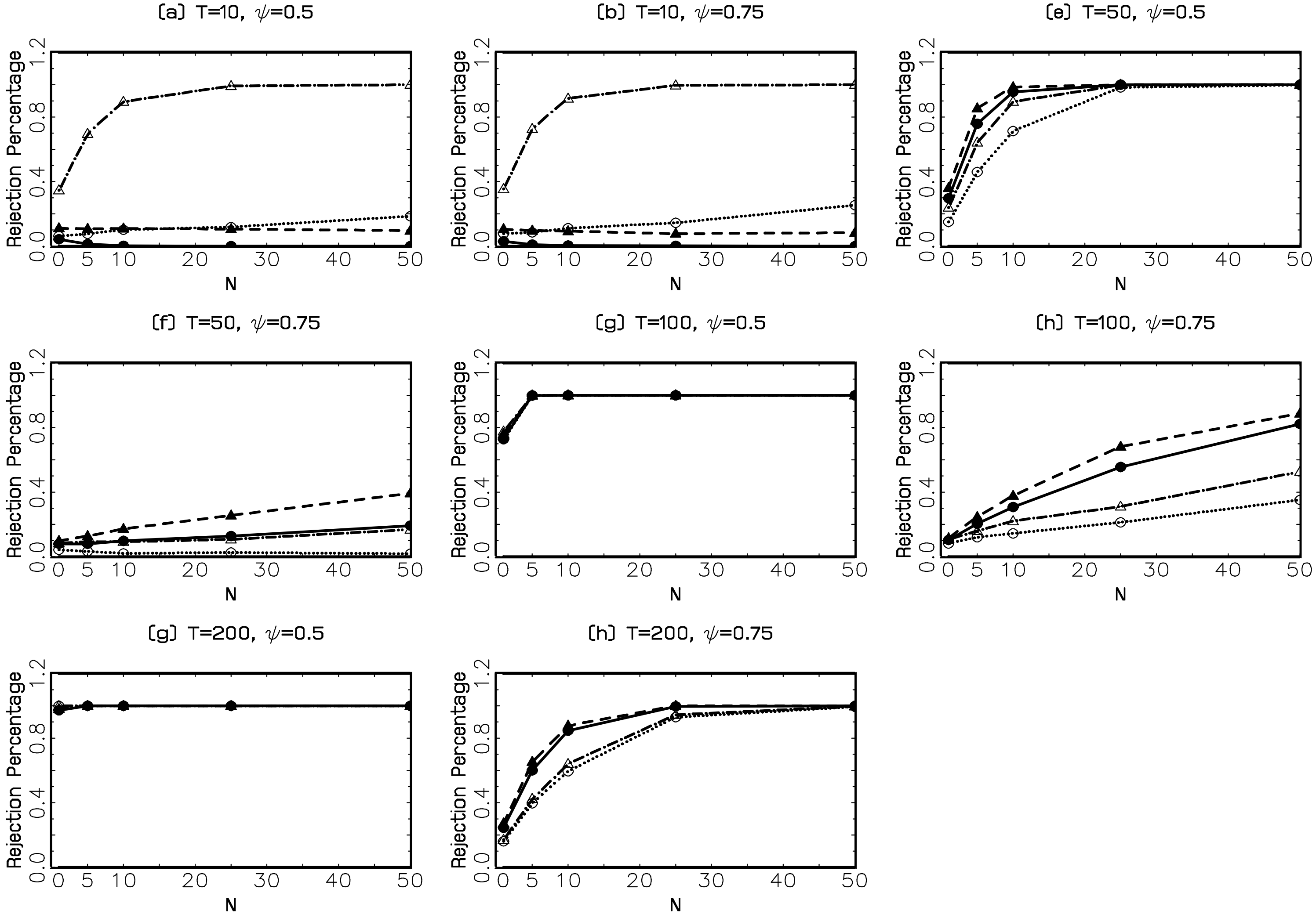

Figure 5: Empirical power results of the tests for DGP B and true cointegrating rank of two when hypothesized rank is one. ●——— panel SL-asymp, ▲ – – – – panel SL-VAR(1), ○........ LLL-asymp, △ – · – · –· LLL-VAR(1).

to unity. Moreover, the LLL test is the least powerful test for $T = 50, 100$ and $\psi_a = 0.95$.

