Nonparametric Rank Tests for Non-Stationary Panels

Peter L. Pedroni
Williams College
Williamstown

Timothy J. Vogelsang
Michigan State University
East Lansing

Martin Wagner
Faculty of Statistics
Technical University Dortmund
and
Institute for Advanced Studies
Vienna

Joakim Westerlund*
Deakin University
Melbourne

July 19, 2014

Abstract

We develop a set of nonparametric rank tests for non-stationary panels based on multivariate variance ratios which use untruncated kernels. As such, the tests do not require the choice of tuning parameters associated with bandwidth or lag length and also do not require choices with respect to numbers of common factors. The tests allow for unrestricted cross-sectional dependence and dynamic heterogeneity among the units of the panel, provided simply that a joint functional central limit theorem holds for the panel of differenced series. We provide a discussion of the relationships between our setting and the settings for which first- and second generation panel unit root tests are designed. In Monte Carlo simulations we illustrate the small-sample performance of our tests when they are used as panel unit root tests under the more restrictive DGPs for which panel unit root tests are typically designed, and for more general DGPs we also compare the small-sample performance of our nonparametric tests to parametric rank tests. Finally, we provide an empirical illustration by testing for income convergence among countries.

JEL Classification: C12; C22; C23.

Keywords: Cointegration; Cross-sectional dependence; Nonparametric rank tests; Time series panel; Unit roots

*Corresponding author: Deakin University, Faculty of Business and Law, School of Accounting, Economics and Finance, Melbourne Burwood Campus, 221 Burwood Highway, VIC 3125, Australia. Telephone: +61 3 924 46973. Fax: +61 3 924 46283. E-mail address: j.westerlund@deakin.edu.au.
1 Introduction

This paper develops rank tests for the number of common stochastic trends present in a time series panel. The tests are designed to perform well in situations where the cross-sectional dimension of the panel is too large for traditional multivariate cointegration methods to be used successfully. The paper also investigates the relationships between these rank tests and other key tests in the non-stationary panel literature, such as panel unit root tests and tests for the fraction of individual series which are $I(1)$ versus $I(0)$ processes.

Much of the recent non-stationary panel literature has focused on permitting increasingly general forms of cross-sectional dependencies among members of the panel (see, for example, Breitung and Pesaran, 2008, and Banerjee and Wagner, 2009, for recent overviews). However, as we discuss in Section 2 of this paper, the extent of the cross-sectional dependency that one permits under the data generating process (DGP) is inherently tied to the types of hypotheses that one can successfully test with asymptotic size control. In particular, as we will see, the absence or presence of (cross unit) cointegration among the series is often a key feature in this regard.\(^1\) This is particularly important in relation to the ability to determine the overall number of individual series that are $I(1)$ versus $I(0)$, as well as the ability to determine which particular series are $I(1)$ versus $I(0)$. Granted, in situations where the time series dimension is large enough, one might consider using time series methods alone rather than panel methods to determine the number of individual series which follow $I(1)$ versus $I(0)$ processes. However, often one is interested not only in the individual series properties, but also the implications of the linkages among the individual series, and most importantly the cross-sectional dependencies that are driven by the common stochastic trends.

An important component of the non-stationary panel literature has been the literature on testing for unit roots in panels. A popular approach to accommodating what have been considered fairly general forms of cross-sectional dependence within this literature has been the factor model approach. An underlying assumption of this approach is the decomposition of the series of the panel into what are assumed to be independent common and idiosyncratic components. The idiosyncratic components are then tested for unit roots and the common components are tested either for unit roots in the single factor case, or cointegration in the

\(^1\)In the case of panels of univariate time series, cointegration and cross-unit cointegration are essentially synonyms (see the discussion in Section 2). We use the term cross-unit cointegration to conform with the conventions of the panel unit root literature. For multivariate time series panels there are, however, conceptual differences (see Wagner and Hlouskova, 2010).
multiple factor case. However, in many applications, such as the income convergence illustration we provide in Section 5, one is not interested to know the unit root versus stationarity properties of these separate components, but rather one is interested to know these properties about the individual raw series. For such cases we argue that our rank test approach is the best suited and most general approach available for panels with moderate to large cross-sectional dimensions.

In this regard, our tests do not impose any restrictions on the cross-sectional dependencies of the series. The only restriction on the series’ behavior is that a joint functional central limit theorem must hold for the first differences of the $N$-vector of series. In such a general setup, except for the extreme cases when the rank is either full (so that all series are $I(1)$ and not cointegrated), or zero (so that all series are $I(0)$), all series will in general be $I(1)$ and cointegrated. In particular, our computationally simple tests are based on multivariate variance ratios computed from tuning parameter free estimates of the respective components that do not require, and are thus not affected by, choices with respect to kernel and bandwidth, lag augmentation or the number of factors to be extracted. These estimators are based on advances in long-run variance estimation pioneered by Kiefer and Vogelsang (2002a,b).

Because the tests are cointegration rank tests, they can be used to infer any rank, and not just the null hypothesis of full rank, which is standard in the literature on non-stationary panels. In fact, the forms of hypotheses considered within this literature are very limited. Specifically, while the null hypothesis is almost always taken to be that all $N$ series are $I(1)$, the alternative hypothesis is usually formulated as that at least some series are $I(0)$. This leaves a rejection of the null somewhat uninformative as it does not indicate how many $I(0)$ series there are. This issue is discussed to some extent by Pesaran (2012), who recommends “the (panel unit root) test outcome to be augmented with an estimate of the proportion of the cross-section units for which the individual unit root tests are rejected” (see page 545). Motivated in part by this recommendation, we also propose a sequential rank testing procedures that compare favorably with for example the Johansen (1995) vector autoregression (VAR) based approach for moderate values of $N$, and increasingly so for larger values of $N$. We also show that for the special case of panels without the presence of cross-unit cointegration, our rank tests can be used to test the same null hypothesis usually tested by conventional panel unit root tests, with the additional advantage that any rank can be tested. Accordingly, in

---

2According to one of the referees, “cointegration rank tests are of no use in testing for unit roots in panels”
this setting our tests allow one to determine the fractions of $I(1)$ versus $I(0)$ series present in the panel.\footnote{The fact that our asymptotic theory is for $N$ fixed and only $T \to \infty$ is another indication that our tests are closely related to the multivariate time series literature. However, there are also close links to the standard first-generation panel unit root literature for the case of cross-sectionally independent panels. For this case both pooled and group-mean versions of our tests are available, which are asymptotically normally distributed in a $T, N \to \infty$ setting, and which have been explored in previous drafts (see, for example, Pedroni and Vogelsang, 2005).}

The new tests have good small-sample properties, which is demonstrated in a series of Monte Carlo simulation experiments. The proposed sequential rank test procedure is also shown to compare well to the Johansen (1995) trace test. In terms of sample size, the comparative advantage of our test occurs when $N$ is moderately sized, in that it is smaller than the $T$ dimension, but larger than one can handle well with parametric based multivariate cointegration methods. Similarly, we also show that when used in place of panel unit root tests the new tests outperform widely-used first-generation as well as state-of-the-art second generation tests under the conditions for which these other tests were designed.\footnote{Note however that, by construction, since they are based on variance ratios, in contrast to panel unit root tests, our tests require that the cross-sectional dimension $N$ of the sample is not larger, and is preferably smaller, than the number of observations over time, $T$.}

The remainder of the paper is organized as follows. In Section 2 we first discuss the DGP's and assumptions used in our approach and then discuss the relationships of our setup to the assumptions and DGPs used in the existing non-stationary panel literature. Section 3 presents the rank tests, provides critical values and discusses the local asymptotic power (LAP) properties of our tests for the special case of cross-sectionally independent panels. In this way, by comparing the LAP of our tests with the LAP of two widely-used first-generation panel unit roots, we seek to demonstrate that there is no cost to the generality of our approach even when the more restrictive assumption of cross-section independence is true. Next, in Section 4 we study the small-sample performance of our tests and compare and “any rank test can never be considered as a panel unit test since they do not test the same hypothesis”. We take a different view. One can analyze the properties of rank tests when the null and alternative hypotheses align with those that are typically of interest when panel unit root tests are used. As we show in the paper, the asymptotic null distributions of the rank tests are well defined and pivotal when all series are $I(1)$ and there is no cointegration (a typical assumption that is made under the null hypothesis of panel unit root tests). In this case the rank tests have power to reject this null when all or some of the series are $I(0)$ (the typical alternative of panel unit root tests). Suppose that all series are $I(1)$ but there is cointegration among these $I(1)$ series. If one wanted to test the panel unit root null that “all series are $I(1)$”, then the rank tests are not informative and we agree with the referee. To test this null hypothesis one would need to use a panel unit root test that tests “all series are $I(1)$” and permits cointegration among $I(1)$ series. Our view is that sometimes ranks tests can be used to test “all series are $I(1)$” and sometimes not depending on what one can, or is willing, to assume about cointegration among $I(1)$ variables. We disagree that ranks tests can never be used to test for unit roots in panels.
our tests with several second-generation tests under the conditions for which these tests were designed. Finally, we also compare the small-sample performance of our sequential rank test procedure with the Johansen trace test in this section. Section 5 in turn contains a brief empirical illustration of the rank tests taken from the growth and convergence literature. Section 6 offers concluding remarks.

2 Assumptions and Model Discussion

2.1 Assumptions

The DGP is stated in terms of the $N$-dimensional vector of time series $y_t = [y_{1t}, \ldots, y_{Nt}]'$, and is given by

$$y_t = \alpha_p d_p^t + u_t, \quad (1)$$

with observations available for $t = 1, \ldots, T$. Here $d_p^t = [1, t, \ldots, t^p]'$, for $p \geq 0$, is a polynomial trend function (with $d_0^t = 1$) and $\alpha_p$ is the associated matrix of trend coefficients.$^5$ The typical specifications considered for $d_p^t$ include a constant ($p = 0$) or a constant and a linear time trend ($p = 1$). These are also the specifications considered in the current paper; and for which critical values are provided.

Clearly, the unit root and cointegration properties of $y_t$ are governed by the respective properties of the stochastic component $u_t = [u_{1t}, \ldots, u_{Nt}]'$. In order to describe the unit root and cointegration properties of $u_t$, we introduce an $N \times N$ ortho-normal matrix $C = [C_1, C_2]$, whose component matrices $C_1$ and $C_2$ are of dimensions $N \times (N - c)$ and $N \times c$. The columns of the matrix $C_1$ span a basis of the cointegrating space of $u_t$, while the columns of $C_2$ span the space of the common stochastic trends.$^6$ The matrix $C$ thus rotates $u_t$ into $I(0)$ and $I(1)$ subsystems as

$$w_t = C'u_t = \begin{bmatrix} C_1'u_t \\ C_2'u_t \end{bmatrix} = \begin{bmatrix} w_{1t} \\ w_{2t} \end{bmatrix}, \quad (2)$$

where $w_{1t}$ is $I(0)$, while $w_{2t}$ is $I(1)$ with no cointegration. The corresponding vector of $I(0)$

---

$^5$Clearly, more general deterministic components can be considered to accommodate for example seasonal dummies or breaks in trends.

$^6$Note that we use the term common stochastic trends as it is used in the cointegration rank literature. This is not to be confused with the number of common $I(1)$ factors which one might consider in a factor model context; see Section 2.2 for further discussion.
errors is given by
\[ v_t = \begin{bmatrix} w_{1t} \\ \Delta w_{2t} \end{bmatrix} = \begin{bmatrix} v_{1t} \\ v_{2t} \end{bmatrix}, \] (3)
whose long-run covariance matrix is the key quantity for the testing approach used in this paper. For two jointly I(0) vector processes \( a_t \) and \( b_t \) with mean zero and absolutely summable covariance function, their long-run covariance matrix is defined as
\[ \Omega_{ab} = \sum_{s=-\infty}^{\infty} E(a_t b'_{t-s}) = \Sigma_{ab} + \Gamma_{ab} + \Gamma'_{ab}, \]
where \( \Sigma_{ab} = E(a_t b'_{t}) \) and \( \Gamma_{ab} = \sum_{s=1}^{\infty} E(a_t b'_{t-s}) \). The long-run covariance matrix of \( v_t \) is partitioned in the following way:
\[ \Omega_{vv} = \begin{bmatrix} \Omega_{v1v1} & \Omega_{v1v2} \\ \Omega_{v2v1} & \Omega_{v2v2} \end{bmatrix}. \]

The following high level Assumption 1 is enough to obtain our results.

