Asymptotically UMP Panel Unit Root Tests

Becheri, I.G.; Drost, F.C.; van den Akker, R.

Publication date:
2013

Citation for published version (APA):
Becheri, I. G., Drost, F. C., & van den Akker, R. (2013). Asymptotically UMP Panel Unit Root Tests. (CentER Discussion Paper; Vol. 2013-017). Tilburg: Econometrics.

CentER

Discussion Paper

No. 2013-017

ASYMPTOTICALLY UMP PANEL UNIT ROOT TEST - THE EFFECT OF HETEROGENEITY IN THE ALTERNATIVES -

By

I. Gaia Becheri, Feike C. Drost, Ramon van den Akker

March 18, 2013

ISSN 0924-7815
ISSN 2213-9532

Tilburg University

Asymptotically UMP Panel Unit Root Tests
- the effect of heterogeneity in the alternatives -

I. Gaia Becheri[a], Feike C. Drost[b,c], Ramon van den Akker[b,*]

[a] *Delft Institute of Applied Mathematics (DIAM), Delft University of Technology*
[b] *Econometrics Group, CentER, Tilburg University*
[c] *Finance Group, CentER, Tilburg University*

Abstract

This paper considers optimal unit root tests for a Gaussian cross-sectionally independent heterogeneous panel with incidental intercepts and heterogeneous alternatives generated by random perturbations. We derive the (asymptotic and local) power envelope for two models: an auxiliary model where both the panel units and the random perturbations are observed, and the second one, the model of main interest, for which only the panel units are observed. We show that both models are Locally Asymptotically Normal (LAN). It turns out that there is an information loss: the power envelope for the auxiliary model is above the envelope for the model of main interest. Equality only holds if the alternatives are homogeneous. Our results exactly identify in which setting the unit root test of Moon, Perron, and Phillips (2007) is asymptotically UMP and, in fact, they show it is not possible to exploit possible heterogeneity in the alternatives, confirming a conjecture of Breitung and Pesaran (2008). Moreover, we propose a new asymptotically optimal test and we extend the results to a model allowing for cross-sectional dependence.

JEL classification: C22; C23

Keywords: panel unit root test, Local Asymptotic Normality, limit experiment, asymptotic power envelope, information loss

*Corresponding author
Email addresses: I.G.Becheri@tudelft.nl (I. Gaia Becheri), F.C.Drost@TilburgUniversity.edu (Feike C. Drost), R.vdnAkker@TilburgUniversity.edu (Ramon van den Akker)

March 8, 2013

1. Introduction

Just as for univariate time series, the presence or absence of unit roots in a panel data model can have crucial economic policy implications. Therefore, in the last two decades, a lot of attention has been given to testing for unit roots in panel data. We refer to Banerjee (1999), Baltagi and Kao (2000), Choi (2006), Breitung and Pesaran (2008), and Westerlund and Breitung (2013) for surveys.

It is, of course, important to know what the (asymptotic) power of panel unit root tests is. For cross-sectionally independent panels, local asymptotic powers of tests have been considered in, e.g., Breitung (2000), Moon and Perron (2008), Harris, Harvey, Leybourne, and Sakkas (2010), and Madsen (2010); (asymptotic) optimality for Gaussian panels has been studied in Moon, Perron, and Phillips (2007). For models with cross-sectional dependence local asymptotic powers have been derived in, e.g., Moon and Perron (2004) and Breitung and Das (2005).

The main contributions of this paper are threefold. First, we reconsider the Gaussian heterogeneous cross-sectionally independent panel from Moon, Perron, and Phillips (2007) allowing for incidental intercepts. In this setup, alternatives to the unit root are possibly heterogeneous and generated by random perturbations to the unit root null hypothesis. We derive the limit experiment[1] à la Le Cam of two models: an auxiliary model where both the panel units and the random perturbations are observed and a second one, the model of main interest, for which only the panel units are observed. To determine the limit experiment we need to derive an asymptotic expansion of likelihood ratios. For the auxiliary model the likelihood ratio and the corresponding expansion can be computed explicitly. However, for the model of interest this is not feasible. We derive a general result, of independent interest, inspired by the information loss result of Le Cam and Yang (1988) for i.i.d. models. This information loss idea relates the expansion of the likelihood ratio for the model of interest to the likelihood ratio of the auxiliary model. We exploit this information loss result to show that both the auxiliary model and the model of interest are Locally Asymptotically Normal (LAN). From this LAN structure we obtain attainable (asymptotic) power envelopes for both models. It turns

[1] See, e.g., Le Cam (1986), Le Cam and Yang (1990), Van der Vaart (1991), or Van der Vaart (2000).

out that there is a "loss of information", i.e. the power envelope for the model of interest is below the power envelope for the auxiliary model and equality holds if and only if the alternatives are homogeneous. If, in a model with heterogeneous alternatives one only observes the panel units, one ends up with the same power envelope as for a model with homogeneous alternatives.

Second, we use our asymptotic power envelopes to clarify the role of the heterogeneous alternatives in Moon, Perron, and Phillips (2007). That paper derived an upper bound to the asymptotic (and local) power of unit root tests and proposed a point-optimal test. This test depends on a sequence of random variables that has to be specified by the researcher (using the test) and can be seen as a likelihood ratio test, where the alternative is determined by the specified random variables. In case these random variables are the same as the random perturbations generating the alternatives to the unit root, the asymptotic power of this test attains the upper bound and the test will be (asymptotically) optimal. We show, in case one does not observe the random perturbations, the upper bound of Moon, Perron, and Phillips (2007) is not attainable and that one, in fact, ends up with the same power envelope as for homogeneous alternatives. This result confirms a conjecture by Breitung and Pesaran (2008) stating it is not possible to exploit the possible heterogeneity of the alternative in case one only observes the panel units. We show that choosing constants in the test of Moon, Perron, and Phillips (2007) yields an asymptotically UMP test and we also introduce a new optimal test.

Third, we consider an extension to a Gaussian panel with cross-sectional dependence. By using a rotation argument we show that the asymptotic power envelope is not affected by the cross-sectional dependence. In case one considers cross-sectional dependence generated by a factor structure (without serial dependence), it turns out that the t-tests of Moon and Perron (2004), based on a modified pooled OLS estimator, are asymptotically UMP in case the idiosyncratic shocks in the model have homogeneous variances.

For the unit root problem in (univariate) time series limit experiment theory has been exploited by Jansson (2008) and Hallin, Van den Akker, and Werker (2011) (for the continuous-valued case) and Drost, Van den Akker, and Werker (2009) (for the integer-valued case). To our best knowledge, this paper is the first to consider the limit experiment for panel data models with unit roots. We stress that the limit experiment approach is not only useful to study optimal inference. Using "Le Cam calculus", in particular Le Cam's third lemma (see, e.g., Van der Vaart (2000, p.90)), the local and asymptotic power of any test having an asymptotically linear expansion under the null hypothesis can be obtained. In this way one can avoid the use of "triangular array arguments" to calculate the local asymptotic power of a test.