Please note that the size and size-adjusted power results remain the same if a cross-section invariant trend parameter is assumed, i.e. $\delta_i = 1$, $i = 1, ..., N$, instead of a heterogeneous one. This outcome coincides with the simulation results of Saikkonen & Lütkepohl (2000) and Trenkler (2002).

In general, for DGP A, the panel SL test has better size properties in comparison to the LLL test under both approximations. On the contrary, the power of the LLL test is the highest when the test statistic is approximated with VAR(1) moments.

5.4.2 Simulation Results for DGP B

Table 6 demonstrates that the panel SL test is oversized for $T \leq 50$ and its size increases with an increase in N. For $T \geq 100$ the size of the panel SL test ranges from 0.057 (for $T = 500$, $N = 10$) to 0.093 (for $T = 100$, $N = 50$). But if the test statistic is approximated with VAR(1) moments, the test is oversized for $T \leq 200$, and the size is around the 5% nominal significance level only for $T \geq 500$. The LLL test is always more distorted than the panel SL test independent of the chosen approximation. Moreover, if asymptotic moments are used, the size of the panel SL test approaches the 5% level for $T \geq 100$ and $N < 10$. If the true cointegrating rank of zero the panel SL test has the most reasonable size among the two tests and approximations.

To compare the size of the panel SL and LLL tests for the true cointegrating rank of one, just the results related to the two cases $\psi = 0.5$ and $\psi = 0.75$ are presented because the results for $\psi = 0.6$ and $\psi = 0.7$ lie in between these two cases. In Table 7 both tests exhibit similar behavior with the approximation based on asymptotic moments, i.e. they are both undersized for small T and slightly oversized for large T. The size of the LLL test is precisely 0.050 if $T = 100$ and $N = 1$ (no panel data). If the test statistics are approximated with VAR(1) moments, the properties of the tests are different for small T. The panel SL test is undersized for $T = 10$, whereas the LLL test is badly oversized. If $T \geq 100$, the size of the panel SL test ranges from 0.050 (for $T = 200$, $N = 5$) to 0.137 (for $T = 100$, $N = 50$), whereas if $T \geq 25$ the size of the LLL test lies between 0.010 (for $T = 50$, $N = 50$) and 0.106 (for $T = 200, N = 50$).

In general, with VAR(1) moments the panel SL test has better size properties for $T \leq 50$, and with asymptotic moments the test exhibits a reasonable size for $T \geq 100$. With the increase in T once more the size results of the tests based on two different approximations converge to each other. It is apparent from Table 7 that both tests are undersized when ψ increases from 0.5 to 0.75, whereas for $T = 1000$ the sizes of the tests converge to the 5% nominal level. When the test statistics are approximated with VAR(1) moments the tests are undersized if $T \leq 500$, except for $T = 10$, and their sizes approach the 5% level for $T = 1000$.

Figures 4-5 display the power results for DGP B. The size-adjusted powers of both tests for the true cointegrating rank of one approach unity with increasing N even for small T. Moreover, for small T the power of the LLL test is slightly higher than the power of the panel SL test and the approximation with VAR(1) moments delivers higher power than the approximation with asymptotic moments. With the increase in T and N the powers convergence to unity (see Figure 4).

In Figure 5 it is presented that the power of the LLL test is higher than the power of the panel SL test for $T = 10$. The false hypothesis of one cointegrating relation cannot be rejected for the panel SL test when it is based on asymptotic moments. In addition to this,

Table 8: Empirical size results of the tests for DGP C and true cointegrating rank of one.