**Assumption 1.** As \( T \to \infty \),
\[ \frac{1}{\sqrt{T}} \sum_{t=1}^{[sT]} v_t \to_w \Omega_{vv}^{1/2} W(s), \]
where \( \Omega_{vv} \) is positive definite, \( \to_w \) and \( [x] \) signify weak convergence and the integer part of \( x \), respectively, and \( W(s) = [W_1(s), \ldots, W_N(s)]' \) is an \( N \times 1 \) vector of independent standard Brownian motions that is partitioned conformably with \( v_t \).

Assumption 1 is stated directly in terms of the required invariance principle rather than a specific set of underlying primitive assumptions, since for the purposes of this paper we are not specifically interested in which set of technical assumptions leads to the required functional convergence results. There is a variety of setups that lead to Assumption 1 (see, for example, Phillips and Durlauf, 1986; Phillips and Solo, 1992).

### 2.2 Discussion

The cointegration properties of \( y_t \), or, equivalently, \( u_t \), can be read off the long-run covariance matrix of \( \Delta u_t \), \( \Omega_{\Delta u \Delta u} \). The rank of this matrix is equal to the number of (non-cointegrated) common stochastic trends, \( c \), and the cointegrating space is spanned by its left kernel.\(^7\) This

\(^7\)See also the previous footnote.
is evident from
\[
\Omega_{\Delta u \Delta u} = C \Omega_{\Delta v_1 \Delta v_1} C' = C \begin{bmatrix}
\Omega_{\Delta v_1 \Delta v_1} & \Omega_{\Delta v_1 \Delta v_2} \\
\Omega_{\Delta v_2 \Delta v_1} & \Omega_{\Delta v_2 \Delta v_2}
\end{bmatrix}
C' = C \begin{bmatrix}
0_{(N-c) \times (N-c)} & 0_{c \times c} \\
0_{c \times c} & \Omega_{\Delta v_2 \Delta v_2}
\end{bmatrix}
C' = C_2 \Omega_{\Delta v_2 \Delta v_2} C_2',
\]
since the long-run covariance matrices \(\Omega_{\Delta v_1 \Delta v_1}\) and \(\Omega_{\Delta v_2 \Delta v_2}\) are zero by construction. This shows that there is (cross-unit) cointegration, that is, a cointegrating relationship between at least two series, whenever the left-kernel of \(C_2\) is not spanned by vectors of the form \(\beta_i = [0, \ldots, 0, 1, 0, \ldots, 0]'\). Upon appropriate re-ordering of the series, this means that cross-unit cointegration occurs when the matrix \(C_2\) is not of the form \(0_{(N-c) \times c} C_2'\), with \(C_2 \in \mathbb{R}^{c \times c}\) a full rank matrix. This in turn means that, upon re-ordering, the matrix \(\Omega_{\Delta u \Delta u}\) is of the form \(\text{diag}(0_{(N-c) \times (N-c)}, \Omega_{22})\), with \(\Omega_{22}\) positive definite. This also shows the obvious fact that allowing for cross-sectional dependencies whilst excluding cross-unit cointegration is a very special case for \(I(1)\) time series panels.\(^8\)

As soon as one allows for cross-unit cointegration, the rank of the long-run variance matrix \(\Omega_{\Delta u \Delta u}\) is only a lower bound for the number of \(I(1)\) series in the panel. The number of \(I(1)\) series in \(\mathbf{u}_t\) depends upon the entries in the matrix \(C_2\) and can be any number between \(c\) and \(N\). Only in cases where \(N\) is sufficiently small can one perform inference on the matrix \(\mathbf{C}\), or its left-kernel, to determine the number of \(I(1)\) series, or indeed to identify the series that are \(I(1)\) by testing restrictions on the cointegrating space. If this is not possible and one would like to determine the number of \(I(1)\) series, the set of feasible DGPs must be constrained. One such example is the test of Ng (2008), which allows one to estimate the fraction of \(I(1)\) series, but only in the absence of cross-unit cointegration. Cross-unit cointegration must be excluded from the feasible set of DGPs because the test is based on the divergence behavior of the cross-sectional variances. Other approaches in the literature that try to overcome this fundamental limitation rely on multiple testing approaches (see, for example, Chortareas and Kapetanios, 2009; Hanck, 2012; Moon and Perron, 2012). However, by construction, multiple testing will typically lead to biased inference even asymptotically in the case where dependent test statistics are used in the multiple testing procedure. This can occur even when variants of modified Bonferroni procedures are used to treat the dependent test statistics. Accordingly, while multiple testing approaches can allow for similar

\(^8\)By construction, in the case of cross-sectionally independent series, the long-run covariance matrix \(\Omega_{\Delta u \Delta u}\) is diagonal with rank equal to the number of \(I(1)\) series in the panel, and one can set \(C = I_N\).
generalizations with respect to the feasible set of DGP s, the single test framework we adapt here has the advantage of avoiding the potential weakness of asymptotic bias arising from multiple dependent test statistics.

Within the panel unit root literature, the most prominent approach that is used to treat cross-unit cointegration is the factor model approach. The underlying assumption of factor models is that $u_t$ can be decomposed as

$$
\mathbf{u}_t = \Lambda \mathbf{f}_t + \mathbf{e}_t,
$$

where $\mathbf{f}_t$ is an $m \times 1$ vector of common factors with $\Lambda$ being the associated matrix of loading coefficients. A typical assumption used for this decomposition is that $\mathbf{f}_t$ and $\mathbf{e}_t$ are independent of each other, and both are at most $I(1)$ processes. Furthermore, it is typically assumed that the elements of $\mathbf{e}_t$ are either independent of each other or only weakly correlated, an assumption that excludes cointegration.

For the factor model the long-run covariance matrix of $\Delta \mathbf{u}_t$ is given by

$$
\Omega_{\Delta u \Delta u} = \Lambda' \Omega_{\Delta \mathbf{f} \Delta \mathbf{f}} \Lambda + \Omega_{\Delta \mathbf{e} \Delta \mathbf{e}}.
$$

This shows that the long-run covariance matrix of $\Delta \mathbf{u}_t$ is given by the sum of a component of dimension $m \times m$ multiplied from the left and right by $\Lambda'$ and $\Lambda$ and an “essentially” diagonal component $\Omega_{\Delta \mathbf{e} \Delta \mathbf{e}}$. The additive decomposition of the long-run covariance matrix into a reduced rank component, $\Lambda' \Omega_{\Delta \mathbf{f} \Delta \mathbf{f}} \Lambda$, and an essentially diagonal component, $\Omega_{\Delta \mathbf{e} \Delta \mathbf{e}}$, is the key simplifying assumption of factor models. It allows one to analyze the common factors and the idiosyncratic components separately with standard methods when properly taking into account the small-sample decomposition errors. Letting both $N$ and $T$ go to infinity allows for the consistent estimation of the space spanned by the factors, and thereby allows one to study the unit root and cointegration properties of the factors. Under the assumption of cross-sectional independence of the idiosyncratic components, modified standard panel unit root tests can be applied to test the null hypothesis of full rank for $\mathbf{e}_t$ (see, for example, Moon and Perron, 2004; Pesaran, 2007; Westerlund and Larsson, 2012).

Footnotes:

9The restrictions imposed by the factor structure allow one to identify the components separately. To be precise identification in approximate factor models critically rests upon $N, T \to \infty$ and thus when we speak of factor models we implicitly consider a sequence of nested models.

10Allowing for some correlation between the components of $\mathbf{e}_t$ is one of the defining characteristics of Bai and Ng’s (2004, 2010) approximate factor models.
Given that no inferential procedures are available for \( \Lambda \), if even just one of the common factors is \( I(1) \) this can lead one to conclude that all the series in the panel are \( I(1) \). Thus, only if the common factors are restricted a priori to be \( I(0) \) can one potentially use the factor model approach to determine the fraction of the series in a panel that are \( I(1) \) versus \( I(0) \) as could be done for example in an extension of the Ng (2008) test based on the estimated idiosyncratic components. Thus, only under this special case a priori restriction on the DGP, which excludes cross unit cointegration, does the orthogonal factor model restriction allow one to infer the fraction of \( I(1) \) versus \( I(0) \) series.\(^{11}\) By comparison our rank tests do not test properties of the common or idiosyncratic decompositions of the series, but rather test the raw (unfactored) series. Accordingly, when cross unit cointegration is permitted for the DGP, our approach essentially tests for the number of unit root processes driving the panel that are not cointegrated, which, following the cointegration rank literature, we refer to as the number of common stochastic trends, \( c \).\(^{12}\) Nevertheless, it is interesting to compare our approach to the Ng (2008) approach in the scenario for which the Ng approach was designed, namely the absence of cross unit cointegration, as we do in Section 4.

Another more recent promising approach to the treatment of cross-sectional dependencies in the panel unit literature is the block bootstrap approach of Palm et al. (2011). These authors consider bootstrapping the Levin et al. (2002) and Im et al. (2003) panel unit root tests and make the case that their bootstrap approach allows one to test for unit roots in a framework that allows for more general DGPs than does the factor model approach, and is thus fairly unrestricted if one is interested in testing the null hypothesis of full rank for the raw (unfactored) series. However, the approach is based on bootstrapping conventional panel unit root tests, and thus does not directly allow for testing of rank or the fraction of \( I(1) \) versus \( I(0) \) series in a panel. Nonetheless, for the special case of testing for full rank of the unfactored series, it is interesting to compare the small-sample performance of the

\(^{11}\)We should note however that in addition to their conceptual simplicity, factor models also have an advantage in that they can be applied for panels with arbitrary cross-sectional dimensions, thus including the finite samples where \( N > T \).

\(^{12}\)Notice that this does not necessarily refer to the number of common \( I(1) \) factors as one might envision in a factor model setting. For example, under the factor model scenario, if there were no \( I(1) \) components in the common factors and only two of the idiosyncratic series were \( I(1) \), we would conclude that the rank is \( c = 2 \), and hence that there are two common stochastic trends driving the panel. Similarly, if there were two \( I(1) \) components among the common factors and none of the idiosyncratic series were \( I(1) \), we would also conclude that the rank is \( c = 2 \), and hence that there are two common stochastic trends driving the panel. In many applications, as for example in the income convergence illustration we provide in Section 5, one is interested to know whether the number of common stochastic trends driving the raw panel of time series is large or small.
approach as we do in Section 4.

3 The Rank Tests

In this section we develop the tests which are designed to test any null hypothesis \( H_0 : 1 \leq c \leq N \) versus any alternative hypothesis \( H_1 : c_1 < c \).

3.1 The Statistics and Their Limiting Null Distributions

The test statistics are based on two ingredients; regressions involving superfluous deterministic trend terms (see, for example, Park, 1990 and Park and Choi, 1988, and long-run variance estimation based on untruncated kernels (see Kiefer and Vogelsang, 2002a).