The paper is organized as follows. Section 2 describes the model and setup, Section 3 contains the information loss result, Section 4 contains the main results, and Appendix A contains the proofs.

2. Setup and assumptions

We assume that the observations Y_{it}, $i = 1, \ldots, n$ and $t = 1, \ldots, T$, are generated by the component model,

$$Y_{it} = m_i + Y_{it}^0, \tag{1}$$

$$Y_{it}^0 = \rho_i Y_{i,t-1}^0 + \sigma_i \varepsilon_{it}, \tag{2}$$

where m_i is a (deterministic) intercept, i.e. fixed effect, and $Y_{i0}^0 = 0$. The autoregression coefficients ρ_i are assumed to be generated by the random coefficient structure,

$$\rho_i = 1 + \frac{h}{T\sqrt{n}} H_i, \tag{3}$$

where H_i is random with mean 1 and the (deterministic) parameter h describes the (mean) deviation from the unit root. Assumption 1 describes our precise conditions on the perturbations H_i and the idiosyncratic shocks ε_{it}.

Assumption 1

(a) The innovations ε_{it}, $i, t \in \mathbb{N}$, are i.i.d. $\mathcal{N}(0, 1)$.

(b) The perturbations H_i, $i \in \mathbb{N}$, are i.i.d. with mean 1 and independent of the idiosyncratic shocks ε_{it}, $i, t \in \mathbb{N}$. Moreover, the moment generating function of H_1 exists on an open interval containing 0.

(c) The deterministic scale parameters σ_i are strictly positive, i.e. $\sigma_i > 0$ for $i \in \mathbb{N}$.

We are interested in testing the unit root hypothesis[2],

$$\mathrm{H}_0 : h = 0 \text{ versus } \mathrm{H}_a : h < 0. \tag{4}$$

[2]Our results can also be used to derive the (asymptotic) power envelope for the two-sided alternative $\mathrm{H}_a : h \neq 0$ or to test a unit root versus an explosive alternative, i.e. $\mathrm{H}_a : h > 0$.

Under the null hypothesis each individual time series $\{Y_{it},\, t \in \mathbb{N}\}$ contains a unit root. Under the alternative, heterogeneity in the ρ_i is introduced via the perturbations H_i. Assumption 1 allows for $P(H_1 = 0) > 0$, which means that a (random) fraction of the time series $\{Y_{it},\, t \in \mathbb{N}\}$ still contains a unit root under the alternative. By considering $P(H_1 = 1) = 1$ we can also cover homogenous alternatives, i.e. $\rho_i = \rho < 1$ for all i. Finally, note that even $P(H_1 < 0) > 0$ is allowed for, yielding a (random) fraction of explosive time series $\{Y_{it},\, t \in \mathbb{N}\}$. This last setting, however, seems to be less relevant from an empirical point of view.

The normalization by $\sqrt{n}T$ in the specification for ρ_i is motivated by the asymptotic structure of the model as this rate of localization yields contiguous alternatives to the null hypothesis. Throughout we use joint limit theory (i.e. $T, n \to \infty$ jointly); see Phillips and Moon (1999) for precise definitions and a detailed exposition.

3. An information loss result

In Section 4 we will derive two limit experiments associated to the panel unit root model. The first limit experiment corresponds to an auxiliary model in which both Y_{it} and H_i are observed. And the second limit experiment corresponds to the model of interest in which only Y_{it} is observed. To determine a limit experiment we need to derive an asymptotic expansion of likelihood ratios. For the auxiliary model the likelihood ratio and the corresponding expansion can be computed explicitly. However, for the model of main interest this is not possible. The following lemma, inspired by general results by Le Cam and Yang (1988) on information loss in i.i.d. models, will allow us to derive the limit experiment for the model of main interest from the limit experiment of the auxiliary model.

Lemma 3.1 *Let $(\mathcal{G}_n)_{n\in\mathbb{N}}$ a sequence of σ-fields and let, for all $n \in \mathbb{N}$, $(\mathcal{X}_{ni})_{1\le i\le n}$ and $(\mathcal{J}_{ni})_{1\le i\le n}$ be $\mathcal{G}_n$-measurable random variables. And let $(U_i)_{i\in\mathbb{N}}$ be i.i.d. random variables independent of $\mathcal{G}_n$ for all n. Suppose*

i) the moment generating function of U_1 exists;

ii)

$$\frac{1}{n}\sum_{i=1}^{n} \mathcal{X}_{ni}^2 = O_P(1) \text{ and } \max_{i=1,\dots,n} \frac{|\mathcal{X}_{ni}|}{\sqrt{n}} = o_P(1); \tag{5}$$

iii) for all $n \in \mathbb{N}$ *and* $i = 1, \ldots, n$, $\mathcal{J}_{ni} \geq 0$,

$$\frac{1}{n}\sum_{i=1}^{n} \mathcal{J}_{ni} - \frac{1}{n}\sum_{i=1}^{n} \mathcal{X}_{ni}^2 = o_P(1) \text{ and } \max_{i=1,\ldots,n} \frac{\mathcal{J}_{ni}}{n} = o_P(1). \tag{6}$$

Then, defining

$$L_n^{\mathcal{X},U} = \exp\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} U_i \mathcal{X}_{ni} - \frac{1}{2n}\sum_{i=1}^{n} U_i^2 \mathcal{J}_{ni}\right),$$

and

$$L_n^{\mathcal{X}} = \mathbb{E}\left[L_n^{\mathcal{X},U} \mid \mathcal{G}_n\right],$$

we have

$$L_n^{\mathcal{X}} = m_n \exp\left((\mathbb{E}U_1)\frac{1}{\sqrt{n}}\sum_{i=1}^{n} \mathcal{X}_{ni} - (\mathbb{E}U_1)^2 \frac{1}{2n}\sum_{i=1}^{n} \mathcal{J}_{ni}\right),$$

with $m_n \xrightarrow{p} 1$.

See Appendix A.1 for the proof. We note that the result is of independent interest and could, for example, be exploited in other (panel) unit root models (with alternatives generated by random unobservable perturbations).