		$\vartheta_i = 0$				$\vartheta_i \sim U(0, 0.5)$			
		panel SL		LLL		panel SL		LLL	
T-1	N	asymp	VAR(1)	asymp	VAR(1)	asymp	VAR(1)	asymp	VAR(1)
10	1	0.046	0.116	0.027	0.225	0.046	0.123	0.033	0.204
	5	0.010	0.090	0.007	0.412	0.011	0.083	0.005	0.320
	10	0.007	0.083	0.004	0.566	0.005	0.091	0.005	0.472
	25	0.000	0.065	0.002	0.816	0.000	0.062	0.000	0.671
	50	0.000	0.055	0.000	0.968	0.000	0.040	0.000	0.885
25	1	0.047	0.080	0.045	0.099	0.044	0.074	0.027	0.073
	5	0.018	0.064	0.024	0.142	0.020	0.066	0.006	0.035
	10	0.016	0.072	0.015	0.155	0.021	0.081	0.002	0.042
	25	0.012	0.082	0.011	0.236	0.011	0.072	0.000	0.025
	50	0.004	0.095	0.007	0.347	0.002	0.062	0.000	0.015
50	1	0.043	0.061	0.067	0.106	0.077	0.105	0.037	0.061
	5	0.053	0.089	0.053	0.121	0.069	0.104	0.013	0.039
	10	0.045	0.086	0.056	0.139	0.059	0.115	0.011	0.028
	25	0.020	0.070	0.036	0.156	0.059	0.126	0.001	0.019
	50	0.024	0.093	0.041	0.218	0.054	0.173	0.000	0.002
100	1	0.060	0.074	0.060	0.074	0.076	0.089	0.047	0.056
	5	0.065	0.078	0.068	0.098	0.078	0.112	0.021	0.032
	10	0.053	0.087	0.066	0.115	0.108	0.140	0.009	0.019
	25	0.042	0.073	0.055	0.109	0.108	0.174	0.003	0.008
	50	0.039	0.079	0.063	0.149	0.149	0.250	0.001	0.006
200	1	0.073	0.077	0.091	0.099	0.082	0.085	0.050	0.058
	5	0.074	0.082	0.071	0.089	0.107	0.117	0.024	0.027
	10	0.055	0.063	0.078	0.093	0.108	0.120	0.012	0.016
	25	0.065	0.074	0.071	0.099	0.150	0.170	0.006	0.011
	50	0.052	0.064	0.074	0.106	0.197	0.232	0.001	0.001
500	1	0.069	0.069	0.055	0.056	0.090	0.091	0.057	0.057
	5	0.073	0.073	0.064	0.065	0.115	0.116	0.028	0.029
	10	0.058	0.059	0.058	0.059	0.107	0.108	0.019	0.019
	25	0.060	0.063	0.057	0.057	0.151	0.160	0.007	0.007
	50	0.061	0.062	0.081	0.079	0.242	0.255	0.000	0.000
1000	1	0.074	0.075	0.073	0.073	0.089	0.089	0.044	0.044
	5	0.064	0.064	0.055	0.054	0.129	0.129	0.018	0.017
	10	0.067	0.067	0.061	0.059	0.135	0.139	0.011	0.011
	25	0.069	0.072	0.073	0.072	0.183	0.191	0.002	0.002
	50	0.069	0.075	0.080	0.074	0.259	0.269	0.001	0.001