We begin with long-run variance estimation. Consider the ordinary least squares (OLS) residuals of the regression of \( y_t \) on \( d_t \), which gives estimates of the stochastic component \( u_t \):\[
\hat{u}_t = y_t - \sum_{t=1}^{T} y_t d_t \left( \sum_{t=1}^{T} d_t d_t' \right)^{-1} d_t'.
\]
Formally, abstracting for the moment from the fact that \( u_t \) is an I(1) process, at least under the null hypothesis, a typical long-run variance estimator is given by

\[
\hat{\Omega}_{p,M} = \frac{1}{T} \sum_{t=1}^{T} \hat{u}_t^p \hat{u}_t'^p + \sum_{j=1}^{M} k(j/M) \sum_{t=j+1}^{T} \left( \hat{u}_t^p \hat{u}_{t-j}^p + \hat{u}_t^p \hat{u}_{t-j}^p \right),
\]

where \( k(x) \) is a kernel function, normalized to \( k(0) = 1 \), and \( M = M(T) \) is the bandwidth, using the observed residuals \( \hat{u}_t \) in place of the unobserved \( u_t \). For such an estimator to have a well-defined limit, the process \( u_t \) needs to possess a finite long-run variance and certain conditions on the kernel and bandwidth have to be in place (see, for example, Jansson, 2002). The last decade has seen a shift in long-run variance estimation to consider bandwidths that are proportional to the sample size, which under appropriate assumptions, leads to asymptotic results that account for bandwidth and kernel choices (see Kiefer and Vogelsang, 2002a,b). In this respect the use of untruncated kernels, that is, when \( M = T \), leads to particularly simple results useful for our purposes. For example, for the Bartlett kernel, \( k(x) = 1 - |x|/T \), Kiefer and Vogelsang (2002b) show that the expression for \( \hat{\Omega}_{p,T} \) simplifies to

\[
\hat{\Omega}_{p,T} = \hat{\Omega}_p = \frac{2}{T^2} \sum_{t=1}^{T} \hat{S}_t^p\hat{S}_t^p',
\]
where \( \hat{\Sigma}_p = \sum_{t=1}^{T} \hat{u}_p^t \). This is the only long-run variance estimator considered, implying that our tests do not require user specific choices concerning bandwidth or kernel function.

In our case, where at least under the null hypotheses considered, \( u_t \) is an \( I(1) \) process, a long-run variance estimator as given above will diverge for any kernel and bandwidth choice. Convergence towards a well-defined limiting distribution requires additional scaling. Based on the above discussion, the appropriate rescaling of the estimator requires rotating the series into the asymptotically \( I(0) \) and \( I(1) \) components by considering \( \hat{\mathbf{w}}_p^t = \mathbf{C}' \hat{u}_p^t \). Define \( \mathbf{D}_T = \text{diag}(\mathbf{I}_{N_L}, T^{1/2}\mathbf{L}_c) \). By using Assumption 1, rotation by \( \mathbf{C} \), and using standard results for OLS detrended variables it then follows that

\[
\frac{1}{T} \mathbf{D}_T^{-1} \hat{\Omega}_p \mathbf{D}_T^{-1} = 2\mathbb{C} \frac{1}{T^3} \sum_{t=1}^{T} \begin{bmatrix} \hat{\mathbf{R}}_1^p \hat{\mathbf{R}}_2^p & \hat{\mathbf{R}}_1^p \hat{\mathbf{R}}_2^p & \hat{\mathbf{R}}_1^p \hat{\mathbf{R}}_2^p \\
\hat{\mathbf{R}}_2^p \hat{\mathbf{R}}_2^p & \hat{\mathbf{R}}_2^p \hat{\mathbf{R}}_2^p & \hat{\mathbf{R}}_2^p \hat{\mathbf{R}}_2^p \\
\hat{\mathbf{R}}_1^p \hat{\mathbf{R}}_2^p & \hat{\mathbf{R}}_1^p \hat{\mathbf{R}}_2^p & \hat{\mathbf{R}}_1^p \hat{\mathbf{R}}_2^p \\
\end{bmatrix} \mathbf{C}'
\]

\[
\rightarrow_{\mathbb{W}} 2\mathbb{C} \begin{bmatrix} 0 & 0 \\
0 & 0 & \int_0^1 \mathbf{R}_2^p(s)\mathbf{R}_2^p(s)'ds \\
\end{bmatrix} \mathbf{C}'
\]

as \( T \to \infty \), where \( \hat{\mathbf{R}}_1^p = \sum_{t=1}^{T} \hat{\mathbf{w}}_p^t \), \( \mathbf{R}_1^p(s) = \int_0^s \mathbf{B}_1^p(r)dr \) and \( \mathbf{B}_1^p(s) = \int_0^1 \mathbf{B}_1^p(s)'ds \) are partitioned conformably with \( \mathbf{C} \), similarly to \( \hat{\mathbf{w}}_p^t \). The usual contemporaneous variance estimator\(^{13}\),

\[ \hat{\Sigma}_p = \frac{1}{T} \sum_{t=1}^{T} \hat{\mathbf{u}}_p^t \hat{\mathbf{u}}_p^t, \]

also needs to be normalized in order to achieve convergence, such that

\[
\mathbf{D}_T^{-1} \hat{\Sigma}_p \mathbf{D}_T^{-1} = \mathbb{C} \frac{1}{T} \sum_{t=1}^{T} \begin{bmatrix} \hat{\mathbf{w}}_1^p \hat{\mathbf{w}}_2^p & \hat{\mathbf{w}}_1^p \hat{\mathbf{w}}_2^p & \hat{\mathbf{w}}_1^p \hat{\mathbf{w}}_2^p \\
\hat{\mathbf{w}}_2^p \hat{\mathbf{w}}_2^p & \hat{\mathbf{w}}_2^p \hat{\mathbf{w}}_2^p & \hat{\mathbf{w}}_2^p \hat{\mathbf{w}}_2^p \\
\hat{\mathbf{w}}_1^p \hat{\mathbf{w}}_2^p & \hat{\mathbf{w}}_1^p \hat{\mathbf{w}}_2^p & \hat{\mathbf{w}}_1^p \hat{\mathbf{w}}_2^p \\
\end{bmatrix} \mathbf{C}'
\]

\[
\rightarrow_{\mathbb{W}} \mathbb{C} \begin{bmatrix} \Sigma_{\mathbf{v}_1 \mathbf{v}_1} & 0 \\
0 & 0 & \int_0^1 \mathbf{B}_2^p(s)\mathbf{B}_2^p(s)'ds \\
\end{bmatrix} \mathbf{C}'
\]

The results in (5) and (6) imply that test statistics with nuisance parameter free limiting distributions can be constructed by simply using appropriately normalized matrix ratios of \( \hat{\Sigma}_p \) and \( \hat{\Omega}_p \). The first test statistic of this type that we consider can be seen as motivated by the \( \hat{\delta}_T \) statistic introduced by Breitung (2002) in a time series context and is given by

\[ MB = \frac{1}{2T} \text{tr}(\hat{\Omega}_p \hat{\Sigma}_p^{-1}). \]

\(^{13}\)Of course, with \( I(1) \) processes, this quantity is not an estimator of a variance, just as \( \hat{\Omega}_{p,T} \) is not a long-run variance estimator in the \( I(1) \) case.
The asymptotic distribution of this statistic under any null hypothesis $H_0 : 1 \leq c \leq N$ is easily derived from the above results:

$$MB = \text{tr} \left( \frac{1}{2T} \tilde{D}_T^{-1} \tilde{\Omega}_T \tilde{D}_T^{-1} (\tilde{D}_T^{-1} \tilde{\Sigma}_T \tilde{D}_T^{-1})^{-1} \right)$$

$$\rightarrow_w \text{tr} \left( C \left[ \begin{array}{cc} 0 & 0 \\ 0 & \int_0^1 R_2^p(s)R_2^p(s)'ds \end{array} \right] \right) C' \left( C \left[ \begin{array}{cc} \Sigma_{\tilde{\varepsilon}T} & 0 \\ 0 & \int_0^1 B_2^p(s)B_2^p(s)'ds \end{array} \right] \right)^{-1} \right)$$

$$= \text{tr} \left( \int_0^1 R_2^p(s)R_2^p(s)'ds \left( \int_0^1 B_2^p(s)B_2^p(s)'ds \right)^{-1} \right)$$

$$= \text{tr} \left( \int_0^1 Q_2^p(s)Q_2^p(s)'ds \left( \int_0^1 W_2^p(s)W_2^p(s)'ds \right)^{-1} \right),$$

where $Q_2^p(s) = \int_0^s W^p(r)dr$ is partitioned conformably with $C$.

The second test statistic that we consider is based on the properties of regressions that include superfluous deterministic trend regressors. Towards this end, suppose that the data are generated as before via (1) but that the trend polynomial used in the OLS detrending is now of degree $q > p$. If $u_t$ is $I(0)$, then the coefficients corresponding to the superfluous trends $t^{q+1}, \ldots, t^q$ are estimated consistently to be zero as $T \rightarrow \infty$. Therefore, a coefficient restriction test such as the Wald test will have a well-defined limiting distribution in this case. On the other hand, if $u_t$ is $I(1)$, then the estimated trend coefficients corresponding to the superfluous trend regressors do not converge to zero. The associated Wald statistic is in this case $O_p(T)$, suggesting that the statistic divided by $T$ can be used as a rank test statistic. The following test statistic, which exploits these differences, can be seen as a multivariate analogue of the Wald-type $J$-statistic of Park and Choi (1988):

$$MJ = \text{tr} (\tilde{\Sigma}_T \tilde{\Sigma}_q^{-1} - I_N),$$

where $\tilde{\Sigma}_q$ is the estimated residual covariance matrix from (1) when the fitted trend polynomial is of degree $q > p$. Vogelsang (1998) uses the $J$-statistic as a means of testing deterministic trends in the univariate case and finds that the choice $q = 9$ leads to powerful tests. Unreported simulation results suggest that the same is true also in the present context. Therefore, for the simulation of critical values, the small-sample simulations and the empirical application we only consider $q = 9$. Similarly to before, under any null hypothesis
\( H_0 : 1 \leq c \leq N \) the limiting distribution is found to be

\[
MJ = \text{tr} \left( D_T^{-1} \hat{\Sigma}_q D_T^{-1} (D_T^{-1} \hat{\Sigma}_q D_T^{-1})^{-1} - I_N \right)
\]

\[
\rightarrow_w \text{tr} \left( C \begin{bmatrix} \Sigma_{v_1 v_1} & 0 \\ 0 & \int_0^1 B_2^p(s)B_2^p(s)' ds \end{bmatrix} C' \right)^{-1} - I_N
\]

\[
\times \left( C \begin{bmatrix} \Sigma_{v_1 v_1} & 0 \\ 0 & \int_0^1 B_2^q(s)B_2^q(s)' ds \end{bmatrix} C' \right)^{-1} - I_N
\]

\[
= \text{tr} \left( I_{N-c} \begin{bmatrix} 0 \\ 0 & \int_0^1 B_2^q(s)B_2^q(s)' ds (f_0^1 B_2^q(s)B_2^q(s)' - I) \end{bmatrix} - I_N \right)
\]

\[
= \text{tr} \left( \int_0^1 W_2^p(s)W_2^q(s)' ds (f_0^1 W_2^q(s)W_2^q(s)' - I) \right)
\]

as \( T \to \infty \), with \( B^q(s) \) and \( W^q(s) \) defined similarly to \( B^p(s) \) and \( W^p(s) \) above.

Let us now consider the behavior of the test statistics under the conventional \( I(0) \) zero rank alternative \( H_1 : c = 0 \). Given that under this alternative \( \hat{R}_{11}^p/\sqrt{T} \rightarrow_w B_1^p(s) \) and \( C = C_1 = I_N \) we obtain \( \hat{\Sigma}_p = \frac{2}{T} \sum_{t=1}^T \hat{R}_{11}^p \hat{R}_{11}^{pp'} \rightarrow_w 2\int_0^1 B_1^p(s)B_1^p(s)' ds \) and \( \hat{\Sigma}_q = \frac{1}{T} \sum_{t=1}^T \hat{w}_{11}^p \hat{w}_{11}^{pp'} \rightarrow_p \Sigma_{v_1 v_1} \), and therefore

\[
TMJ \rightarrow_w \text{tr} \left( \int_0^1 B_1^p(s)B_1^p(s)' ds \Sigma_{v_1 v_1}^{-1} \right)
\]

so that \( MB = O_p(1/T) \), while, because of the consistency of both \( \hat{\Sigma}_p \) and \( \hat{\Sigma}_q \),

\[
MJ \rightarrow_w \text{tr} (\Sigma_{v_1 v_1} \Sigma_{v_1 v_1}^{-1} - I_N) = 0.
\]

Hence, in this case both statistics degenerate to zero.\(^{14}\) Thus, the null hypothesis will be rejected for these tests when the test statistics assumes a value smaller than a critical value, which depends upon the null rank as well as the deterministic components.

