Journal of Statistical Planning and Inference 173 (2016) 1–30

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

A simple test for nonstationarity in mixed panels: A further investigation[☆]

Joakim Westerlund [*]

Lund University, Sweden
Centre for Economics and Financial Econometrics Research, Deakin University, Australia

ARTICLE INFO

Article history:
Received 28 January 2015
Received in revised form 20 December 2015
Accepted 6 January 2016
Available online 13 January 2016

JEL classification:
C13
C33

Keywords:
Unit root test
Panel data
Local asymptotic power

ABSTRACT

The test of Ng (2008) is one of the few that enables general inference regarding the proportion of non-stationary units in panel data. The current paper furthers the investigation of Ng (2008) in two directions. First, the existing sequential limit analysis is generalized to a very flexible asymptotic framework in which the number of time periods, T, can be either fixed or tending to infinity jointly with the number of cross-section units, N. Second, the test statistic is evaluated not only under the null hypothesis, but also under alternatives that can be either fixed or local.

1. Introduction

Consider the panel data variable $y_{i,t}$, observable for $t = 1, \ldots, T$ time series and $i = 1, \ldots, N$ cross-section units. While applications of panel unit root tests to such variables are now commonplace, there are still ambiguities as how best to interpret the test results. In a recent note, Pesaran (2012) emphasizes that a rejection of the panel unit root hypothesis should be interpreted as evidence that a proportion $\theta < 1$ of the cross-sectional units are unit root non-stationary, which is not very informative. He therefore recommends augmenting the test outcome with an estimate of θ. Unfortunately, most existing panel unit root tests do not lead to such an estimate. One exception is the test of Ng (2008), which is based on a direct estimator of $\theta \in (0, 1]$.[1] Hence, unlike other tests (see Breitung and Pesaran, 2008; Baltagi, 2008, Chapter 12, for surveys of the panel unit root and cointegration literatures), the test of Ng (2008) is appropriate in general when wanting to infer θ, and not just when testing $H_0 : \theta = \theta_0 = 1$ versus $H_1 : \theta \in (0, 1)$. What is more, this advantage seems to come at no expense in terms of test construction. In fact, it is difficult to imagine a simpler test.

In this paper we extend the work of Ng (2008) in two directions. First, while very convenient, as is well known (see, for example, Westerlund and Breitung, 2013; Phillips and Moon, 1999), sequential asymptotics are unlikely to be enough to capture actual behavior. Ng (2008) assumes that $N \to \infty$ before $T \to \infty$, which means that her results are subject to this

☆ The author would like to thank one anonymous referee and Serena Ng for many valuable comments and suggestions. Thanks also to the Knut and Alice Wallenberg Foundation for financial support through a Wallenberg Academy Fellowship, and to the Jan Wallander and Tom Hedelius Foundation for financial support under research grant number P2014–0112:1.

* Correspondence to: Department of Economics, Lund University, Box 7082, 220 07 Lund, Sweden. Tel.: +46 46 222 8997; fax: +46 46 222 4613.
E-mail address: joakim.westerlund@nek.lu.se.

[1] Another approach that can be used to infer θ is that of Pesaran (2007). However, this approach only allows for consistent estimation of θ under the null hypothesis that $\theta = 1$, and cannot be used when wanting to infer $\theta < 1$.

http://dx.doi.org/10.1016/j.jspi.2016.01.004
0378-3758/

same critique. In the current paper we take this as our starting point to develop a new more general theory. The new theory is based on a finite-sample expansion of the test statistic that retains not only the first order terms but also higher order terms (see Westerlund and Larsson, 2015, for a discussion). The expansion is evaluated in two ways; (i) as $N \to \infty$ with T held fixed, and (ii) as $N, T \to \infty$ jointly. The reason for this is that we want to understand not only the observed test behavior for a given T, but also how that behavior changes as T is allowed to increase. Except for Im et al. (2003), to the best of our knowledge this is the only panel unit root study to consider both the fixed-T and large-T cases, an undertaking that is shown to be very rewarding. Indeed, the theoretical predictions are shown to be very accurate, even in very small samples.

Second, the sequential asymptotic analysis of Ng (2008) only covers the behavior under the null hypothesis, and there is no analysis of power. Therefore, in order to compensate for this, in the current paper we evaluate power against two types of alternatives. On the one hand, if the alternative is "local-to-unity" in the sense that the deviation from the unit root null goes to zero as $N \to \infty$, then we show that while power is non-negligible, θ is no longer estimable, not even if $N, T \to \infty$ jointly. On the other hand, if the alternative is "non-local" in the sense that the deviation from the null does not depend on the sample size, then we show that power is increasing in N and that θ is again estimable. These results complement nicely the discussion in Pesaran (2012, page 546), who states that: "To identify the exact proportion of the sample for which the null hypothesis is rejected one requires country-specific data sets with T sufficiently large". While in principle correct, in light of the new results provided here, it is clear that having T large enough is not a sufficient condition for identification of θ; for this to happen the deviation from the null must also not be "too small".