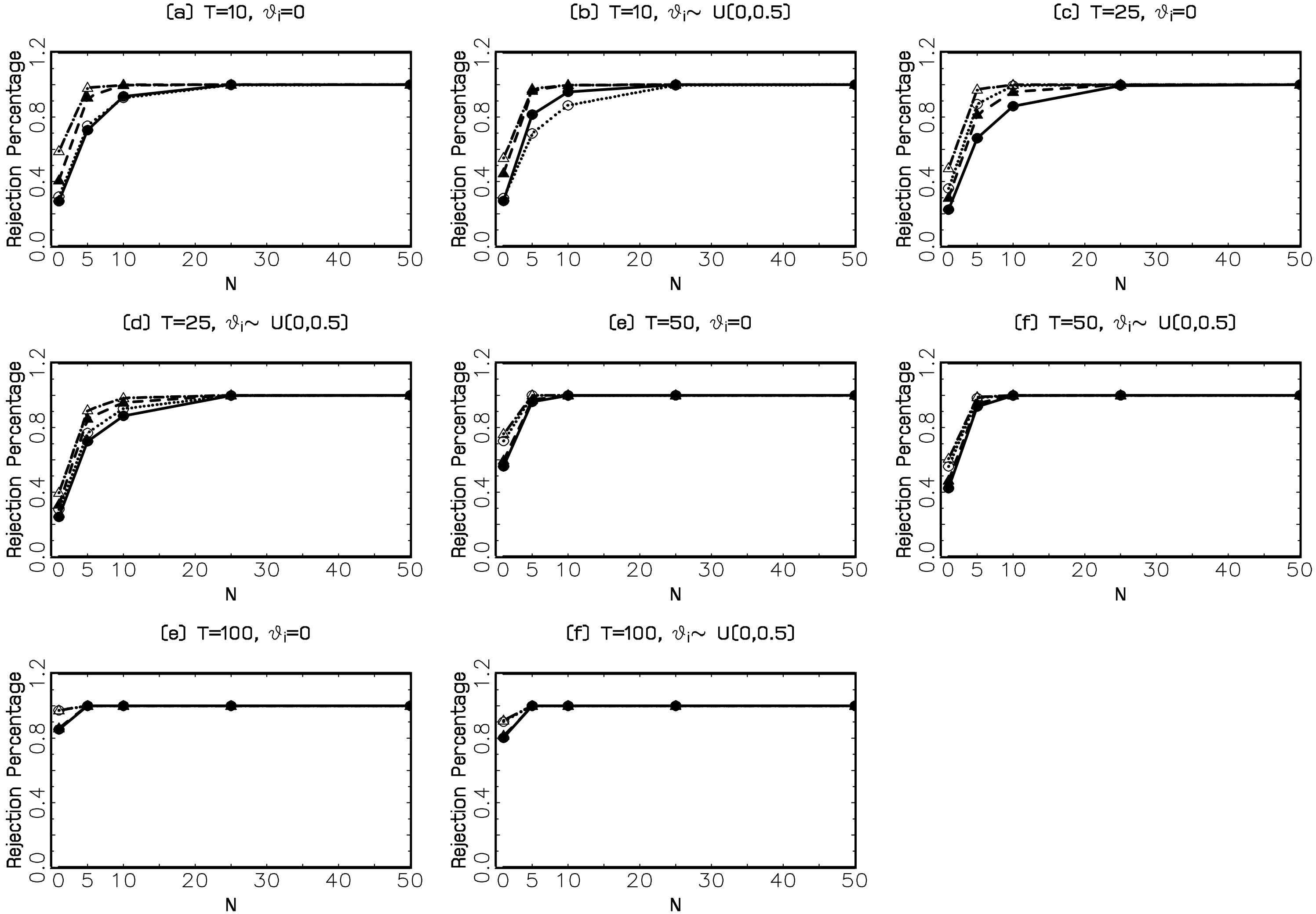

Figure 6: Empirical size-adjusted power results of the tests for DGP C and true cointegrating rank of one when the hypothesized rank is zero. ●——— panel SL-asymp, ▲ – – – – panel SL-VAR(1), ○........ LLL-asymp, △ – · – · –· LLL-VAR(1).

if VAR(1) moments are used the power of the LLL test for $T = 10$ approaches unity, which is not the case for the panel SL test. On the contrary, the panel SL test shows better power than the LLL test with an increase in T to 50. Furthermore, the power of both tests decreases if ψ increases, which is in line with the simulation results of DGP A.

Hence, for DGP B we can conclude that the panel SL test shows better size properties than the LLL test. As outlined above the power of the LLL test based on the approximation with VAR(1) moments is the highest among the considered tests and approximations.

5.4.3 Simulation Results for DGP C

If there is no correlation and the test statistics are approximated with asymptotic moments, both panel tests are undersized for small T and their sizes are around the 5% level for large T (see Table 8). Based on VAR(1) moments, the size of the panel SL test ranges from 0.055 (for $T = 10$, $N = 50$) to 0.095 (for $T = 25$, $N = 50$), especially for $N \leq 5$. On the contrary, if the LLL test statistic is approximated with VAR(1) moments, the test is oversized for $T \leq 200$, and its size is close to the 5% level for $T \geq 500$. Based on the approximation with asymptotic moments, the panel SL test has slightly better size properties than the LLL test for $T = 100, 200$.