For any other alternative \( H_1 : 0 < c_1 < c \), the \( MB \) and \( MJ \) statistics converge to the traces of similar random matrices as under the null, but with dimensions corresponding to the rank under the alternative. Accordingly, these test statistics can be used flexibly to test for any value of the rank. It should be noted, however, that it is only when the alternative is \( H_1 : c_1 = 0 \) that \( MB \) and \( MJ \) are consistent. For all other alternatives, while power is retained, the tests are not consistent since their power does not tend to one asymptotically. This is, of course, a common feature of cointegration rank tests.

\(^{14}\)The exact rate at which \( MJ \) converges to zero depends on the number of finite moments of \( w_t \). With finite eighth moments of \( v_t \) it can be shown that \( MJ = O_p(1/\sqrt{T}) \).
These observations lead us to consider a variant of the MB statistic that is consistent when testing the null hypothesis $H_0 : c = N$ against any alternative $H_1 : 0 \leq c_1 < N$. Specifically, the inverse of the Breitung (2002) $\hat{\varrho}_T$ statistic is considered:

$$MIB = 2 T \text{tr} (\hat{\Sigma}_p \hat{\Omega}_p^{-1}).$$

Under the full rank null hypothesis it follows directly from our previous results that

$$MIB \rightarrow_w \text{tr} \left( \int_0^1 W^p_2(s) W^p_2(s)' ds \left( \int_0^1 Q^p_2(s) Q^p_2(s)' ds \right)^{-1} \right).$$

By construction, the $MIB$ statistic diverges should the rank be less than full. To show this, we first rewrite $MIB$ equivalently as

$$MIB = 2 T \sum_{i=1}^N \hat{\lambda}_i,$$

where $\hat{\lambda}_1 \geq \ldots \geq \hat{\lambda}_N$ are the ordered eigenvalues of $\hat{\Sigma}_p \hat{\Omega}_p^{-1}$ ordered according to decreasing absolute value.\(^{15}\) Suppose that the rank is less than full such that $N - c > 0$. Then, using $\hat{\lambda}_1, \ldots, \hat{\lambda}_{N-c}$ to denote the eigenvalues corresponding to the $I(0)$ series,

$$\frac{1}{T} MIB = 2 \sum_{i=1}^{N-c} \hat{\lambda}_i + o_p(1) \rightarrow_w 2 \sum_{i=1}^{N-c} \lambda_i = \text{tr} \left( \Sigma_{v_1 v_1} \left( \int_0^1 B^1_1(s) B^1_1(s)' ds \right)^{-1} \right),$$

where $\lambda_1, \ldots, \lambda_{N-c}$ are the eigenvalues of $\Sigma_{v_1 v_1} (\int_0^1 B^1_1(s) B^1_1(s)' ds)^{-1}$. Thus, $MIB = O_p(T)$, suggesting that, unlike the other tests, $MIB$ is consistent against all alternatives $c_1 < N$, and not just against $c_1 = 0$. For this test statistic the critical region consists of all values larger than a critical value, which itself depends upon the null rank as well as the deterministic components.

The $MIB$ test statistic can be used, by slight modification, to test any null hypothesis $H_0 : 1 \leq c < N$. In this case the largest $N - c$ eigenvalues in the $MIB$ statistic and thus the $MIB$ statistic diverge. To circumvent this we use a modified statistic that is based on only the eigenvalues $\hat{\lambda}_{N-c+1}$ to $\hat{\lambda}_N$, which we refer to as modified $MIB$, briefly $MMIB$ test statistic:

$$MMIB = 2 T \sum_{i=N-c+1}^N \hat{\lambda}_i.$$

\(^{15}\)In the event that one wants to ensure real eigenvalues, one can use a symmetrized version of the test statistic, such as $2 T \text{tr} (\hat{\Sigma}_p \hat{\Omega}_p^{-1} + \hat{\Omega}_p^{-1} \hat{\Sigma}_p)$. The corresponding implications for the limiting distributions are straightforward.
which in fact coincides with the $\Lambda_q$ statistic studied by Breitung (2002). Under the null hypothesis that the rank is equal to $c$,

$$\text{MMIB} \rightarrow \text{tr} \left( \int_0^1 W_2^p(s)W_2^p(s)'ds \left( \int_0^1 Q_2^p(s)Q_2^p(s)'ds \right)^{-1} \right),$$

whereas under any alternative $c_1 < c$, $\text{MMIB} = O_p(T)$. By construction, for $H_0 : c = N$, the $\text{MIB}$ and $\text{MMIB}$ test statistics coincide.$^{16}$

The $\text{MMIB}$ statistic naturally lends itself to a recursive procedure for determining the rank $c$, which is a common approach in the cointegration literature (see, for example, Johansen, 1995). Specifically, one begins the testing sequence with an initial null hypothesis, such as for example $c = N$. If this null hypothesis is not rejected, one concludes that all the cross-section units are $I(1)$, and that there is consequently no cross-unit cointegration. In this case the procedure is stopped. If, on the other hand, the null $c = N$ is rejected, the procedure continues by testing the null $c = N - 1$, this time using the $\text{MMIB}$ statistic based on all but the largest eigenvalue. The testing procedure then continues by sequentially dropping the largest eigenvalue until the null hypothesis cannot be rejected, or until zero rank is reached.

Note that although a similar sequential procedure can in principle also be performed with the $\text{MB}$ and $\text{MJ}$ tests, this is generally not recommended. The reason is that the resulting procedures will only be able to discriminate between full and zero rank with unit asymptotic power, but will have asymptotically diminished power for intermediate cases.$^{17}$

### 3.2 Critical Values

To effectively summarize the critical values which are simulated to reflect the dependencies upon the null rank $c$, the deterministic components and the time series dimension $T$, we rely on the use of response surface regressions.$^{18}$ Experimenting with a variety of

$^{16}$Clearly, one can also base the other test statistics on only the subset of eigenvalues corresponding to the null rank, that is, one could neglect the eigenvalues converging to zero under the null for the $\text{MB}$ and $\text{MJ}$ statistics. Since we do not recommend using these test statistics in a sequential fashion we do not pursue this avenue here. See the discussion below for further details.

$^{17}$Another possibility is to consider maximum eigenvalue rather than trace statistics. However, unreported simulation results suggest that the trace statistics perform better in small samples, and in what follows we therefore only consider these. Similar observations have been made for VAR cointegration analysis, for which trace-type statistics tend to be preferred.

$^{18}$The convergence to stochastic integrals occurs for $T \to \infty$ and thus suggests only using large $T$ values for the simulation. It has, however, been found in the literature that using finite-sample critical values may lead to improved test performance of panel unit root and cointegration tests (see, for example, Hlouskova and Wagner, 2009).
specifications leads to linear regression models of the form \( \tau = \delta'Z + \eta \), where \( \tau \) is the simulated 5% critical value and \( \eta \) is an error term. The choice of regressors \( Z \) included is dictated by overall significance subject to the requirement that the \( R^2 \) of the regression be no smaller than 0.999. The set of regressors retained for the MIB (MMIB) and MJ tests is \( Z = (1,c^{1/4},c,c^2,c^3,c^2/T,c^3/T,1/T,1/T^2,c/T^2,c^2/T^2)' \), while for MB we retain \( Z = (1,1/c^{1/4},1/c,1/c^2,1/c^3,1/Te^2,1/Te^3,1/T,1/T^2,1/T^2c,1/T^2c^2)' \).

The simulated critical values are based on 1,000 draws from finite-\( T \) discrete counterparts of the stochastic integral limit distributions of each of the three test statistics, with normal random walks of dimensions \( c = 1,2,\ldots,50 \) and lengths \( T = \max\{30,2c\}, \max\{30,2c\} + 5,\ldots,300 \) in place of the vector Brownian motion \( W(s) \). This implies that there are a total of 2,165 observations available for each regression. The resulting estimated response surface coefficients are reported in the top panel of Table 1.

Unreported simulation results suggest that the fit of the response surface regressions can be poor when \( c \) is close to the sample endpoints of the response surface regression. To compensate for this we simulate critical values for \( c = 1,2,\ldots,5 \) and \( T = 1,000 \). These are reported in the bottom panel of Table 1.

### 3.3 Local Asymptotic Power

In this section we examine the LAP of our tests by considering the asymptotic behavior in the case of near-\( I(1) \) processes (see Phillips, 1988). In particular, we consider \( w_t = w_{2t} \) and replace \( \Delta w_{2t} = v_{2t} \) in (2) with

\[
\Delta w_{2t} = \frac{1}{T} g w_{2t-1} + v_{2t},
\]

where \( g \) is an \( N \times N \) drift parameter matrix that in our case is considered to be diagonal, that is, \( g = \text{diag}(g_1,\ldots,g_N) \). Clearly, setting \( g = 0 \) leads us back to the case of non-cointegrated \( I(1) \) series.\(^{19}\)

By using the invariance principle for near-integrated processes given in Phillips (1988, Lemma 3.1), we obtain

\[
\frac{1}{\sqrt{T}} w_{2t} \to_w \Omega_{1/2}^{1/2} J_g(s)
\]

\(^{19}\)Setting \( w_t = w_{2t} \) is not necessary. One can also include \( I(0) \) components \( w_{1t} \) to have a mix of \( I(0) \) and near-\( I(1) \) series.
as $T \to \infty$, where $J_{g}(s) = \int_{0}^{s} \exp((s - r)g) dW_{2}(r)$ is a vector standard Ornstein–Uhlenbeck process. This means that in order to obtain the LAP functions of the rank statistics, the process $W_{2}(s)$ in the limiting null distributions should be replaced by $J_{g}(s)$. For example, in case of the MB statistic,

$$MB \to_{w} \text{tr} \left( \int_{0}^{1} K_{g}^{p}(s)K_{g}^{p}(s)' ds \left( \int_{0}^{1} J_{g}^{p}(s)J_{g}^{p}(s)' ds \right)^{-1} \right)$$

as $T \to \infty$, where $K_{g}^{p}(s) = \int_{0}^{s} J_{g}^{p}(r) dr$ with $J_{g}^{p}(r)$ being the detrended version of $J_{g}(r)$.

It is instructive to compare LAP of our rank tests with LAP of some existing first-generation panel unit root tests, like the Im et al. (2003) and Levin and Lin (1992) statistics, henceforth denoted by IPS and LL, for cross-sectionally independent data. Such a comparison allows us to gauge the “price” to be paid for our more general setup even when the generality of our setup is not required by the DGP. For $\Omega_{\text{cov}} = I_{N}$ and $g = g_{l}I_{N}$ it can be shown that as $T \to \infty$ for fixed $N$, the limiting distributions of these tests are given by:

$$IPS \to_{w} \frac{1}{\sigma \sqrt{N}} \sum_{i=1}^{N} \left( 8 \sqrt{\int_{0}^{1} (I_{g_{i}}^{p})(s)^{2} ds + \frac{\int_{0}^{1} I_{g_{i}}^{p}(s) dW_{2}(s)}{\sqrt{\int_{0}^{1} (I_{g_{i}}^{p})(s)^{2} ds}} - \mu \right)$$

$$LL \to_{w} \frac{1}{\sigma} \left( 8 \sqrt{\int_{0}^{1} J_{g}^{p}(s)J_{g}^{p}(s)' ds + \frac{\int_{0}^{1} J_{g}^{p}(s)' dW_{2}(s) ds}{\sqrt{\int_{0}^{1} J_{g}^{p}(s)' J_{g}^{p}(s) ds}} - \mu \right),$$

where $\mu$ and $\sigma^{2}$ are mean and variance adjustment terms, and $I_{g_{i}}^{p}(s)$ and $dW_{2}(s)$ are the elements of $J_{g}^{p}(s)$ and $dW_{2}(s)$, respectively.

Given the above results, local asymptotic power can be simulated in a similar fashion as the critical values, that is, by simulating discrete versions of $J_{g}(s)$. The results for $c = N = 10$ and varying $g_{i} = g \in [-8, 0]$ are reported in Figure 1 for the case when $p = 0$ and in Figure 2 for the case $p = 1$ for $g_{i} = g \in [-16, 0]$.