In our applications of the lemma we relate $L_n^{\mathcal{X},U}$ to a likelihood ratio for the auxiliary model and $L_n^{\mathcal{X}}$ to a likelihood ratio for the model of interest (with U_i corresponding to H_i and $\mathcal{X}_{ni}$ and $\mathcal{J}_{ni}$ corresponding to certain statistics of Y_{it}). As the perturbations H_i are not observed in the model of interest, there is an information loss: the model of interest does not contain as much "statistical information" as the auxiliary model. Intuitively this follows by comparing the probability limits of $n^{-1}\sum_{i=1}^{n} U_i^2 \mathcal{J}_{ni}$ and $(\mathbb{E}U_1)^2 n^{-1}\sum_{i=1}^{n} \mathcal{J}_{ni}$, which can be interpreted as the respective Fisher-informations in the two models. Since $\mathbb{E}U_1^2 > (\mathbb{E}U_1)^2$, if U is not degenerated, we see that there is indeed an information loss. This heuristic result will be formalized in Section 4 where we show that both the auxiliary and the model of interest are Locally Asymptotically Normal (LAN).

4. Main results

This section presents our main results. Subsection 4.1 contains the results on the limit experiments, subsection 4.2 discusses (asymptotically) optimal tests for the model (1)-(3), and subsection 4.3 considers an extension to a model allowing for cross-sectional dependence. Throughout we use the 'Δ-operator' to denote differencing across the time dimension, i.e. $\Delta Y_{it} = Y_{it} - Y_{i,t-1}$.

4.1. Limit experiments and Power envelopes

In this section we assume that the parameters m_i and σ_i are known. The main objective is to determine the limit experiment of the model corresponding to observing Y_{it}, $i = 1, \ldots, n$ and $t = 1, \ldots, T$, where h is the parameter of interest. To this end we first derive the limit experiment for the auxiliary model corresponding to observing both Y_{it} and H_i.

Denote the joint law of Y_{it} and H_i, $i = 1, \ldots, n$ and $t = 1, \ldots, T$, by $\tilde{P}_h^{(n,T)}$. The following proposition shows that the corresponding model is of the Locally Asymptotical Normal (LAN) type. To formulate the proposition we first introduce the partial sum processes[3],

$$W_i^{(T)}(u) := \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tu]} \frac{1}{\sigma_i} \Delta Y_{it}, \quad u \in [0, 1].$$

Note that these partial sum processes are measurable with respect to the observations Y_{it}, $i = 1, \ldots, n$ and $t = 1, \ldots, T$ (recall that $Y_{i0} = m_i$ is known in this section). Moreover, as already suggested by the notation, the partial sum processes approximate independent Brownian motions under $\tilde{P}_0^{(n,T)}$.

Proposition 4.1 *Let Assumption 1 hold and let $h_n \to h \in \mathbb{R}$. Then, under $\tilde{P}_0^{(n,T)}$,*

$$\log \frac{\mathrm{d}\tilde{P}_{h_n}^{(n,T)}}{\mathrm{d}\tilde{P}_0^{(n,T)}} = h_n \tilde{\Delta}_{n,T} - \frac{1}{2} h_n^2 \tilde{J}_{n,T}, \qquad (7)$$

where the central-sequence $\tilde{\Delta}_{n,T}$ and the finite-sample Fisher information $\tilde{J}_{n,T}$ are given by

$$\tilde{\Delta}_{n,T} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} H_i X_i^{(T)} \text{ and } \tilde{J}_{n,T} = \frac{1}{n} \sum_{i=1}^{n} H_i^2 J_i^{(T)},$$

with

$$X_i^{(T)} = \int_0^1 W_i^{(T)}(u-)\, \mathrm{d}W_i^{(T)}(u) \text{ and } J_i^{(T)} = \int_0^1 \left(W_i^{(T)}(u-) \right)^2 \mathrm{d}u.$$

Moreover, still under $\tilde{P}_0^{(n,T)}$, $\tilde{\Delta}_{n,T} \xrightarrow{d} \mathcal{N}(0, \mathbb{E}H_1^2/2)$ and $\tilde{J}_{n,T} \xrightarrow{p} \mathbb{E}H_1^2/2$ as $(T, n \to \infty)$.

[3] We adopt the convention that an empty sum is equal to 0.

See Appendix A.2 for a proof.

In empirical applications we, of course, only observe Y_{it}. Denote the law of Y_{it}, $i = 1, \dots, n$ and $t = 1, \dots, T$, by $P_h^{(n,T)}$. Then we have the following relation between the likelihood ratios of the model of interest and the auxiliary model discussed above:

$$\frac{\mathrm{d}P_h^{(n,T)}}{\mathrm{d}P_0^{(n,T)}} = \mathbb{E}\left[\left.\frac{\mathrm{d}\tilde{P}_h^{(n,T)}}{\mathrm{d}\tilde{P}_0^{(n,T)}}\right| \mathcal{F}_{n,T}\right], \tag{8}$$

with $\mathcal{F}_{n,T} := \sigma(Y_{it},\, i = 1, \dots, n,\, t = 1, \dots, T)$. To determine the limit experiment we need to obtain the asymptotic behavior of the conditional expectation as $(T, n \to \infty)$. Recall that an expansion holds under the joint limit $(T, n \to \infty)$ if and only if it holds for all non-decreasing sequences $T = T(n) \to \infty$ as $n \to \infty$. Therefore we can consider an arbitrary sequence $T(n) \to \infty$ in the following. Inserting (7) in (8) we obtain, under $P_0^{(n,T)}$,

$$\frac{\mathrm{d}P_{h_n}^{(n,T)}}{\mathrm{d}P_0^{(n,T)}} = \mathbb{E}\left[\left.\exp\left(\frac{h_n}{\sqrt{n}}\sum_{i=1}^n H_i X_i^{(T)} - \frac{h_n^2}{2n}\sum_{i=1}^n H_i^2 J_i^{(T)}\right)\right| \mathcal{F}_{n,T}\right].$$

This conditional expectation cannot be calculated explicitly as we do not impose explicit distributional assumptions on H_i. Exploiting Lemma 3.1, using $\mathcal{G}_n = \mathcal{F}_{n,T}$, $\mathcal{X}_{ni} = X_i^{(T)}$, $\mathcal{J}_{ni} = J_i^{(T)}$, and $U_i = H_i$, we show that the model of interest is also Locally Asymptotically Normal (LAN). We organize the result in the following proposition; see Appendix A.3 for the proof.