In addition to this, if the asymptotic moments are used and there is correlation between the components of the DGP, the panel SL test is undersized for $T = 10$, 25 and it becomes oversized with an increase in T and N, e.g. 0.259 (for $T = 1000$ and $N = 50$). However, the size of the panel SL test is 0.054 for $T = 50$, $N = 50$. If asymptotic moments are used, the LLL test is undersized for almost all T values. Furthermore, based on VAR(1) moments for almost all combinations of T and N, the panel SL test is oversized, whereas the LLL test is just oversized if $T = 10$, and it becomes undersized as T and N rise. Thus, the size of the LLL test does not approach the 5% level, except for $T \geq 50$ and $N = 1$. However, then the LLL test is just the standardized version of the multivariate Johansen trace test which allows a linear time trend in the data.

The size-adjusted power results are similar for both tests, independent of which approximation method is used (see Figure 6). The power of the tests converge to unity with an increase in N, even for small T. This means that the probability of rejecting the false hypothesis of no cointegrating relation is one. If $T \geq 50$, the powers of the tests converge to unity even for $N = 1$. The panel SL test has slightly lower power than the LLL test, but the difference disappears as T rises.

For DGP C the panel SL test has again the best size properties. Both tests are size distorted when there is correlation between the components of the process. Hence, the power of the LLL test is slightly higher than for the panel SL test.

6 Conclusions

In this study a new likelihood-based panel cointegration test (i.e. the panel SL test) was introduced. It allows for a linear time trend in the DGP and is an extension of the multivariate cointegration test ($\mathrm{LR}^{\mathrm{GLS}}_{\mathrm{trace}}$ test) of Saikkonen & Lütkepohl (2000). To find out the finite sample properties of the panel SL test, in a Monte Carlo study three different DGPs were considered and the results were compared with those for the Larsson et al. (2001) test (i.e. the LLL test), which allows a linear time trend in the data.

The simulation results indicate size distortions for small T. The sizes of both tests come close to the nominal 5% significance level as T increases. In general the panel SL test has better size properties than the LLL test, especially if there is no correlation between the components of the DGP. Also for small T, if VAR(1) moments are used the panel SL test delivers better size properties in comparison to the LLL test, which is severely oversized for small T independent of the approximation chosen. Moreover, the sizes of both tests with different approximations converge to each other with an increase in T.

With the introduction of correlation between stationary and nonstationary components of the process, size distortions are observed, however the panel SL test has still reasonable size for large T. In addition to this we found that, if the DGP consists of a nearly nonstationary component, then the tests become size distorted.

In general, the powers of both panel cointegration tests approach unity with an increase in N even when T is small. Additionally, for small T the approximation based on VAR(1) moments delivers tests with higher power than the approximation based on asymptotic moments. When there is a nearly stationary component in the DGP, than larger T and N are necessary so that the test has high power.

7 Appendix

Our proof of Theorem 1 relies on the following Lemma 1, which states that the fourth moments of the statistic $Z_{T,d}$ defined in (18) are uniformly bounded in T. For the sake of simplicity, we present here the proof only for $d = 1$. The extension to the general case $d \geq 1$ is similar as in Karaman Örsal & Droge (2009) for a related statistic and therefore omitted.

Lemma 1. *Let $Z_{T,d}$ be defined as in* (18)*. Then there exist some constants a and b such that, for all T,*

(i.) $\mathbb{E}[Z_{T,d}^2] < a,$

(ii.) $\mathbb{E}[Z_{T,d}^4] < b.$

Proof of Lemma 1. Let $\varepsilon_t \sim N_d(0, \Omega)$ *i.i.d*, $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_T)'$ and $\widetilde{\varepsilon} = (\varepsilon_1 - \overline{\varepsilon}, \ldots, \varepsilon_T - \overline{\varepsilon})'$ with $\overline{\varepsilon} = T^{-1} \sum_{t=1}^{T} \varepsilon_t$. Then the statistic $Z_{T,d}$ may be rewritten as

$$Z_{T,d} \quad = \quad \mathrm{tr}[B_T' A_T^{-1} B_T] = \mathrm{tr}[\widetilde{\varepsilon}' \widetilde{P}_{\widetilde{Y}} \widetilde{\varepsilon}], \tag{27}$$