The first observation that is striking in Figures 1 and 2 is that the tests designed for cross-sectionally independent data do not do better than our more general tests even when the

---

20 Note that in the scalar case the above limiting distribution coincides with the one given in Appendix B of Breitung (2002) for his $\hat{r}_{T}$ test.

21 Since the rate of shrinking of the local alternative is given here by $1/T$, which is the same as in the time series literature, the LAP functions of IPS and LL are simply normalized cross-section averages of the LAP function of the augmented Dickey–Fuller test (see, for example, Phillips, 1988). The reason for considering the Levin and Lin (1992) test statistic, which is slightly different than the one that appears in Levin et al. (2002), is that it has been shown to lead to higher power (see Westerlund and Breitung, 2013). For an analysis of the LAP of the Im et al. (2003) test statistic, including an assessment of the impact of initial observations, see Harris et al. (2010).

22 Results for other values of $N$ are very similar and are therefore not presented.
simulated DGP conforms to the case of cross-sectional independence, for which the first-generation tests were designed. The $MJ$ test performs particularly well and has the highest LAP for almost all values of $g$. The largest gain in LAP occurs when $g$ is close to zero, for example when $p = 0$ and $g \in [-3, 0]$, the LAP of $MJ$ is almost twice as large as that of the $LL$ test, and many times higher than that of the $IPS$ test. The second main observation is that $MIB$ performs worst, which will be commented upon below.

The inclusion of a linear trend leads to a reduction of LAP for all tests. This is a reflection of the well-known incidental trends problem, from which all panel unit root tests are known to suffer (see, for example, Moon et al., 2007). Also, while $MJ$ is still ranked first, the presence of a linear trend changes the ranking of some of the other tests. For example while the $MB$ test ranks second when $p = 0$, when $p = 1$, $LL$ ranks second. The fact that $IPS$ is always dominated by $LL$ is due to the homogenous alternative considered here (see, for

\[ \text{Note: The horizontal and vertical axes display the drift parameter } g \text{ and power, respectively. The test abbreviations are as used in the main text.} \]
The local alternatives considered above are such that, as long as $g < 0$, all the eigenvalues that form the $MIB$ statistic are bounded as $T \to \infty$. This lack of divergence of eigenvalues explains the poor performance of $MIB$ in Figures 1 and 2. For $MIB$ to be powerful one should consider local alternatives of the form $g = \text{diag}(I_{N-c}, g I_c)$ for $c < N$, that is, a mix of $I(1)$ and near-$I(1)$ series. This is done in unreported simulations, which we describe here briefly. Choosing $c < N$ causes $MIB$ to outperform all other tests considered, and this is true for all values of $g$ and for both $p = 0$ and $p = 1$. Another observation is that the LAP of $IPS$ and $LL$ drops significantly when $c < N$. For example, when $c = 5$, $N = 10$ and $p = 1$, the LAP of $LL$ never goes above 46%.

Thus, from the perspective of LAP, the choice concerning which test to use ($MB$, $MJ$ or $MIB$) will depend upon the type of alternative with which one is concerned. If under the alternative a mix of $I(1)$ and $I(0)$ series is considered, the $MIB$ test is the best choice. If under the alternative all units are $I(0)$, the $MJ$ test is the best choice. In both cases, the best performing rank test outperforms the best performing first-generation test. This happens example, Westerlund and Breitung, 2013).
for a simple DGP that corresponds exactly to the situation for which the first-generation tests were originally developed. The case for the new tests is further strengthened when one considers that for practical application of the first-generation tests, potential performance deteriorating choices must also be made, such as lag length selection, whereas for the new tests no such choices need to be made.

4 Small-Sample Performance

In this section, we use Monte Carlo simulations to evaluate the small-sample properties of the new tests. The results are organized in two subsections. In Section 4.2, we investigate size and power, as well as the accuracy of the estimated ranks based on the sequential MMIB test. In Section 4.3, we compare the performance of the new tests with that of the second-generation tests of Ng (2008) when one wishes to test for the fraction of series that follow $I(1)$ processes, and with Bai and Ng (2004), and of Palm et al. (2011) when one wishes to test the specialized hypothesis of full rank. In each case we make comparisons under DGPs that correspond to the frameworks for which these comparison tests were designed. We also compare the performance of our sequential MMIB test for rank with the conventional trace-based cointegration rank test procedure of Johansen (1995).

4.1 Simulation Design

The simulation DGP is given by a restricted version of (1)-(3) that sets $a_p = 0$, $C_1 = [I_{N-c}, 0_{(N-c)\times c}]'$ and $C_2 = [0_{c \times (N-c)}, I_c]'$, such that $y_t = w_t$. The vector of $I(0)$ innovations is generated as an ARMA(1,1) process;

$$
\begin{bmatrix}
\Delta w_{1t} \\
\Delta w_{2t}
\end{bmatrix}
= \begin{bmatrix}
v_{1t} \\
v_{2t}
\end{bmatrix}
= \begin{bmatrix}
\rho I_{N-c} & 0 \\
0 & 0
\end{bmatrix}
\begin{bmatrix}
v_{1t-1} \\
v_{2t-1}
\end{bmatrix}
+ \epsilon_t + \theta \epsilon_{t-1},
$$

with $|\rho| < 1$, $v_{10} = 0_{(N-c)\times 1}$, $v_{20} = 0_{c\times 1}$ and $\epsilon_t \sim N(0_{N\times 1}, \Sigma)$. Three values of $\theta$ are considered: $\theta = 0$ (AR(1)), and $\theta = -0.8, 0.8$ (ARMA(1,1)). To ensure that $\Sigma$ is a symmetric positive definite matrix, we follow Chang (2004) and set $\Sigma = PVP'$, where $V = \text{diag} (\lambda_1, \ldots, \lambda_N)$ is a matrix of eigenvalues such that $\lambda_1 = 0.1$, $\lambda_N = 1$ and $\lambda_2, \ldots, \lambda_{N-1} \sim U(0.1, 1)$. Also, $P = U(U'U)^{-1/2}$, where the elements of the $N \times N$ matrix $U$ are all drawn from $U(0,1)$. The number of replications is set to 3,000, and for each unit of the panel we generate 100 pre-sample values that are then discarded. For brevity, we only report size and power at the.
5% level.\footnote{The power results are not size corrected because such a correction is generally not available in practice. Therefore, a test is useful for applied work only if it respects roughly the nominal significance level.}

ARMA(1,1) processes are the workhorses for simulation analysis of unit root and also cointegration tests (see Perron and Ng, 1996; Vogelsang and Wagner, 2013), since by variation of the autoregressive and moving average parameter many aspects relevant for the performance of such tests can be assessed. The specific DGP used here is the “canonical” DGP considered by Toda (1994) and Breitung (2002) among others.

4.2 Size, Power and Rank Selection Frequencies

In Tables 2, 3 and 4 we report size and power for the rank tests for $T = 100, 200$ and $N = 10, 20$. Consider first the upper panel of Table 2 where we report the size results for the null hypothesis of full rank $c = N$, for which the MIB and MMIB tests coincide. All simulation results reported in this table are generated with $\rho = 0.1$. The tests have good size accuracy when $\theta = 0$, but tend to be undersized for $\theta = 0.8$, reflecting the widely-documented detrimental effects that large positive MA coefficients exert on unit root and cointegration tests. Our tests are also not immune to this general weakness of (panel) unit root and cointegration tests. Increasing the time dimension beyond $T = 200$ alleviates the reported size distortions, whereas increasing the cross-sectional dimension $N$ does not have the same clear beneficial effects. In fact it is known for many panel unit root tests that size distortions may grow with $N$ especially for small values of $T$ (see Hlouskova and Wagner, 2006). Against this background it is comforting to see that the rank tests are not adversely affected by increasing values of $N$.

It is well known that large negative moving average coefficients are particularly troublesome in time series and panel time series contexts, and more troublesome than are large positive coefficients (see below). This is also the case here. Setting $\theta = -0.8$ causes all tests considered, including the Ng (2008), Bai and Ng (2004), Palm et al. (2011) and Johansen (1995) tests to be seriously distorted, with sizes that are close to 100% in most cases. The distortions decrease with increasing $T$. However, for the sample sizes considered here they are still unacceptably large, which is not surprising given that most, if not all, existing panel unit root tests have been shown to suffer from the same problem (see, for example, Hlouskova and Wagner, 2006).
The lower panel of Table 2 displays the results for testing $H_0 : c = 0.5N$. We know from the discussion in the previous section that in this case the $MIB$ statistic is divergent for $T \to \infty$, which is reflected in non-surprising rejection rates of 100% throughout all experiments. Compared to the full rank null, the other tests appear to be slightly more conservative. This holds true especially for the $MMIB$ test, which is severely undersized already in case $\theta = 0$.

We next consider power of the tests in Tables 3 and 4. In Table 3 we consider power against different alternative ranks $c_1$ given by various fractions of $N$. Preliminary simulations confirm the ex ante expectations that power is higher in case of $\rho = 0$ compared to $\rho = 0.9$. Consequently, in case $\rho = 0.9$ (upper panel) we consider smaller fractions of ranks given by $0.1N$, $0.3N$ and $0.7N$. For $\rho = 0$ (lower panel) the ranks considered under the alternative are $0.5N$, $0.8N$ and $0.9N$. In Table 4 we consider power for varying values of $\rho$ for given alternative ranks of either $c_1 = 0.8N$ (upper panel) or $c_1 = 0$ (lower panel). Similarly, to the setup considered in Table 3, in Table 4 we consider $\rho$ values closer to one when considering $c_1 = 0$ compared to $c_1 = 0.8$. In both tables $\theta = 0$ and, since the null hypothesis is that of full rank, the $MIB$ and $MMIB$ test statistics coincide.

The results can be summarized as follows: First, power increases with $T$, with the gap between the full null rank and the true alternative rank $c_1$, and as $\rho$ gets smaller. Increasing $N$ also leads to power gains, albeit this effect is not as pronounced and general as for increasing $T$. Second, power is affected by the inclusion of a linear trend in the model, which is in agreement with the LAP results depicted in Figures 1 and 2. Interestingly, there are several instances in which the power drop is not sizeable for the rank tests when a linear time trend is included. For the $MJ$ test in many of the experiments power is even higher when a linear trend is included (see the lower panel in Table 3 and the upper panel in Table 4), which is to a certain extent surprising given the LAP results.

Third, perfectly in line with the LAP findings, power behavior is somewhat complementary between the $MIB$ test on the one hand and the $MB$ and $MJ$ tests on the other. Often, and in particular when small deviations between null and alternative rank are considered, the $MIB$ test has highest power. When $H_1 : c_1 = 0$ is considered, the $MB$ and $MJ$ tests have higher power than the $MIB$ test. Fourth, the consistency property of the $MIB$ test under the alternative is “less clearly visible” when the autoregressive parameter $\rho$ is close to one. Under these circumstances, $MIB$ is dominated by the other tests. Consequently, for testing the null hypothesis of full rank against the alternative of rank zero, the $MB$ and $MJ$ tests are
preferably used.

Figure 3: Correct rank selection frequencies of the MMIB test when $N = 10$, $T = 100$, $p = 0$, $\theta = 0$ and $\Sigma = I_N$.

Note: $c$ and $\rho$ refer to the number of common trends and the identical autoregressive root of the $I(0)$ series, respectively.

We now consider the performance of sequential rank testing using the MMIB test statistics. The correct rank selection frequencies are depicted in Figures 3 ($T = 100$) and 4 ($T = 200$) with the procedure initiated at a first null hypothesis of full rank. Both figures display results for the intercept only case, $p = 0$, for $\theta = 0$, $\Sigma = I_N$ and $N = 10$. The figures display the results for varying true ranks $c$ ranging from zero to ten and different values of $\rho$ ranging from zero to 0.9. As expected, due to the sequential nature, the correct rank selection frequencies decrease with the true rank $c$ and this drop can be quite substantial. Consider as an example the case $T = 100$ and $\rho = 0.9$, in which case the correct rank selection frequency drops from about 95% to 0% when $c$ decreases from ten to eight. Clearly, a value of $\rho$ close to one makes it hard to distinguish $I(0)$ from $I(1)$ behavior. The problem is less pronounced
Figure 4: Correct rank selection frequencies of the MMIB test when $N = 10$, $T = 200$, $p = 0$, $\theta = 0$ and $\Sigma = I_N$.