Proposition 4.2 *Let Assumption 1 hold and $h_n \to h$. Then, under $P_0^{(n,T)}$ as $(T, n \to \infty)$,*

$$\log\frac{\mathrm{d}P_{h_n}^{(n,T)}}{\mathrm{d}P_0^{(n,T)}} = h\Delta_{n,T} - \frac{1}{2}h^2 J + o_P(1), \tag{9}$$

with central-sequence $\Delta_{n,T} = n^{-1/2}\sum_{i=1}^n X_i^{(T)}$ and Fisher-information $J = 1/2$. Furthermore, still under $P_0^{(n,T)}$, $\Delta_{n,T} \xrightarrow{d} \mathcal{N}(0, J)$ as $(T, n \to \infty)$. □

Remark 1 *If $P(H_1 = 1) < 1$, i.e. the perturbations are non-homogeneous, the Fisher information J is strictly less than the Fisher information $\tilde{J}$ for the model in which the perturbations H_i are also observed, i.e. there is "loss of information". This can also be observed from the relation between the central-sequences: $\Delta_{n,T} = \mathbb{E}\left[\tilde{\Delta}_{n,T} \mid \mathcal{F}_{n,T}\right]$.*

The proposition implies that the model is of the LAN type, i.e. the limit experiment is a Gaussian shift experiment[4] $X \sim \mathcal{N}(Jh, J)$. An application of the Asymptotic Representation Theorem, see, e.g., Van der Vaart (1991, Theorem 7.2) or Van der Vaart (2000, Chapter 15), yields the (asymptotic) power envelope for testing the hypothesis (4).

Corollary 4.1 *Let Assumption 1 hold and let $\alpha \in (0,1)$. Let $T_{n,T} = t_{n,T}(Y_{11}, \ldots, Y_{nT})$ be a sequence of level α tests, i.e. $\limsup_{n,T\to\infty} \pi_{n,T}(0) \leq \alpha$, where $\pi_{n,T}(h)$ denotes the power of $T_{n,T}$ under $P_h^{(n,T)}$. Then we have, for all $h \leq 0$,*

$$\limsup_{n,T\to\infty} \pi_{n,T}(h) \leq \Phi\left(-z_\alpha - \frac{h}{\sqrt{2}}\right), \tag{10}$$

where $z_\alpha = \Phi^{-1}(1-\alpha)$. And the test statistic $T^\star_{n,T} = 1\{\sqrt{2}\Delta_{n,T} \leq -z_\alpha\}$ attains this upper bound (for all h).

Remark 2 *In a similar fashion it follows that the power envelope for the auxiliary model in which we also observe the perturbations H_i is given by $\Phi\left(-z_\alpha - h\sqrt{\mathbb{E}H_1^2/2}\right)$ (and is attained by a test based on $\tilde{\Delta}_{n,T}$). We, of course, have*

$$\Phi\left(-z_\alpha - \frac{h}{\sqrt{2}}\right) \leq \Phi\left(-z_\alpha - h\sqrt{\frac{\mathbb{E}H_1^2}{2}}\right), \quad h \leq 0,$$

with equality if and only if H_1 is degenerated, i.e. when the alternatives are homogeneous. This inequality should be compared to Display (10), and the discussion thereafter, in Moon, Perron, and Phillips (2007). The implications of our results are as follows.

(a) In case one observes both Y_{it} and H_i it is possible to attain the right-hand-side of the preceding display.

(b) Using a test that is only based on the observations Y_{it} it is not possible, in presence of heterogeneous alternatives, to achieve the right-hand-side of the previous display. In fact, we end up with the same power envelope as for the model with homogeneous alternatives.

The corollary completes the picture for the model in which m_i and σ_i are known. Of course, in empirical applications the parameters m_i and σ_i are (unknown) nuisance parameters and the statistic $T^\star_{n,T}$ is infeasible. Then, (10) still yields an upper bound to the (asymptotic) power of valid tests. In the next subsection we show that the bound is sharp, i.e. not knowing the nuisance parameters m_i and σ_i does not affect the bound.

[4]Local Asymptotic Normality and convergence of experiments is in most books (see, e.g., Van der Vaart (2000)) presented for 'univariate' limits $n \to \infty$. The abstract theory, however, is formulated in terms of nets and the results can thus be applied in the joint-limit setting as well; see, e.g., Van der Vaart (1991).

4.2. Asymptotically UMP tests

In this section we consider the situation where the incidental intercepts m_i and the scale parameters σ_i are (unknown) nuisance parameters and discuss two test statistics that attain the bound (10) and thus are (asymptotically) UMP. In fact, this demonstrates that the Gaussian panel unit root problem is adaptive with respect to the parameters m_i and σ_i.

We impose the following additional assumption, which is standard in the panel unit root literature and is needed to control the heterogeneity in σ_i.

Assumption 2 $n/T \to 0$.

Remark 3 *Proposition 4.2 is, of course, not affected by Assumption 2. Corollary 4.1 still holds true when we read all limits $(n, T \to \infty)$ as limits for which Assumption 2 holds. When we refer to Corollary 4.1 in the following, we actually refer to this adapted version.*

We first introduce a new optimal test. Estimate σ_i^2, $i = 1, \dots, n$, by

$$\hat{\sigma}_{i,nT}^2 = \frac{1}{T-1} \sum_{t=2}^{T} (\Delta Y_{it})^2 \tag{11}$$

and introduce

$$\hat{\Delta}_{n,T} = \frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=3}^{T} \left(\sum_{s=2}^{t-1} \frac{1}{\hat{\sigma}_{i,nT}} \Delta Y_{is} \right) \frac{1}{\hat{\sigma}_{i,nT}} \Delta Y_{it}. \tag{12}$$

Note that $\hat{\Delta}_{n,T}$ only depends on the data via ΔY_{it} for $t = 2, \dots, T$ and thus its distribution is invariant with respect to the incidental intercepts m_i. Under the null hypothesis its distribution does not depend on the scale parameters σ_i.

Proposition 4.3 *Let Assumptions 1 and 2 hold. Then we have, for all $h \in \mathbb{R}$ and under $P_h^{(n,T)}$ as $(T, n \to \infty)$,*

$$\hat{\Delta}_{n,T} = \Delta_{n,T} + o_P(1), \tag{13}$$

and the test $T_{n,T}^f = 1\{\sqrt{2}\hat{\Delta}_{n,T} \le -z_\alpha\}$ is (asymptotically) optimal.

Remark 4 *The optimality of $T_{n,T}^f$ immediately follows from (13) and Corollary 4.1. As we have shown that our model is LAN, an application of Le Cam's first lemma (see, e.g., Van der Vaart (2000, Lemma 6.4)) shows that we only have to prove (13) under $P_0^{(n,T)}$ (and thus avoid "triangular array calculations").*

Next we reconsider the point optimal invariant test based on the statistic $\hat{V}_{fe1,nT}$ proposed by Moon, Perron, and Phillips (2007, Section 4). The implementation of this test requires, besides the data Y_{it}, the specification of a sequence C_i, $i \in \mathbb{N}$, of random variables. Reason for this is that $\hat{V}_{fe1,nT}$ is based on a likelihood ratio statistic for testing the null hypothesis $\rho_i = 1$ versus the alternative $\rho_i = 1 - C_i/(\sqrt{n}T)$. Under our Assumptions 1-2, and some additional regularity, an application of Theorems 9-10 in Moon, Perron, and Phillips (2007) shows

$$\hat{V}_{fe1,nT} \xrightarrow{d} \mathcal{N}\left(h\mathbb{E}C_1H_1, 2\mathbb{E}C_1^2\right).$$

A combination with Corollary 4.1 yields the following conclusions.