where

$$A = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 1 & \cdots & \cdots & \cdots & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix},$$

and

$$\widetilde{P}_{\widetilde{Y}} \quad = \quad BA\widetilde{\varepsilon}(\widetilde{\varepsilon}' A' B' BA\widetilde{\varepsilon})^{-1} \widetilde{\varepsilon}' A' B' \tag{28}$$

is the random projection matrix onto the column space of $\widetilde{Y} = BA\widetilde{\varepsilon}$. Moreover, if J_T denotes the $(T \times T)$ matrix of ones, then $P = \frac{1}{T} J_T$ and $Q = (I_T - P)$ are orthogonal projections with

$PQ = 0$ and it follows $\widetilde{\varepsilon} = Q\varepsilon$. Thus,

$$0 \leq Z_{T,d} \;=\; \mathrm{tr}[\widetilde{\varepsilon}'\widetilde{P}_{\widetilde{Y}}\widetilde{\varepsilon}] \leq \mathrm{tr}[\widetilde{\varepsilon}'\widetilde{\varepsilon}] = \mathrm{tr}[\varepsilon' Q\varepsilon] \leq \mathrm{tr}[\varepsilon'\varepsilon]. \tag{29}$$

This shows that all moments of $Z_{T,d}$ exists because $\varepsilon'\varepsilon$ is Wishart distributed [9], more precisely $\varepsilon'\varepsilon \sim W_d(T, \Omega)$, and all moments of a Wishart distributed matrix exist (see Letac & Massam, 1999; Graczyk et al., 2005).

However, to establish the result we have to verify that the second (fourth) moments of $Z_{T,d}$ are uniformly bounded in T. To accomplish this, we write, with $D = BA$,

$$Z_{T,d} \;=\; \varepsilon' QDQ\varepsilon \left[\varepsilon' QD'DQ\varepsilon\right]^{-1} \varepsilon' QD'Q\varepsilon. \tag{30}$$

Consider now the case $d = 1$ and assume, without loss of generality, $\Omega = 1$, i.e. $\varepsilon_t \sim N(0, 1)$ *i.i.d.* Then it holds

$$\varepsilon' QDQ\varepsilon = \varepsilon' QSQ\varepsilon \;\text{ with }\; S = \frac{D + D'}{2} = \frac{1}{2}(J_T - I_T) = \left(\frac{T-1}{2}\right) P - \left(\frac{1}{2}\right) Q. \tag{31}$$

Because of $PQ = 0$, it follows

$$R = QSQ = -\frac{1}{2}Q, \tag{32}$$

so that (30) may be expressed as

$$Z_{T,1} = \frac{U_1^2}{U_2} \;\text{ with }\; U_1 = \varepsilon' R\varepsilon \;\text{ and }\; U_2 = \varepsilon' H\varepsilon, \tag{33}$$

where $H = QD'DQ = QFQ$ and $F = D'D$. Clearly,

$$U_1 = -\frac{1}{2}\zeta_1 \;\text{ with }\; \zeta_1 = \chi^2_{T-1}. \tag{34}$$

Moreover, let $\lambda_1, \lambda_2, \ldots, \lambda_T$ and $\mathrm{v}_1, \mathrm{v}_2, \ldots, \mathrm{v}_T$ be the (ordered) eigenvalues and the associated orthonormal eigenvectors, respectively, of the symmetric and positive semidefinite matrix H. Then $\xi_t := \varepsilon'\mathrm{v}_t \sim N(0, 1)$ *i.i.d.* $(t = 1, \ldots, T)$ and thus

$$U_2 = \sum_{t=1}^{T} \lambda_t \varepsilon' \mathrm{v}_t \mathrm{v}_t' \varepsilon = \sum_{t=1}^{T} \lambda_t \xi_t^2, \tag{35}$$

where $\xi_t^2 \sim \chi_1^2$ *i.i.d* for $t = 1, \ldots, T$.