Note: See Figure 3.

for larger $T$ (and for smaller values of $\rho$), for example with $T = 200$ and $\rho = 0$, the correct rank selection frequency never falls below 75% for any value of $\epsilon$.

4.3 Comparison to Some Existing Tests

The first test with which we compare our tests is the Ng (2008) $t$-test designed to infer the fraction of unit root time series in a panel without cross-unit cointegration, that is, in a panel where the number of common trends coincides with the number of $I(1)$ series. The Ng (2008) test is designed for factor model DGPs with only $I(0)$ factors, which excludes cross-unit cointegration.

The results for the case $\rho = 0$, $\theta = 0$ and $\Sigma = I_N$ are given in Table 5. The cross-sectional

\[^{25}\text{At the end of the following subsection we compare the MMIB rank selection performance with the performance of the Johansen (1995) trace test.}\]
dimensions are $N = 10, 20$ and we consider various (combinations of) null and alternative ranks. We present results for this simplest DGP only, since for more complicated DGPs, the relative performance of the Ng (2008) test compared to our tests is very poor. The results presented are thus the relatively most favorable for the Ng (2008) test.

Comparing first the size of the tests (upper block of Table 5), we see that, in contrast to $MB$ and $MJ$, the Ng (2008) $t$-test is severely oversized. This happens for a very simple DGP with homoskedastic and serially uncorrelated innovations and increasing $T$ only leads to marginal improvements. Power of the Ng (2008) test is quite low, especially when considering the fact that the reported powers are not size adjusted (and the test is oversized under the null). By contrast, the $MMIB$ test, which is now not used sequentially but is only evaluated once under the null rank, exhibits very high power, and this in spite of its conservativeness under the null.

We now turn to a comparison with two approaches that allow for cross-unit cointegration, the PANIC approach of Bai and Ng (2004), and the block-bootstrap tests of Palm et al. (2011). The DGP is modified for this comparison to fit the factor structure that these tests are designed for. In particular we consider

$$y_t = \lambda f_t \mathbf{1}_{N \times 1} + w_t,$$

with $f_t = f_{t-1} + \eta_t$ and $\eta_t \sim N(0, 1)$ independent of $w_t$. We consider two specifications for $w_t$, the AR(1) case with uncorrelated innovations ($\theta = 0$, $\Sigma = I_N$) and the ARMA(1,1) case with correlated innovations ($\theta = 0.8, -0.8$, $\Sigma$ generated as described in Section 4.1). The fact that $y_t$ has a factor structure means that the simulation design is again tailored to fit the tests used for comparison. The factor loading $\lambda$ is set equal to either zero or one.

PANIC requires the extraction of factors, which is done by principal components analysis with the number of factors determined using the $IC_{p2}$ information criterion of Bai and Ng (2002), where the maximum number of factors set to five. The cointegration properties of the extracted factors are analyzed using the $MQ_f$ and $MQ_c$ common trends tests of Bai and Ng (2004), with very poor results. First, the information criterion leads in a vast majority of replications to (the maximum of) five factors. Second, the factors are often found to be all $I(1)$ when using the common trends tests. An exception is the DGP in the third panel of Table 6, in which $MQ_f$ ($MQ_c$) correctly finds one $I(1)$ factor in about 40% (90%) of the

---

26The fact that the $MIB$ test is oversized when $c < N$ is clear, as it is divergent when the null rank is not full.
cases. Hence, with this approach all the series will be considered as I(1), since the series are deemed to contain five non-cointegrated I(1) factors. Of course, this is incorrect since the series either contain no common factor or one I(1) factor, depending upon the experiment.²⁷

Keeping the above results in mind, in what follows we consider the pooled panel unit root test \( P_\hat{e} \) of Bai and Ng (2004), applied to the estimated idiosyncratic components. This test is a combination of \( p \)-value-type unit root test in the spirit of Maddala and Wu (1999), and Choi (2001). The lag lengths used in the computations of the individual augmented Dickey–Fuller statistics are chosen using the Campbell-Perron sequential test rule based on the \( t \)-statistic of the highest lag, where the maximum number of lags is set to \( \lfloor 4(T/100)^{2/9} \rfloor \) following Ng and Perron (1995).

The other approach with which we compare our tests is due to Palm et al. (2011), which relies on bootstrap inference to deal with cross-sectional dependencies. These authors propose using block bootstrap versions of the pooled Levin et al. (2002), \( \tau_{p} \), and the group-mean \( \tau_{gm} \), panel unit root tests. It is the blocking of residuals over the whole cross-section in the resampling scheme that allows one to handle cross-sectional dependencies. The block length for the simulations is set equal to \( \lfloor 1.75T^{1/3} \rfloor \) (see Palm et al., 2011, Section 5) and the number of bootstrap replications is set to 499. Just as with our tests, the Palm et al. (2011) tests are applied to the original series and no decomposition into factors and idiosyncratic components is performed.

The size results are presented in Table 6 for four DGPs. In the top panel, we consider the no factor case with \( \lambda = 0 \), such that \( y_t = w_t \), where \( w_t \) follows an AR(1) process (\( \theta = 0 \)) with uncorrelated innovations (\( \Sigma = I_N \)). In the second panel, \( w_t \) is as before, but now \( \lambda = 1 \), so there is an I(1) factor present. In the bottom two panels, \( \lambda = 1 \) and \( w_t \) is an ARMA(1,1) process (\( \theta = 0.8, -0.8 \)) with correlated innovations (\( \Sigma \) is non-diagonal). For the first and second DGP our rank test have sizes close to the nominal level, whereas for the third DGP our tests are undersized. The \( P_\hat{e} \) test is oversized for these DGPs, and the two bootstrap tests are oversized for the first two DGPs and undersized for the third. In the latter case the bootstrap tests are less undersized than the rank tests. In the fourth DGP with \( \theta = -0.8 \) it is not surprising that all tests are substantially over-sized although the bootstrap tests have less severe distortions than the other tests.

²⁷Note for completeness that the correct number of common trends is selected with slightly higher probabilities in the \( p = 1 \) case compared to the \( p = 0 \) case. The results tend to improve for larger values of \( N \) and the cross-sectional dimensions considered here are very small for factor models.
Table 7 contains some power results when the tests are set up against the alternatives of rank equal to 0.5N or zero. The data are generated as an AR(1) process (θ = 0) without a common factor (λ = 0) and with uncorrelated innovations (Σ = IN). This is done for different values of of ρ and either 0.5N or no common trends. The results suggest that the MIB test is least powerful, and that the powers of MB and MJ are roughly comparable to those of P̂e, τp and τgm. These results illustrate the value of our tests, which achieve comparable power properties whilst being more accurate in terms of size.

Finally we compare the sequential rank testing procedure using the MMIB statistics with cointegration rank testing using the trace test of Johansen (1995), referred to as λtrace, which uses VAR models fitted to the data.28 We use the same DGPs as when generating Figures 3 and 4. The only difference is that we now also display results for N = 6. In general, and as expected, when N = 6, the Johansen approach dominates, while when N = 10 it is the other way round, especially for T = 100. Unreported simulations for larger values of N confirm the advantage of our MMIB sequential procedure for larger cross-sectional dimensions, that is, for medium sized time series panels. The MMIB test shows especially good comparative performance for large true ranks, that is, in situations “close” to the classical panel unit root null hypothesis.

5 Empirical Illustration

In this section we pursue a brief illustration taken from the empirical growth and convergence literature. Our analysis follows in large part the interpretation of Evans (1998) (see Banerjee and Wagner, 2009, for a more detailed discussion, including the role of deterministic components). Specifically, suppose that yit, log per capita GDP, which we refer to as income, in country i at time t, is I(1). Then the income panel is said to exhibit absolute convergence if, for any pair of countries i ≠ j, yit − yjt is I(0), and yit and yjt are thereby cointegrated with cointegrating vector [1, −1]′. Thus, by this definition, a necessary condition for absolute convergence is that the rank of the income panel must be one, since in this case there is only one common trend driving the whole panel.

Clearly, absolute convergence as defined is equivalent to having the panel of cross-section demeaned series ỹit = yit − \sum_{j=1}^{N} yjt/N form a panel with a rank of zero. Thus, one could

28The lag lengths for the VAR models are specified by minimizing the BIC with a maximum lag length equal to \lfloor 8(T/100)^{1/4} \rfloor.
consider testing the rank of the cross-sectionally demeaned data. Let $\Delta \tilde{u}_t$ denote the vector of stacked $\Delta \tilde{u}_{it} = \Delta u_{it} - \frac{\sum_{j=1}^{N} \Delta u_{jt}}{N}$. Clearly, the rank of $\Omega_{\Delta \tilde{u}_t \Delta \tilde{u}_t}$ is at most $N - 1$, even if the rank of $\Omega_{\Delta u \Delta u}$ is full. Similarly the rank of the appropriately scaled $\Sigma_{\Delta \tilde{u}_t \Delta \tilde{u}_t}$ is bounded by $N - 1$. This example shows that cross-sectional demeaning, often advocated in the literature as a simple device to remove cross-sectional dependencies, does in general affect the unit root and cointegration properties of panels of time series.\textsuperscript{29}

The given definition is the strictest definition of economic convergence, and it can be loosened in several ways, such as by not restricting the cointegrating vectors to be $[1, -1]'$ for all cointegrated pairs, or by allowing for several so called convergence clubs, that is, subsets of countries that are pair-wise cointegrated. A large rank of the panel, that is, a large number of common trends, implies that there cannot be a small number of convergence clubs. Thus, a finding of a large rank in a panel of log per-capita GDP data is evidence against convergence in the strict sense, and also against club convergence with a reasonable, small number of clubs.

Given the open-ended nature of growth theories, the researcher typically does not have much knowledge about the origins and nature of dependencies in income panels, which makes our tests well suited for this application since they do not place any restrictions on the dependencies. The data are taken from Maddison (2007), and comprise annual observations of log per capita GDP for 22 countries over the period 1870–2001.\textsuperscript{30} The results of performing our tests for both $p = 0$ and $p = 1$ are given in Table 9.\textsuperscript{31} The null hypothesis of full rank $c = 22$ is rejected by both the MIB and MB tests and not rejected by the MJ test for both specifications of the deterministic components. Performing the sequential MMIB testing sequence leads to a rank of 20. Thus, there is neither evidence for economic convergence nor evidence for a small number of convergence clubs in the Maddison data over the period 1870–2001.

\textsuperscript{29}Technically, for our tests it implies that the null limiting distributions as given are not well defined, due to the necessary singularity of the components of the test statistics.

\textsuperscript{30}The included countries are Australia, Austria, Belgium, Brazil, Canada, Denmark, Finland, France, Germany, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Spain, Sri Lanka, Sweden, United Kingdom, United States, Switzerland and Uruguay.

\textsuperscript{31}The inclusion of deterministic trends, driven by the observation of positive growth rates on average, leads to further implications with respect to economic convergence, as convergence then also requires, in the simplest case, equal trend slopes across countries. These and other aspects related to the presence of deterministic trends in convergence analysis are discussed in Banerjee and Wâgner (2009). Here we only focus on the rank aspect of the problem.
6 Conclusions

We have introduced new cointegration rank tests for time series panels. The tests have several important and partly distinctive features: First, conceptually, the extent of cross-sectional dependencies is not limited or restricted in shape, apart from some standard technical assumptions that lead to limiting distributions that are functionals of Wiener processes. Second, computationally, the tests do not require the estimation of any nuisance parameters and are free of choices concerning kernels and bandwidths, lag lengths or decisions concerning the number of factors to be extracted. Third, with our tests it is possible to flexibly test any null rank against any lower rank alternative. Fourth, the tests – with the details depending upon null and alternative as well as test considered – have good power properties, even when a linear time trend is included.