(a) In case one observes both Y_{it} and H_i choosing $C_i = H_i$ yields an (asymptotically) optimal test.

(b) In case one only observes Y_{it} (and possibly auxiliary data whose distribution, conditional on the panel units, does not depend on h) choosing $C_i = 1$ yields an (asymptotically) optimal test.

Our results clarify the role of the random perturbations H_i and the interaction with the C_i variables used in the $\hat{V}_{fe1,nT}$ statistic. These results confirm a conjecture in Breitung and Pesaran (2008): "This suggests that if there is no information about variation of H_i, then a test cannot be improved by taking into account a possible heterogeneity of the alternative.". This concludes the discussion of optimal tests for the unit root hypothesis in the model (1)-(3).

4.3. Extension to cross-sectional dependence case

This section discusses an extension of the model allowing for cross-sectional dependence. For convenience we ignore the possible presence of fixed effects in this extension. As model for observations $Z_{\cdot,t} = (Z_{1t}, \ldots, Z_{nt})'$ we consider

$$Z_{\cdot,t} = \Xi Z_{\cdot,t-1} + u_{\cdot,t}, \quad t = 1, \ldots, T, \tag{14}$$

where $Z_{\cdot,0} = 0_n$ and $u_{\cdot,1}, \ldots, u_{\cdot,T}$ are i.i.d. $\mathcal{N}(0_n, \Omega_n)$ with Ω_n non-singular. And the $n \times n$ matrix of autoregression coefficients is assumed to be generated by the random coefficient structure

$$\Xi = I_n + \frac{h}{\sqrt{n}T}\Omega_n^{1/2}\mathcal{H}_n\Omega_n^{-1/2}, \tag{15}$$

where $\mathcal{H}_n = \text{diag}(H_1, \ldots, H_n)$ and the random perturbations H_i satisfy Assumption 1(b). Note that for $\Omega_n = \text{diag}(\sigma_1^2, \ldots, \sigma_n^2)$ (14)-(15) indeed reduce to (2)-(3). The specification for Ξ is motivated by a rotation of the cross-sectional independence model as will become clear below.

If we consider the model in which the $n \times n$ matrix Ω_n is known, observing Z_{it}, $t = 1, \ldots, T$ and $i = 1, \ldots, n$, is equivalent to observing $\tilde{Y}_{\cdot,t} = \Omega_n^{-1/2} Z_{\cdot,t}$. Using (14) and (15) we obtain

$$\tilde{Y}_{\cdot,t} = \left(I_n + \frac{h}{\sqrt{n}T} \mathcal{H}_n \right) \tilde{Y}_{\cdot,t-1} + \varepsilon_{\cdot,t},$$

where $\varepsilon_{\cdot,t} = \Omega_n^{-1/2} u_{\cdot,t} \sim \mathcal{N}(0_n, I_n)$. The observations $\tilde{Y}_{it}$ can thus be seen as observations from the panel unit root model (1)-(3). A combination with Proposition 4.2 and Corollary 4.1 yields the following result (here $Q_h^{(n,T)}$ denotes the underlying probability measure corresponding to (14)-(15)).

Corollary 4.2 *Let Assumption 1(b) hold and let $\alpha \in (0,1)$. Let $T_{n,T} = t_{nT}(Z_{11}, \ldots, Z_{nT})$ be a sequence of level α tests, i.e. $\limsup_{n,T\to\infty} \pi_{n,T}(0) \leq \alpha$, where $\pi_{n,T}(h)$ denotes the power of $T_{n,T}$ under $Q_h^{(n,T)}$. Then we have, for all $h \leq 0$,*

$$\limsup_{n,T\to\infty} \pi_{n,T}(h) \leq \Phi\left(-z_\alpha - \frac{h}{\sqrt{2}} \right). \tag{16}$$

Remark 5 *The right-hand-side of (16) still yields an upper bound to the asymptotic power of valid tests in case Ω_n is a matrix of nuisance parameters.*

(a) In case one would restrict to using sequential asymptotics (see Phillips and Moon (1999) for precise definitions), $n \to \infty$ after $T \to \infty$, then the arguments in Breitung and Das (2005) can be followed to obtain an optimal test under the additional assumption that the maximum eigenvalue of Ω_n is bounded.

(b) By using a factor structure

$$u_{it} = \gamma_i F_t + \eta_{it}, \quad i, t \in \mathbb{N},$$

where (the unobserved) factors F_t are i.i.d. $\mathcal{N}(0,1)$ independent of the idiosyncratic shocks η_{it} which are assumed to be independent and $\mathcal{N}(0, \sigma_i^2)$ distributed, one obtains $\Omega_n = \gamma^{(n)}\gamma^{(n)\prime} + \text{diag}(\sigma_1^2, \ldots, \sigma_n^2)$ with $\gamma^{(n)} = (\gamma_1, \ldots, \gamma_n)'$. Assume that $P(H_1 = 1) = 1$, i.e. consider homogeneous alternatives. Then the t-statistics t_a^* and t_b^* proposed by Moon and Perron (2004), which are based on a modified pooled OLS estimator, behave as, see Theorem 2 in Moon and Perron (2004) (and under the assumptions to this theorem), under $Q_h^{(n,T)}$,

$$t_a^*,\ t_b^* \xrightarrow{d} \mathcal{N}\left(h\sqrt{\frac{\omega^2}{2\phi}}, 1 \right),$$

where the existence of $\omega = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \sigma_i^2$ *and* $\phi = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \sigma_i^4$ *is part of their assumptions. Corollary 4.2 shows that rejecting the null if* t_a^*, $t_b^* \leq -z_\alpha$ *is (asymptotically) optimal only if* $\phi = \omega^2$, *i.e. these tests are optimal only if the heterogeneity in the scale parameter is very limited or non-existent. This gives a partial answer to the question, raised by Moon, Perron, and Phillips (2007), if "defactoring" the residuals as proposed in Moon and Perron (2004) yields optimal tests.*

We conjecture that it is possible to modify the t-statistics of Moon and Perron (2004) to obtain a test attaining the right-hand-side of (16).