The eigenvalues of the positive semidefinite matrix $H = QFQ$ are[10]

$$\lambda_t = \frac{1}{2 - 2\cos\left(\frac{t\pi}{T}\right)} \quad \text{for} \quad t = 1, \ldots, T-1, \quad \lambda_T = 0, \tag{36}$$

[9] If $X = Y'Y$, in which the $(n \times m)$ matrix Y is $N(0, I_n \otimes \Sigma)$, then X follows a *Wishart distribution* with n degrees of freedom and covariance matrix Σ; i.e. $X \sim W_m(n, \Sigma)$ where m denotes the size of the matrix X. Note that $I_n \otimes \Sigma$ is the covariance matrix of $\mathbf{y} = \mathrm{vec}(Y')$. Moreover, the $W_m(n, \Sigma)$ distribution has a density function when $n \geq m$ (see Muirhead, 1982).

[10] The positive eigenvalues of the matrix H are the same as the eigenvalues of the inverse of the following tridiagonal matrix:

$$G = (-1)\begin{pmatrix} -2 & 1 & 0 & 0 & \ldots & 0 & 0 \\ 1 & -2 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & -2 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & 1 & -2 \end{pmatrix}.$$

The eigenvalues of such a matrix are discussed in Yueh (2005).

with $\lambda_1 > \lambda_2 > \ldots > \lambda_{T-1} > \lambda_T$. The series expansion of the cosine function provides, for a fixed t,

$$1 - \cos\left(\frac{t\pi}{T}\right) = \frac{(t\pi)^2}{2T^2} + o(T^{-3}) \quad (\text{as} \quad T \to \infty) \tag{37}$$

and therefore

$$\frac{1}{\lambda_t} = \frac{c_1}{T^2} + o(T^{-3}) \quad \text{for some} \quad c_1 \in (0, \infty) \quad (\text{as} \quad T \to \infty). \tag{38}$$

To prove (i.), we first apply the Cauchy-Schwarz inequality, which leads to

$$\mathbb{E}(Z_{T,1}^2) = \mathbb{E}\left(\frac{U_1^4}{U_2^2}\right) \leq \sqrt{\mathbb{E}(U_1^8)\mathbb{E}\left(\frac{1}{U_2^4}\right)} \quad (\text{see (33)}). \tag{39}$$

Because of (34) we obtain

$$\mathbb{E}(U_1^8) = \frac{1}{2^8}\mathbb{E}(\zeta_1^8) = \frac{1}{2^8}2^8\frac{\Gamma\left(8 + \frac{T-1}{2}\right)}{\Gamma\left(\frac{T-1}{2}\right)} = c_2T^8 + o(T^8) \quad \text{for some} \quad c_2 \in (0, \infty). \tag{40}$$

On the other hand, (35) implies for $T > 9$ the lower bound

$$U_2 \geq \lambda_9 z_1 \quad \text{with} \quad z_1 = \sum_{t=1}^{9} \xi_t^2 \sim \chi_9^2 \tag{41}$$

and consequently, on account of (38) and $\mathbb{E}(z_1^{-4}) = \frac{1}{105}$,

$$\mathbb{E}\left(\frac{1}{U_2^4}\right) \leq \frac{1}{\lambda_9^4}\mathbb{E}\left(\frac{1}{z_1^4}\right) = \frac{c_3}{T^8} + o(T^{-8}) \quad \text{for some} \quad c_3 \in (0, \infty) \quad (\text{as} \quad T \to \infty). \tag{42}$$

In view of (40), (42) and recalling that all moments of $Z_{T,d}$ exist, (39) yields

$$\mathbb{E}(Z_{T,1}^2) \leq a \quad \text{for some} \quad a \in (0, \infty) \quad \text{and all} \quad T. \tag{43}$$

The proof of (ii.) is analogous and thus omitted. The main difference is to replace in equality (41) λ_9 by λ_{17} (which is also of order T^2), because the eighth moment of the inverse-χ_{17}^2 distribution is finite.

Proof of Theorem 1. Because of (18) the result follows if $\{Z_{T,d}^2\}$ is uniformly integrable (see Theorem A on p.14 in Serfling, 1980). A sufficient condition for the uniform integrability of $\{Z_{T,d}^2\}$ is that $\mathbb{E}|Z_{T,d}|^{2+\delta}$ is uniformly bounded for some $\delta > 0$, i.e $\sup_T \mathbb{E}|Z_{T,d}|^{2+\delta} < \infty$. But this is an immediate consequence of Lemma 1 (ii.), which completes the proof.