2. Model and assumptions

The DGP of $y_{i,t}$ is similar to the one considered in Ng (2008), and is given by

$$y_{i,t} = \lambda_i + u_{i,t}, \tag{1}$$

$$u_{i,t} = \alpha_i u_{i,t-1} + \epsilon_{i,t}, \tag{2}$$

where $u_{i,0} = 0$. The unit-specific intercept λ_i and the error $\epsilon_{i,t}$ satisfy the following assumptions.

Assumption 1. $\epsilon_{i,t}$ independently and identically distributed (iid) across both i and t with $E(\epsilon_{i,t}) = 0$, $E(\epsilon_{i,t}^2) = \sigma_\epsilon^2 > 0$ and $E(\epsilon_{i,t}^4)/\sigma_\epsilon^4 = \kappa_\epsilon < \infty$.

Assumption 2. λ_i can be random or non-random, provided that $\sigma_{\lambda,N}^2 = \sum_{i=1}^N (\lambda_i - \overline{\lambda})^2/N \to_p \sigma_\lambda^2 < \infty$ as $N \to \infty$, where $\overline{\lambda} = \sum_{i=1}^N \lambda_i/N$ and $\to_p$ signifies convergence in probability. If λ_i is random, it should be independent of α_i and $\epsilon_{i,t}$.

Our first main departure from the setup of Ng (2008) is the modeling of α_i. Let us therefore assume without loss of generality that the first $N_1 \geq 1$ units have $\alpha_i = 1$ and that the remaining $N_0 = N - N_1$ units have $\alpha_i < 1$. Thus, in this notation,

$$\theta = \frac{N_1}{N} \in (0, 1]. \tag{3}$$

The null hypothesis of interest is that $H_0 : \theta = \theta_0$, which is equivalent to requiring $\alpha_1 = \cdots = \alpha_{N_1} = 1$. A common way to set up the alternative hypothesis is to assume that $\alpha_{N_1+1}, \ldots, \alpha_N$ are "non-local" (or fixed) in the sense that the degree of mean reversion is not allowed to depend on the sample size. However, with such a specification we only learn if the test is consistent and, if so, at what rate. To be able to evaluate the power analytically, we therefore have to consider an alternative in which α_i is local-to-unity as $N \to \infty$. Assumption 3 nests both types of alternatives.

Assumption 3.

$$\alpha_i = \exp\left(\frac{c_i}{N^\eta}\right), \tag{4}$$

where $\eta \geq 0$ and $c_i \leq 0$ is a random drift parameter such that $c_1 = \cdots = c_{N_1} = 0$ and $c_{N_1+1}, \ldots, c_N$ independent with $E(|c_i|^p) < \infty$ for all p and $i = N_1 + 1, \ldots, N$. Also, c_i is independent of $\epsilon_{i,t}$.

Let us denote by $\mu_{1,p}$ and $\mu_{0,p}$ the p-order moments of $c_1, \ldots, c_{N_1}$ and $c_{N_1+1}, \ldots, c_N$, respectively, and let μ_p be the corresponding moment of $c_1, \ldots, c_N$. The above specification with both non-stationary and stationary units implies that c_i has a mixture distribution, whose moments are weighted sums of the moments of the two component distributions. Hence, since $\mu_{1,p} = 0$ for all $p \geq 1$, we have that $\mu_p = \theta\mu_{1,p} + (1-\theta)\mu_{0,p} = (1-\theta)\mu_{0,p}$. If $p = 0$, then we define $\mu_{1,0} = \mu_{0,0} = 1$, suggesting that $\mu_0 = 1$. Although this is not strictly necessary, to simplify the analysis, we assume that $E(|c_i|^p) < \infty$ for all p and $i = N_1 + 1, \ldots, N$, such that all the moments of $c_{N_1+1}, \ldots, c_N$ exist.[2] The "closeness" of the local alternative to

[2] Most of the existing literature (see, for example, Moon and Perron, 2008; Moon et al., 2007) supposes that the support of c_i is bounded, which implies finite moments. In our case, strictly speaking the assumption of finite moments is only required in the case when $\eta = 0$ (α_i is non-local), which is not restrictive in the sense that the case with α_i explosive does not seem very realistic.