A. Proofs

A.1. Proof of Lemma 3.1

The result is immediate if $\sigma_U^2 = \text{var}(U_1) = 0$. So we consider $\sigma_U^2 > 0$. Let, for some $\tilde{\eta} > 0$, ϕ denote the moment generating function of $U_1 - \mathbb{E}(U_1)$ on the interval $(-\tilde{\eta}, \tilde{\eta})$. We have $m_n = m_{n1} m_{n2}$ for

$$m_{n1} = \exp\left(-\frac{\sigma_U^2}{2n}\sum_{i=1}^n \mathcal{J}_{ni}\right),$$

$$m_{n2} = \mathbb{E}\left[\exp\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n (U_i - \mathbb{E}U_1)\mathcal{X}_{ni}\right)\exp\left(-\frac{1}{2n}\sum_{i=1}^n (U_i^2 - \mathbb{E}U_1^2)\mathcal{J}_{ni}\right)\middle| \mathcal{G}_n\right].$$

Introduce, for $a \in \mathbb{R}$,

$$M_n(a) = \mathbb{E}\left[\exp\left(\frac{a}{\sqrt{n}}\sum_{i=1}^n (U_i - \mathbb{E}U_i)\mathcal{X}_{ni}\right)\middle| \mathcal{G}_n\right],$$

and decompose $m_n = m_{n1}m_{n2} = m_{n1}M_n(1) + m_{n1}(m_{n2} - M_n(1))$. To enhance readability we organize the proof of $m_n \xrightarrow{p} 1$ in two steps. In Step A we show $m_{n1}M_n(1) \xrightarrow{p} 1$, and in Step B we establish $m_{n2} - M_n(1) \xrightarrow{p} 0$. This also yields $m_{n1}(m_{n2} - M_n(1)) \xrightarrow{p} 0$ since $m_{n1} \leq 1$.

Step A By (6) it suffices (as $\phi'(0) = 0$ and $\phi''(0) = \sigma_U^2$) to show $\log(M_n(1)) - \frac{1}{2}\phi''(0)\frac{1}{n}\sum \mathcal{X}_{ni}^2 \xrightarrow{p} 0$. Let $\epsilon > 0$ arbitrary. Choose $\eta \in (0, \tilde{\eta})$ such that, for all $|x| \leq \eta$,

$$\begin{aligned} |\log(1+x) - x| &\leq \epsilon|x|, \\ \left|\phi(x) - 1 - \frac{1}{2}\phi''(0)x^2\right| &\leq \epsilon x^2, \end{aligned}$$

and $(\frac{1}{2}\phi''(0) + \epsilon)\eta \leq 1$. Then on the event $\mathcal{A}_n = \{\max_{i=1,\ldots,n} |\mathcal{X}_{ni}|/\sqrt{n} \leq \eta\}$,

$$\left|\phi\left(\frac{\mathcal{X}_{ni}}{\sqrt{n}}\right) - 1\right| \leq \left(\frac{1}{2}\phi''(0) + \epsilon\right)\frac{\mathcal{X}_{ni}^2}{n} \leq \left(\frac{1}{2}\phi''(0) + \epsilon\right)\eta^2 \leq \eta.$$

Thus, still on the event $\mathcal{A}_n$,

$$\left|\log(M_n(1)) - \frac{1}{2}\phi''(0)\sum_{i=1}^{n}\frac{\mathcal{X}_{ni}^2}{n}\right| \leq \sum_{i=1}^{n}\left|\log\left(\phi\left(\frac{\mathcal{X}_{ni}}{\sqrt{n}}\right)\right) - \frac{1}{2}\phi''(0)\frac{\mathcal{X}_{ni}^2}{n}\right|$$
$$= \sum_{i=1}^{n}\left|\log\left(\phi\left(\frac{\mathcal{X}_{ni}}{\sqrt{n}}\right)\right) - \left\{\phi\left(\frac{\mathcal{X}_{ni}}{\sqrt{n}}\right) - 1\right\}\right| + \sum_{i=1}^{n}\left|\phi\left(\frac{\mathcal{X}_{ni}}{\sqrt{n}}\right) - 1 - \frac{1}{2}\phi''(0)\frac{\mathcal{X}_{ni}^2}{n}\right|$$
$$\leq \epsilon\left(1 + \frac{1}{2}\phi''(0) + \epsilon\right)\frac{1}{n}\sum_{i=1}^{n}\mathcal{X}_{ni}^2.$$

Using $n^{-1}\sum_{i=1}^{n}\mathcal{X}_{ni}^2 = O_P(1)$ and $\mathbb{P}(\mathcal{A}_n) \to 1$, we obtain the desired convergence to zero.

Step B For each $\epsilon > 0$ choose $M > 0$ and N such that, for the event

$$\mathcal{B}_n = \left\{\frac{1}{n}\sum_{i=1}^{n}\mathcal{J}_{ni} \leq M\right\} \cap \left\{\max_{i=1,\ldots,n}\frac{\mathcal{J}_{ni}}{n} \leq 1\right\} \cap \left\{\max_{i=1,\ldots,n}\frac{|\mathcal{X}_{ni}|}{\sqrt{n}} < \tilde{\eta}/2\right\} \in \mathcal{G}_n,$$

we have $\mathbb{P}(\mathcal{B}_n^c) \leq \epsilon$ when $n \geq N$. Part B follows if $|m_{n2} - M_n(1)|1_{\mathcal{B}_n} \xrightarrow{p} 0$. We have, using Cauchy-Schwarz,

$$\mathbb{E}|m_{n2} - M_n(1)|1_{\mathcal{B}_n} \leq \sqrt{\mathbb{E}r_n 1_{\mathcal{B}_n}}\sqrt{\mathbb{E}M_n(2)1_{\mathcal{B}_n}},$$

where

$$r_n = \mathbb{E}\left[\left(\exp\left(-\frac{1}{2n}\sum_{i=1}^{n}(U_i^2 - \mathbb{E}U_1^2)\mathcal{J}_{ni}\right) - 1\right)^2 \middle| \mathcal{G}_n\right].$$

On the event $\mathcal{B}_n$ $M_n(2)$ is uniformly bounded (which follows by similar arguments as in Step A), so $\mathbb{E}M_n(2)1_{\mathcal{B}_n}$ is bounded in n. We show that $\mathbb{E}r_n 1_{\mathcal{B}_n} \to 0$. The conditional Markov inequality implies, for all $\delta > 0$,

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n}(U_i^2 - \mathbb{E}U_1^2)\mathcal{J}_{ni}\right| > \delta \middle| \mathcal{G}_n\right) \leq \frac{1}{\delta^2}\operatorname{var}(U_1^2)\left(\max_{i=1,\ldots,n}\frac{|\mathcal{J}_{ni}|}{n}\right)\frac{1}{n}\sum_{i=1}^{n}\mathcal{J}_{ni},$$

which converges to 0 in probability by (5) and (6). As $r_n 1_{\mathcal{B}_n}$ is uniformly bounded an application of the continuous mapping theorem, the law of iterated expectations, and the bounded convergence theorem yields $\mathbb{E}r_n 1_{\mathcal{B}_n} \to 0$. Conclude that $\mathbb{E}|m_{n2} - M_n(1)|1_{\mathcal{B}_n} \to 0$, which concludes Step B and the proof.