References

Anderson, R., Qian, H., & Rasche, R. (2006). Analysis of panel vector error correction models using maximum likelihood, bootstrap, and canonical-correlation estimators. Working Paper Series 2006-050A, Federal Reserve Bank of St. Louis.

Banerjee, A., Marcellino, M., & Osbat, C. (2004). Some cautions on the use of panel methods for integrated series of macroeconomic data. *Econometrics Journal*, **7**, 322–340.

Box, G. E. P. & Tiao, G. C. (1977). A canonical analysis of multiple time series. *Biometrika*, **64**, 355–365.

Breitung, J. (2005). A parametric approach to the estimation of cointegration vectors in panel data. *Econometric Reviews*, **24**, 1–20.

Graczyk, P., Letac, G., & Massam, H. (2005). The hyperoctahedral group, symmetric group representations and the moments of the real Wishart distribution. *Journal of Theoretical Probability*, **18**, 1–42.

Johansen, S. (1988). Statistical analysis of cointegrating vectors. *Econometric Reviews*, **12**, 231–254.

Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press: Oxford.

Karaman Örsal, D. D. & Droge, B. (2009). On the existence of the moments of the asymptotic trace statistic. Discussion paper, SFB 649, Humboldt Universität zu Berlin.

Larsson, R. (1999). The order of the asymptotic error term for moments of the log likelihood ratio test for cointegration. unpublished manuscript.

Larsson, R., Lyhagen, J., & Löthgren, M. (2001). Likelihood-based cointegration tests in heterogeneous panels. *Econometrics Journal*, **4**, 109–142.

Letac, G. & Massam, H. (1999). All the moments and the inverse moments of the Wishart distribution. *Seminarberichte aus dem Fachbereich Mathematik*, **68**, 309–339.

Lütkepohl, H. (2005). *New Introduction To Multiple Time Series Analysis*. Berlin: Springer.

Lütkepohl, H. & Saikkonen, P. (2000). Testing for the cointegrating rank of a VAR process with a time trend. *Journal of Econometrics*, **95**, 177–198.

Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*. New York: Wiley.

Phillips, P. C. B. & Moon, H. R. (1999). Linear regression limit theory for nonstationary panel data. *Econometrica*, **67**, 1057–1111.

Saikkonen, P. (1999). Testing the normalization and overidentification of cointegrating vectors in vector autoregressive processes. *Econometric Reviews*, **18**, 235–257.

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

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. New York: Wiley.

Stock, J. H. & Watson, M. W. (1988). Testing for common trends. *Journal of American Statistical Association*, **83**, 1097–1107.

Toda, H. Y. (1994). Finite sample properties of likelihood ratio tests for cointegrating ranks when linear trends are present. *Review of Economics & Statistics*, **66**, 66–79.

Toda, H. Y. (1995). Finite sample performance of likelihood ratio tests for cointegrating ranks in vector autoregressions. *Econometric Theory*, **11**, 1015–1032.

Trenkler, C. (2002). *Testing for the Cointegrating Rank in the Presence of Level Shifts*. Aachen: Shaker Verlag.

Yueh, W.-C. (2005). Eigenvalues of several tridiagonal matrices. *Applied Mathematics E-Notes*, **5**, 66–74.

SFB 649 Discussion Paper Series 2009

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "Implied Market Price of Weather Risk" by Wolfgang Härdle and Brenda López Cabrera, January 2009.
002 "On the Systemic Nature of Weather Risk" by Guenther Filler, Martin Odening, Ostap Okhrin and Wei Xu, January 2009.
003 "Localized Realized Volatility Modelling" by Ying Chen, Wolfgang Karl Härdle and Uta Pigorsch, January 2009.
004 "New recipes for estimating default intensities" by Alexander Baranovski, Carsten von Lieres and André Wilch, January 2009.
005 "Panel Cointegration Testing in the Presence of a Time Trend" by Bernd Droge and Deniz Dilan Karaman Örsal, January 2009.
006 "Regulatory Risk under Optimal Incentive Regulation" by Roland Strausz, January 2009.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".