When testing the more specialized hypothesis of full rank, the generality of our tests does not come with a cost when compared to standard first generation panel unit root tests such as Im et al. (2003) and Levin and Lin (1992) even for the specialized DGPs with cross-sectional independence for which first generation tests were designed. Similarly, when testing for full rank our tests also perform well compared to state-of-the-art second generation panel unit root approaches, such as those of Bai and Ng (2004) or Palm et al. (2011) under the DGPs for which those tests were designed. In terms of power little or no price is paid for the good size accuracy, simplicity, general applicability and flexibility of the new tests. Finally, under more general DGPs, the sequential rank testing procedure based on our MMIB test performs well compared to the Johansen (1995) trace test.

Acknowledgements

We thank seminar participants at Brown University, Cornell University, Maastricht University, Texas AM, University of Montreal and Williams College, and conference participants at the Econometric Society World Congress, the Midwest Econometrics Group at the University of Chicago, the Unit Root and Cointegration Conference in Faro, the Conference on Factor Models in Panels at Goethe University Frankfurt, the 11th International Panel Data Conference, the NY Econometrics Camp, and in particular Robert Kunst, Peter Robinson and two anonymous referees for helpful comments and suggestions. All remaining errors and shortcomings are, of course, entirely ours. An earlier draft version of this paper circu-
lated with a similar title by the first two authors. Wagner is grateful for support by funds of the Oesterreichische Nationalbank (Anniversary Fund, project number 13398) and Westerlund thanks the Jan Wallander and Tom Hedelius Foundation for financial support under research grant P2009–0189:1.
References


Table 1: Critical values for the nonparametric tests.

<table>
<thead>
<tr>
<th>Value</th>
<th>MIB, MMIB</th>
<th>MB</th>
<th>MJ</th>
<th>MIB, MMIB</th>
<th>MB</th>
<th>MJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_1$</td>
<td>169037.15804</td>
<td>0.28636</td>
<td>45.12093</td>
<td>173133.91134</td>
<td>0.02454</td>
<td>4.86511</td>
</tr>
<tr>
<td>$\delta_2$</td>
<td>-375970.47567</td>
<td>-0.74096</td>
<td>-156.67769</td>
<td>-386036.11573</td>
<td>0.25342</td>
<td>-33.63947</td>
</tr>
<tr>
<td>$\delta_3$</td>
<td>236742.44174</td>
<td>1.30607</td>
<td>88.62252</td>
<td>242753.06689</td>
<td>-0.42490</td>
<td>18.71568</td>
</tr>
<tr>
<td>$\delta_4$</td>
<td>-28714.71985</td>
<td>-1.55572</td>
<td>-0.26967</td>
<td>-29227.10664</td>
<td>0.09250</td>
<td>1.66408</td>
</tr>
<tr>
<td>$\delta_5$</td>
<td>462.26011</td>
<td>1.08594</td>
<td>0.68685</td>
<td>499.44273</td>
<td>0.14925</td>
<td>0.23377</td>
</tr>
<tr>
<td>$\delta_6$</td>
<td>5.07971</td>
<td>-0.37162</td>
<td>-0.00337</td>
<td>4.95667</td>
<td>-0.09111</td>
<td>-0.00119</td>
</tr>
<tr>
<td>$\delta_7$</td>
<td>12978.51792</td>
<td>0.59838</td>
<td>-61.11665</td>
<td>11680.94595</td>
<td>0.88073</td>
<td>-23.24877</td>
</tr>
<tr>
<td>$\delta_8$</td>
<td>-800.55043</td>
<td>-0.63492</td>
<td>-0.26967</td>
<td>-29227.10664</td>
<td>0.09250</td>
<td>1.66408</td>
</tr>
<tr>
<td>$\delta_9$</td>
<td>-531706.18433</td>
<td>0.06527</td>
<td>3193.19031</td>
<td>-435669.56662</td>
<td>0.04085</td>
<td>1199.08007</td>
</tr>
<tr>
<td>$\delta_{10}$</td>
<td>34932714.04854</td>
<td>4.42248</td>
<td>-23173.56365</td>
<td>33299454.36837</td>
<td>5.13046</td>
<td>-10069.57330</td>
</tr>
<tr>
<td>$\delta_{11}$</td>
<td>-7336758.89483</td>
<td>-34.18132</td>
<td>-25072.41143</td>
<td>-7720478.56401</td>
<td>-35.78908</td>
<td>-9009.64022</td>
</tr>
<tr>
<td>$\delta_{12}$</td>
<td>286553.86002</td>
<td>29.48598</td>
<td>3577.85398</td>
<td>334056.64067</td>
<td>32.62937</td>
<td>1349.07676</td>
</tr>
</tbody>
</table>

Specific critical values for null ranks $c \leq 5$

<table>
<thead>
<tr>
<th>Value</th>
<th>MIB, MMIB</th>
<th>MB</th>
<th>MJ</th>
<th>MIB, MMIB</th>
<th>MB</th>
<th>MJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c = 5$</td>
<td>2261.66016</td>
<td>0.10545</td>
<td>32.75513</td>
<td>3397.36193</td>
<td>0.02813</td>
<td>13.62580</td>
</tr>
<tr>
<td>$c = 4$</td>
<td>1393.16502</td>
<td>0.08999</td>
<td>22.98579</td>
<td>2235.34896</td>
<td>0.02254</td>
<td>9.63480</td>
</tr>
<tr>
<td>$c = 3$</td>
<td>751.72488</td>
<td>0.06450</td>
<td>13.96212</td>
<td>1329.77879</td>
<td>0.01639</td>
<td>6.16500</td>
</tr>
<tr>
<td>$c = 2$</td>
<td>323.89602</td>
<td>0.03505</td>
<td>7.01714</td>
<td>697.89714</td>
<td>0.00994</td>
<td>3.17750</td>
</tr>
<tr>
<td>$c = 1$</td>
<td>101.56031</td>
<td>0.00985</td>
<td>1.90737</td>
<td>289.02594</td>
<td>0.00346</td>
<td>0.91449</td>
</tr>
</tbody>
</table>

Notes: The regression model is given by $\tau = \delta^T Z + \epsilon$, where $\tau$ is the 5% critical value and $Z$ is given in Section 3.2. The estimated 5% critical value is computed as the fitted value of the regression. $p = 0$ and $p = 1$ refer to the model with constant, and constant and linear trend, respectively.
Table 2: Size at the 5% level when $\rho = 0.1$.

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>$T$</th>
<th>$N$</th>
<th>$p = 0$</th>
<th>$p = 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>MIB</td>
<td>MMIB</td>
</tr>
<tr>
<td>$c = N$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>10</td>
<td>100</td>
<td>3.3</td>
<td>5.7</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>100</td>
<td>4.9</td>
<td>6.6</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>200</td>
<td>2.5</td>
<td>5.1</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>200</td>
<td>7.4</td>
<td>5.5</td>
</tr>
<tr>
<td>0.8</td>
<td>10</td>
<td>100</td>
<td>0.1</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>100</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>200</td>
<td>0.3</td>
<td>2.6</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>200</td>
<td>0.0</td>
<td>0.3</td>
</tr>
<tr>
<td>$-0.8$</td>
<td>10</td>
<td>100</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>100</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>200</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>200</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>$c = 0.5N$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>10</td>
<td>100</td>
<td>100.0</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>100</td>
<td>100.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>200</td>
<td>100.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>200</td>
<td>100.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.8</td>
<td>10</td>
<td>100</td>
<td>100.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>100</td>
<td>100.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>200</td>
<td>100.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>200</td>
<td>100.0</td>
<td>0.0</td>
</tr>
<tr>
<td>$-0.8$</td>
<td>10</td>
<td>100</td>
<td>100.0</td>
<td>99.6</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>100</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>200</td>
<td>100.0</td>
<td>85.5</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>200</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Notes: $\rho$ and $\theta$ refer to the autoregressive and moving average coefficients in the definition of $w_t$. Since under the null of full rank, $MIB$ and $MMIB$ are identical, in the upper panel we only consider the former statistic.
Table 3: Power at the 5% level when testing $c = N$ versus $c_1 < N$ for varying values of $c_1$.

<table>
<thead>
<tr>
<th>$c_1$</th>
<th>$T$</th>
<th>$N$</th>
<th>$p = 0$</th>
<th>$p = 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>MIB</td>
<td>MB</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: $c_1$ refers to the true number of common trends under the alternative, and $\rho$ refers to the autoregressive coefficient of the $I(0)$ units. See Table 2 for further explanations.
Table 4: Power at the 5% level when testing $c = N$ versus $c_1 < N$ for varying values of $\rho$.

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$T$</th>
<th>$N$</th>
<th>$p = 0$</th>
<th>$p = 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$MIB$</td>
<td>$MB$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$c_1 = 0.8N$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.9</td>
<td>100</td>
<td>10</td>
<td>3.9</td>
<td>12.1</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>10</td>
<td>4.7</td>
<td>16.0</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>20</td>
<td>9.8</td>
<td>21.6</td>
</tr>
<tr>
<td>0.7</td>
<td>100</td>
<td>10</td>
<td>14.9</td>
<td>23.4</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>20</td>
<td>13.0</td>
<td>39.6</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>10</td>
<td>53.0</td>
<td>29.9</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>20</td>
<td>60.3</td>
<td>56.8</td>
</tr>
<tr>
<td>0.5</td>
<td>100</td>
<td>10</td>
<td>53.1</td>
<td>30.9</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>20</td>
<td>44.9</td>
<td>59.4</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>10</td>
<td>97.4</td>
<td>35.4</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>20</td>
<td>99.7</td>
<td>74.1</td>
</tr>
<tr>
<td></td>
<td>$c_1 = 0$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.99</td>
<td>100</td>
<td>10</td>
<td>4.1</td>
<td>13.6</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>20</td>
<td>5.3</td>
<td>16.2</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>10</td>
<td>2.4</td>
<td>26.2</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>20</td>
<td>7.8</td>
<td>34.4</td>
</tr>
<tr>
<td>0.95</td>
<td>100</td>
<td>10</td>
<td>4.6</td>
<td>83.6</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>20</td>
<td>5.8</td>
<td>96.2</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>10</td>
<td>5.9</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>20</td>
<td>12.0</td>
<td>100.0</td>
</tr>
<tr>
<td>0.9</td>
<td>100</td>
<td>10</td>
<td>8.5</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>20</td>
<td>8.4</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>10</td>
<td>26.5</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>20</td>
<td>30.6</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Notes: See Tables 2 and 3 for further explanations.
Table 5: Size and power comparison with the Ng (2008) test at the 5% level for $p = 0$, $\rho = 0$, $\theta = 0$ and $\Sigma = I_N$.