A.2. Proof of Proposition 4.1

We have

$$\log\frac{\mathrm{d}\tilde{P}_{h_n}^{(n,T)}}{\mathrm{d}\tilde{P}_0^{(n,T)}} = \frac{h_n}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{H_i}{\sigma_i^2}(Y_{i,t-1} - m_i)\Delta Y_{it} - \frac{h_n^2}{2}\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{H_i^2}{\sigma_i^2}(Y_{i,t-1} - m_i)^2$$

$$= h_n \tilde{\Delta}_{n,T} - \frac{1}{2} h_n^2 \tilde{J}_{n,T}.$$

An application of the following proposition (with $U_i = H_i$) yields, under $\tilde{P}_0^{(n,T)}$, $\tilde{\Delta}_{n,T} \xrightarrow{d} \mathcal{N}\left(0, \mathbb{E}H_1^2/2\right)$ and $\tilde{J}_{n,T} \xrightarrow{p} \mathbb{E}H_1^2/2$ as $(T, n \to \infty)$. This completes the proof. □

Proposition A.1 *Let Assumption 1 hold and let U_i be i.i.d. with mean 1 and finite fourth moment independent of ε_{it}, $i, t \in \mathbb{N}$. Then we have, with $X_i^{(T)}$ and $J_i^{(T)}$ as defined in Proposition 4.1 and Y_{it} generated under $h = 0$, and as $(T, n \to \infty)$ and for any $\delta > 0$,*

$$\operatorname{plim} \frac{1}{n} \sum_{i=1}^{n} U_i^2 J_i^{(T)} = \operatorname{plim} \frac{1}{n} \sum_{i=1}^{n} U_i^2 \left(X_i^{(T)}\right)^2 = \frac{1}{2} \mathbb{E}U_1^2, \tag{17}$$

$$\lim \mathbb{E}U_1^2 \left(X_1^{(T)}\right)^2 1\left\{\left|U_1 X_1^{(T)}\right| > \delta\sqrt{n}\right\} = 0, \tag{18}$$

and

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} U_i X_i^{(T)} \xrightarrow{d} \mathcal{N}(0, \mathbb{E}U_1^2/2). \tag{19}$$

Proof: The convergence in probability (17) follows from $\mathbb{E}U_1^2 J_1^{(T)} = \mathbb{E}U_1^2 \left(X_1^{(T)}\right)^2 = [T(T-1)/(2T^2)]\mathbb{E}U_i^2$ and an application of the Markov inequality (using that var $\left(X_1^{(T)}\right)^2$ and var $J_1^{(T)}$ are both bounded and using that U_1 has a finite fourth moment).

Let $\delta > 0$. Note that $I_T := \sum_{j=1}^{T} B(t_{j-1}^{(T)}) \left(B(t_j^{(T)}) - B(t_{j-1}^{(T)})\right)$, where $t_j^{(T)} = j/T$ and B is a standard Brownian motion, and $X_1^{(T)}$ have the same distribution. As $T \to \infty$ we have $I_T \to \int_0^1 B(u)\,\mathrm{d}B(u) =: I$ in L_2. This implies $I_T^2 \to I^2$ in L_1. Hence the collection of random variables $\{I_T^2, T \in \mathbb{N}\}$ is uniformly integrable. As uniform integrability of a collection of random variables is determined by the marginal laws, we can conclude that the sequence $\left\{\left(X_1^{(T)}\right)^2, T \in \mathbb{N}\right\}$ is uniformly integrable. This yields $\mathbb{E}\left[\left(X_1^{(T)}\right)^2 1\left\{\left|U_1 X_1^{(T)}\right| > \delta\sqrt{n}\right\} \mid U_1\right] \to 0$. As this conditional expectation is also bounded (we already noted that $\mathbb{E}\left(X_1^{(T)}\right)^2 = O(1)$) an application of the dominated convergence theorem yields (18).

As the $U_i X_i^{(T)}$ are i.i.d. with mean zero and the Lindeberg condition (18) holds, an application of the joint-limit central limit theorem (see Phillips and Moon (1999, Theorem 2)) yields (19). □

A.3. Proof of Proposition 4.2

We first verify the conditions to Lemma 3.1. Recall that $\mathcal{G}_n = \mathcal{F}_{n,T(n)}$, $\mathcal{X}_{ni} = X_i^{(T(n))}$, $\mathcal{J}_{ni} = J_i^{(T(n))}$, and $U_i = H_i$. Condition i) is satisfied because of Assumption 1. The first part of (5) and the first part of (6) follow from Proposition A.1 (using $U_i = 1$). The second part of (5) is a consequence of (18) (using $U_i = 1$). And the second part of (6) follows from, for $\epsilon > 0$ and with probabilities and expectations calculated under $P_0^{(n,T(n))}$,

$$\begin{aligned} P\left(\max_{i=1,\dots,n} \frac{J_i^{(T(n))}}{n} > \epsilon\right) &\leq \frac{1}{n}\sum_{i=1}^n \mathbb{E} J_i^{(T(n))} 1\left\{J_i^{(T(n))} > \epsilon n\right\} \\ &= \mathbb{E}\left[\int_0^1 \left(W_1^{(T(n))}(u-)\right)^2 \mathrm{d}u 1\left\{\int_0^1 \left(W_1^{(T(n))}(u-)\right)^2 \mathrm{d}u > \epsilon n\right\}\right] \\ &\to 0, \end{aligned}$$

which follows along similar lines as used in the proof of Proposition A.1. The lemma thus yields, under $P_0^{(n)}$,

$$\begin{aligned} \log \frac{\mathrm{d}P_{h_n}^{(n)}}{\mathrm{d}P_0^{(n)}} &= \log \mathbb{E}\left[\exp\left(\frac{h_n}{\sqrt{n}}\sum_{i=1}^n H_i X_i^{(T(n)} - \frac{h_n^2}{2n}\sum_{i=1}^n H_i^2 J_i^{(T(n))}\right)\middle| \mathcal{F}_n\right] \\ &= \frac{h_n}{\sqrt{n}}\sum_{i=1}^n X_i^{(T(n))} - \frac{1}{2}\frac{h_n^2}{n}\sum_{i=1}^n J_i^{(T(n))} + o_P(1) = h_n\Delta_n - \frac{h_n^2}{2n}\sum_{i=1}^n J_i^{(T(n))} + o_P(1). \end{aligned}$$

Another application of Proposition A.1 (with $U_i = 1$) yields $\Delta_n \xrightarrow{d} \mathcal{N}(0, 1/2)$ and $n^{-1}\sum_{i=1}^n J_i^{(T(n))} \xrightarrow{p} 1/2$, which concludes the proof.