<table>
<thead>
<tr>
<th>$N$</th>
<th>$c$</th>
<th>$c_1$</th>
<th>MIB</th>
<th>MMIB</th>
<th>MB</th>
<th>MJ</th>
<th>$t$</th>
<th>$T = 100$</th>
<th>MIB</th>
<th>MMIB</th>
<th>MB</th>
<th>MJ</th>
<th>$t$</th>
<th>$T = 200$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>10</td>
<td>10</td>
<td>3.8</td>
<td>3.8</td>
<td>5.8</td>
<td>5.2</td>
<td>20.1</td>
<td>2.6</td>
<td>2.6</td>
<td>5.6</td>
<td>6.0</td>
<td>12.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>20</td>
<td>20</td>
<td>4.7</td>
<td>4.7</td>
<td>6.5</td>
<td>4.1</td>
<td>18.8</td>
<td>7.8</td>
<td>7.8</td>
<td>5.3</td>
<td>5.1</td>
<td>12.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>7</td>
<td>7</td>
<td>100.0</td>
<td>0.4</td>
<td>4.4</td>
<td>5.1</td>
<td>17.4</td>
<td>100.0</td>
<td>0.1</td>
<td>4.0</td>
<td>5.6</td>
<td>10.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>14</td>
<td>14</td>
<td>100.0</td>
<td>0.0</td>
<td>1.8</td>
<td>2.0</td>
<td>18.1</td>
<td>100.0</td>
<td>1.5</td>
<td>3.4</td>
<td>4.6</td>
<td>12.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>4</td>
<td>4</td>
<td>100.0</td>
<td>16.1</td>
<td>4.1</td>
<td>6.8</td>
<td>14.2</td>
<td>100.0</td>
<td>0.2</td>
<td>4.4</td>
<td>2.7</td>
<td>5.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>8</td>
<td>8</td>
<td>100.0</td>
<td>0.0</td>
<td>1.1</td>
<td>1.8</td>
<td>16.5</td>
<td>100.0</td>
<td>0.0</td>
<td>2.3</td>
<td>3.9</td>
<td>9.8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$c_1 = c$ (size)

$c_1 < c$ (power)

Notes: For the size analysis (upper panel) $c = c_1$ denotes the true rank and for the power analysis (lower panel) the hypothesized rank is $c$ and the true rank is given by $c_1$. $t$ refers to the Ng (2008) $t$-test, and $\rho$, $\theta$ and $\Sigma$ refer to the autoregressive and moving average coefficients and the error covariance matrix.
Table 6: Size comparison with the Bai and Ng (2004) and Palm et al. (2011) tests at the 5% level for $H_0 : c = N$.

<table>
<thead>
<tr>
<th>$N$</th>
<th>$T$</th>
<th>$p = 0$</th>
<th>$p = 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$MIB$</td>
<td>$MB$</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>2.8</td>
<td>5.1</td>
</tr>
<tr>
<td>20</td>
<td>100</td>
<td>5.1</td>
<td>6.8</td>
</tr>
<tr>
<td>10</td>
<td>200</td>
<td>2.6</td>
<td>5.4</td>
</tr>
<tr>
<td>20</td>
<td>200</td>
<td>7.5</td>
<td>5.3</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>3.3</td>
<td>5.7</td>
</tr>
<tr>
<td>20</td>
<td>100</td>
<td>4.7</td>
<td>6.6</td>
</tr>
<tr>
<td>10</td>
<td>200</td>
<td>2.2</td>
<td>4.9</td>
</tr>
<tr>
<td>20</td>
<td>200</td>
<td>7.0</td>
<td>5.0</td>
</tr>
</tbody>
</table>

$\lambda = \theta = 0, \Sigma = I_N$

$\lambda = 1, \theta = 0, \Sigma = I_N$

$\lambda = 1, \theta = 0.8, \Sigma $ nondiagonal

$\lambda = 1, \theta = -0.8, \Sigma $ nondiagonal

Notes: $P_e$ refers to the idiosyncratic panel unit root test of Bai and Ng (2004), while $\tau_p$ and $\tau_{gm}$ refer to the bootstrap tests of Palm et al. (2011). $\lambda$ refers to the factor loading and $\theta$ is the moving average coefficient.
Table 7: Power comparison with the Bai and Ng (2004) and Palm et al. (2011) tests at the 5% level for $H_0 : c = N$.

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$N$</th>
<th>$T$</th>
<th>$MIB$</th>
<th>$MB$</th>
<th>$MJ$</th>
<th>$P_{\tau}$</th>
<th>$\tau_{gm}$</th>
<th>$MIB$</th>
<th>$MB$</th>
<th>$MJ$</th>
<th>$P_{\tau}$</th>
<th>$\tau_{gm}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$c_1 = 0.5N$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$c_1 = 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.9</td>
<td>10</td>
<td>100</td>
<td>5.5</td>
<td>49.3</td>
<td>46.4</td>
<td>66.8</td>
<td>54.1</td>
<td>4.9</td>
<td>29.3</td>
<td>33.1</td>
<td>52.8</td>
<td>44.7</td>
</tr>
<tr>
<td>20</td>
<td>100</td>
<td>5.8</td>
<td>65.9</td>
<td>58.7</td>
<td>93.6</td>
<td>72.4</td>
<td>88.8</td>
<td>6.7</td>
<td>44.6</td>
<td>44.2</td>
<td>81.2</td>
<td>62.2</td>
</tr>
<tr>
<td>10</td>
<td>200</td>
<td>11.9</td>
<td>73.8</td>
<td>62.5</td>
<td>96.6</td>
<td>76.0</td>
<td>99.5</td>
<td>10.5</td>
<td>57.2</td>
<td>61.3</td>
<td>93.9</td>
<td>80.4</td>
</tr>
<tr>
<td>20</td>
<td>200</td>
<td>17.5</td>
<td>93.2</td>
<td>76.0</td>
<td>100.0</td>
<td>92.4</td>
<td>100.0</td>
<td>17.2</td>
<td>85.1</td>
<td>79.0</td>
<td>99.9</td>
<td>95.2</td>
</tr>
<tr>
<td>0.7</td>
<td>10</td>
<td>100</td>
<td>50.7</td>
<td>89.9</td>
<td>62.4</td>
<td>99.2</td>
<td>89.0</td>
<td>44.8</td>
<td>83.4</td>
<td>68.4</td>
<td>99.9</td>
<td>92.1</td>
</tr>
<tr>
<td>20</td>
<td>100</td>
<td>38.3</td>
<td>99.5</td>
<td>77.2</td>
<td>100.0</td>
<td>97.5</td>
<td>100.0</td>
<td>37.2</td>
<td>98.7</td>
<td>82.6</td>
<td>100.0</td>
<td>99.0</td>
</tr>
<tr>
<td>10</td>
<td>200</td>
<td>98.9</td>
<td>96.9</td>
<td>71.2</td>
<td>99.9</td>
<td>91.7</td>
<td>100.0</td>
<td>98.3</td>
<td>95.6</td>
<td>79.6</td>
<td>100.0</td>
<td>96.5</td>
</tr>
<tr>
<td>20</td>
<td>200</td>
<td>99.4</td>
<td>100.0</td>
<td>86.2</td>
<td>100.0</td>
<td>98.8</td>
<td>100.0</td>
<td>99.0</td>
<td>100.0</td>
<td>93.1</td>
<td>100.0</td>
<td>99.9</td>
</tr>
<tr>
<td>0.95</td>
<td>10</td>
<td>100</td>
<td>3.9</td>
<td>83.0</td>
<td>93.5</td>
<td>68.0</td>
<td>58.7</td>
<td>4.3</td>
<td>28.3</td>
<td>37.1</td>
<td>42.2</td>
<td>35.6</td>
</tr>
<tr>
<td>20</td>
<td>100</td>
<td>5.4</td>
<td>95.4</td>
<td>99.4</td>
<td>91.4</td>
<td>80.2</td>
<td>71.2</td>
<td>5.6</td>
<td>42.5</td>
<td>51.5</td>
<td>64.0</td>
<td>50.7</td>
</tr>
<tr>
<td>10</td>
<td>200</td>
<td>6.3</td>
<td>99.9</td>
<td>100.0</td>
<td>99.5</td>
<td>99.4</td>
<td>97.6</td>
<td>5.6</td>
<td>83.2</td>
<td>96.4</td>
<td>88.1</td>
<td>93.8</td>
</tr>
<tr>
<td>20</td>
<td>200</td>
<td>11.3</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>99.9</td>
<td>11.1</td>
<td>98.0</td>
<td>99.9</td>
<td>99.9</td>
<td>99.8</td>
</tr>
<tr>
<td>0.9</td>
<td>10</td>
<td>100</td>
<td>7.7</td>
<td>100.0</td>
<td>100.0</td>
<td>99.8</td>
<td>99.4</td>
<td>7.6</td>
<td>87.0</td>
<td>95.4</td>
<td>91.4</td>
<td>91.9</td>
</tr>
<tr>
<td>20</td>
<td>100</td>
<td>8.1</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>7.8</td>
<td>98.4</td>
<td>99.9</td>
<td>99.7</td>
<td>99.5</td>
</tr>
<tr>
<td>10</td>
<td>200</td>
<td>27.3</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>25.6</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>20</td>
<td>200</td>
<td>30.1</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>27.4</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Notes: The data are generated without a common factor ($\lambda = 0$), $\rho = 0$, $\theta = 0$ and $\Sigma = 1_N$ and $c_1$ is the true number of common trends. See Table 6 for further explanations.
Table 8: Correct rank selection frequencies of the MMIB and Johansen (1995) tests for $p = 0$.

<table>
<thead>
<tr>
<th>$N$</th>
<th>$c$</th>
<th>$\rho = 0.1$</th>
<th>$\rho = 0.5$</th>
<th>$\rho = 0.9$</th>
<th>$\lambda_{\text{trace}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\rho = 0.1$</td>
<td>$\rho = 0.5$</td>
<td>$\rho = 0.9$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\rho = 0.1$</td>
<td>$\rho = 0.5$</td>
<td>$\rho = 0.9$</td>
<td></td>
</tr>
<tr>
<td>T = 100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 0</td>
<td>83.7</td>
<td>83.7</td>
<td>83.7</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>2 1</td>
<td>70.1</td>
<td>10.8</td>
<td>0.0</td>
<td>94.5</td>
<td>78.0</td>
</tr>
<tr>
<td>4 1</td>
<td>84.9</td>
<td>23.8</td>
<td>0.2</td>
<td>93.4</td>
<td>71.1</td>
</tr>
<tr>
<td>6 1</td>
<td>98.1</td>
<td>98.1</td>
<td>98.1</td>
<td>90.7</td>
<td>90.7</td>
</tr>
<tr>
<td>10 1</td>
<td>43.6</td>
<td>43.6</td>
<td>43.6</td>
<td>32.9</td>
<td>32.9</td>
</tr>
<tr>
<td>2 1</td>
<td>26.7</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>4 1</td>
<td>29.3</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>6 1</td>
<td>26.4</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>8 1</td>
<td>56.3</td>
<td>3.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>10 1</td>
<td>96.5</td>
<td>96.5</td>
<td>96.5</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>T = 200</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 0</td>
<td>99.7</td>
<td>99.7</td>
<td>99.7</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>2 1</td>
<td>92.5</td>
<td>58.2</td>
<td>0.0</td>
<td>94.3</td>
<td>94.3</td>
</tr>
<tr>
<td>4 1</td>
<td>95.9</td>
<td>75.8</td>
<td>0.4</td>
<td>93.9</td>
<td>93.6</td>
</tr>
<tr>
<td>6 1</td>
<td>99.8</td>
<td>99.8</td>
<td>99.8</td>
<td>92.1</td>
<td>92.1</td>
</tr>
<tr>
<td>10 1</td>
<td>96.3</td>
<td>96.3</td>
<td>96.3</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>2 1</td>
<td>80.5</td>
<td>15.6</td>
<td>0.0</td>
<td>94.3</td>
<td>94.1</td>
</tr>
<tr>
<td>4 1</td>
<td>89.0</td>
<td>10.3</td>
<td>0.0</td>
<td>93.3</td>
<td>93.1</td>
</tr>
<tr>
<td>6 1</td>
<td>74.6</td>
<td>2.1</td>
<td>0.0</td>
<td>92.3</td>
<td>85.2</td>
</tr>
<tr>
<td>8 1</td>
<td>95.2</td>
<td>30.0</td>
<td>0.1</td>
<td>89.9</td>
<td>85.3</td>
</tr>
<tr>
<td>10 1</td>
<td>97.3</td>
<td>97.3</td>
<td>97.3</td>
<td>86.5</td>
<td>86.5</td>
</tr>
</tbody>
</table>

Notes: $\lambda_{\text{trace}}$ refers to the trace test of Johansen (1995). See Figure 3 for further explanations.

Table 9: Empirical results for the Maddison data.

<table>
<thead>
<tr>
<th>$p = 0$</th>
<th>$p = 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$MIB$</td>
<td>$MB$</td>
</tr>
<tr>
<td>Test results</td>
<td>100</td>
</tr>
<tr>
<td>5% critical values</td>
<td>90986.5</td>
</tr>
</tbody>
</table>

Notes: Asterisks indicate rejection of the null hypothesis at the 5% level. The critical values based on the response surface regressions are given in the second row.