A.4. Proof of Proposition 4.3

As noted in Remark 4 we only need to prove (13) for $h = 0$. In the following, all probabilities and expectations are evaluated under $P_0^{(n,T)}$.

Decompose $\hat{\Delta}_{n,T} - \Delta_{n,T} = I_{n,T} - II_{n,T}$ with

$$I_{n,T} = \frac{1}{\sqrt{n}}\sum_{i=1}^n \left(\frac{\sigma_i^2}{\hat{\sigma}_{i,nT}^2} - 1\right)\int_0^1 W_i^{(T)}(u-)\,\mathrm{d}W_i^{(T)}(u),$$

$$II_{n,T} = \frac{1}{\sqrt{nT}}\sum_{i=1}^n \varepsilon_{i1}\sum_{t=2}^T \varepsilon_{it} + \frac{1}{\sqrt{nT}}\sum_{i=1}^n \left(\frac{\sigma_i^2}{\hat{\sigma}_{i,nT}^2} - 1\right)\varepsilon_{i1}\sum_{t=2}^T \varepsilon_{it}.$$

To see that $I_{n,T}$ converges to 0 in probability we first apply Cauchy-Schwarz to obtain

$$|I_{n,T}|^2 \leq \sum_{i=1}^n \left(\frac{\sigma_i^2}{\hat{\sigma}_{i,nT}^2} - 1\right)^2 \times \frac{1}{n}\sum_{i=1}^n \left(\int_0^1 W_i^{(T)}(u-)\,\mathrm{d}W_i^{(T)}(u)\right)^2.$$

As the second term is $O_P(1)$, $I_{n,T} = o_P(1)$ follows from

$$\sum_{i=1}^{n}\left(1-\frac{\hat{\sigma}^2_{i,nT}}{\sigma_i^2}\right)^2 = o_P(1) \text{ and } \lim P\left(\sup_{1\le i\le n}\frac{\sigma_i^2}{\hat{\sigma}^2_{i,nT}} \le 3/2\right) = 1,$$

where the second statement in the display follows from the first one, and the first statement follows from

$$\mathbb{E}\left[\sum_{i=1}^{n}\left(\frac{\hat{\sigma}^2_{i,nT}}{\sigma_i^2}-1\right)^2\right] = \sum_{i=1}^{n}\operatorname{var}\left(\frac{1}{T-1}\sum_{t=2}^{T}\varepsilon_{it}^2\right) = 2\frac{n}{T-1} \to 0.$$

As the first term of $II_{n,T}$ is centered and its variance, given by $n(T-1)/(nT^2)$, tends to 0 it follows that this term converges to 0 in probability. The second term of $II_{n,T}$ is seen to convergence to 0 in probability by similar arguments as for $I_{n,T}$. This completes the proof.

References

Baltagi, B.H. and C. Kao (2000). Nonstationary Panels, Cointegration in Panels and Dynamic Panels, A Survey, in: B. Baltagi (ed.), Nonstationary Panels, Panel Cointegration, and Dynamic Panels, Advances in Econometrics, Vol. 15, JAI Press, Amsterdam, 752.

Banerjee, A. (1999), Panel Data Unit Roots and Cointegration: An Overview, *Oxford Bulletin of Economics and Statistics* 61, 607629.

Breitung, J. (2000). The Local Power of Some Unit Root Tests for Panel Data, in B. Baltagi (ed.), Nonstationary Panels, Panel Cointegration, and Dynamic Panels, Advances in Econometrics, Vol. 15, JAI Press, Amsterdam, 161–178.

Breitung, J. and S. Das (2005). Panel unit root tests under cross-sectional dependence, *Statistica Neerlandica* 59, 414–433.

Breitung J., and M.H. Pesaran (2008). Unit roots and cointegration in panels. In *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*, Mátyás L., Sevestre, P. (eds). Kluwer: Dordrecht.

Choi, I. (2006), Nonstationary Panels, in K. Patterson and T.C. Mills (eds.) Palgrave Handbooks of Econometrics, Vol. 1, 511539. Palgrave Macmillan, New York.

Drost, F.C., R. van den Akker, and B.J.M. Werker (2009). The asymptotic structure of nearly unstable non negative integer-valued AR(1) models, *Bernoulli* 15, 297–324.

Hallin, M., R. van den Akker, and B.J.M. Werker (2011). A class of simple distribution-free rank-based unit root tests, *Journal of Econometrics* 163, 200–214.

Harris, D., D.I. Harvey, S.J. Leybourne, and N.D. Sakkas (2010). Local asymptotic power of the Im-Pesaran-Shin panel unit root test and the impact of initial observations, *Econometric Theory* 26, 311–324.

Jansson, M. (2008). Semiparametric power envelopes for tests of the unit root hypothesis. *Econometrica* 76, 1103–1142.

Le Cam, L. (1986). Asymptotic methods in statistical decision theory. New York: Springer-Verlag.

Le Cam, L. and G.L. Yang (1988). On the preservation of Local Asymptotic Normality under information loss, *Annals of Statistics* 16, 483–520.

Le Cam, L. and G.L. Yang (1990). Asymptotics in statistics - some basic concepts. New York: Springer.

Madsen, E. (2010). Unit root inference in panel data models where the time-series dimension is fixed: a comparison of different tests, *Econometrics Journal* 13, 63–94.

Moon, H. R. and B. Perron (2004). Testing for a unit root in panels with dynamic factors, *Journal of Econometrics* 122, 81–126.

Moon, H. R. and B. Perron (2008). Asymptotic local power of pooled t-ratio tests for unit roots in panels with fixed effects, *Econometrics Journal* 11, 80–104.

Moon, H. R., B. Perron, and P.C.B. Phillips (2007). Incidental trends and the power of panel unit root tests, *Journal of Econometrics* 141, 416–459.

P.C.B. Phillips and H.R. Moon (1999). Limit theory for nonstationary panel data, *Econometrica* 67, 1057–1111.

Van der Vaart, A.W. (1991). An asymptotic representation theorem, *International Statistical Review* 59, 97–121.

Van der Vaart, A.W. (2000). Asymptotic statistics. Cambridge: Cambridge University Press.

Westerlund, J. and J. Breitung (2013). Lessons from a Decade of IPS and LLC, *Econometric Reviews* 32, 547–591.