On the Interpretation of Panel Unit Root Tests*

M. Hashem Pesaran

University of Cambridge and USC

Trinity College, Cambridge, CB2 1TQ, England

MHP1@cam.ac.uk

September 2011

Abstract

Applications of panel unit root tests have become commonplace in empirical economics, yet there are ambiguities as how best to interpret the test results. This note clarifies that rejection of the panel unit root hypothesis should be interpreted as evidence that a statistically significant proportion of the units are stationary. Accordingly, in the event of a rejection, and in applications where the time dimension of the panel is relatively large, it recommends the 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. The economic importance of the rejection can be measured by the magnitude of this proportion.

JEL Classifications: C12, C33, C52

Key Words: Unit Root tests, Panels, Statistical Significance

*Helpful suggestions by Michael Binder, Ron Smith, and Elisa Tosetti are gratefully acknowledged.

Over the last decade considerable work has been carried on unit root testing in panel data models. See, for example, Breitung and Pesaran (2008) for a recent survey of the literature. Most panel unit root tests are designed to test the null hypothesis of a unit root for each individual series in a panel. The formulation of the alternative hypothesis is instead a controversial issue that critically depends on which assumptions one makes about the nature of the homogeneity/heterogeneity of the panel. A number of panel unit root tests proposed in the literature use the following articulation of the alternative hypothesis:

$$H_1^a : \text{Each of the series are stationary as a panel,}$$

while other tests use:

$$\begin{aligned} H_1^b &: \text{At least one of the series in} \\ &\quad \text{the panel is generated by a stationary process.} \end{aligned}$$

The above two formulations of the alternative hypothesis are not satisfactory for carrying inference on the non-stationarity properties of panel data models. To see why this is so, consider the following simple dynamic heterogeneous panel on N cross sections observed over T time periods:

$$y_{it} = (1-\phi_i)\,\mu_i + \phi_i y_{i,t-1} + \varepsilon_{it},\ i = 1,2,...,N;\ t = 1,2,...,T, \tag{1}$$

where initial values, y_{i0}, are assumed to be given. (1) can be expressed as

$$\Delta y_{it} = \alpha_i + \beta_i y_{i,t-1} + \varepsilon_{it}, \tag{2}$$

where $\alpha_i = (1-\phi_i)\,\mu_i,\ \beta_i = -(1-\phi_i)$ and $\Delta y_{it} = y_{it} - y_{i,t-1}$. The null hypothesis of unit roots can then be written as

$$H_0 : \beta_i = 0 \text{ for all } i. \tag{3}$$

Under a homogeneous alternative we have $\beta_i = \beta \neq 0$ for all i, and

$$H_1^a : \beta < 0,$$

is meaningful. Testing for unit roots within homogeneous dynamic panels is considered, for example, by Harris and Tzavalis (1999) and Levin, Lin, and Chu (2002). One drawback of tests based on such alternative hypothesis is that they usually have power also if not all units are stationary; and hence a rejection is not convincing evidence that all series are indeed stationary. In particular, Westerlund and Breitung (2009) show that the local power of the Levin, Lin, and Chu (2002) test is greater than that of the Im, Pesaran, and Shin (2003) test, based on a less restrictive alternative, also when not all individual series are stationary. A further drawback in using H_1^a is that this is likely to be unduly restrictive, particularly for cross-country studies involving differing short-run dynamics. For example, such homogeneous alternative seems particularly inappropriate in the case of the purchasing power parity (PPP) hypothesis, where y_{it} is taken to be the real exchange rate. There are no theoretical grounds for the imposition of the homogeneity hypothesis, $\beta_i = \beta$, under PPP. The alternative hypothesis H_1^b stands at the other extreme and in terms of the above notations states that

$$H_1^b : \beta_i < 0, \text{ for one or more } i.$$

Such alternative hypothesis is at the basis of panel unit root tests proposed by Chang (2002) and Chang (2004). We observe that H_1^b is only appropriate when N is finite, namely within the multivariate model with a fixed number of variables analyzed in the time series literature. On the contrary, in the case of large N and T, panel unit root tests will lack power if the alternative, H_1^b, is adopted. For large N and T panels it is reasonable to entertain alternatives that lie somewhere between the two extremes of H_1^a and H_1^b. In this context, a more appropriate alternative is given by

$$H_1^c : \beta_i < 0, \ i = 1, 2, \ldots, N_1, \ \beta_i = 0, \ i = N_1 + 1, N_1 + 2, \ldots, N, \tag{4}$$

such that

$$\lim_{N \to \infty} \frac{N_1}{N} = \delta, \ 0 < \delta \leq 1. \tag{5}$$

Using the above specification the null hypothesis is $H_0 : \delta = 0$, while the alternative hypotheses H_1^c can be written as

$$H_1^c : \delta > 0.$$

In other words, rejection of the unit root null hypothesis can be interpreted as providing evidence in favour of rejecting the unit root hypothesis for a non-zero fraction of panel members as $N \to \infty$. In cases where T is sufficiently large, δ

can be estimated by application of univariate unit root tests to all the individual time series in the panel. With N and T sufficiently large, a consistent estimate of δ will then be given by the proportion of the cross-section units for which the individual unit root tests are rejected. In applications where N and T are not sufficiently large, δ can not be estimated consistently, although the panel unit root test outcome could still be valid, in the sense that it has the correct size under the null hypothesis, H_0.

A number of recent papers have considered the problem of estimating the proportion of stationary units, δ. In the context of testing for output and growth convergence, Pesaran (2007) suggests to use the proportion of unit root tests applied to pairs of log per-capita output gaps across N economies, for which the null hypothesis of non-stationarity is rejected at a given significance level, α. The author proves that, although the underlying individual unit-root tests are not cross-sectionally independent, under the null hypothesis of non-stationarity such average rejection statistic converges to α, as N and T jointly tend to infinity. Ng (2008) shows that, if a fraction, δ, of the panel is made of stationary units, with the remaining series having unit roots, the cross-sectional variance of the panel will have a linear trend that increases exactly at rate $1-\delta$. Hence, she suggests a statistic for the proportion of non-stationary units $(1-\delta)$ based on the time average of the sample cross-sectional variance.

While these procedures deliver an estimate of the fraction of (non)stationary units, they are not designed to identify which units are stationary. After rejection of the null hypothesis of unit roots for each individual series in a panel, it is often of interest to identify which series can be considered stationary and which can be deemed non-stationary. Kapetanios (2003) and Chortareas and Kapetanios (2009) propose a sequential panel selection method that consists of applying the Im, Pesaran, and Shin (2003)'s panel unit root test sequentially on progressively smaller fractions of the original data set, where the reduction is carried out by dropping series for which there is evidence of stationarity, signalled by low individual t-statistics. A similar approach is taken by Smeeks (2010), who proposes testing on user-defined fractions of the panel, using panel unit root tests based on order statistics and computing the corresponding critical values by block bootstrap. Hanck (2009) and Moon and Perron (2010) apply methods from the literature on multiple testing to classify the individual series into stationary and non-stationary sets. In particular, Moon and Perron (2010) suggest the use of the so-called false discovery rate (FDR), given by the expected fraction of series classified as I(0) that are in fact I(1), as a useful diagnostic on the aggregate decision. In the computation of the FDR, the authors estimate the fraction of true null hypotheses by applying the Ng (2008)'s approach described above.

The heterogeneity of panel data models used in cross-country analysis introduces a new kind of asymmetry in the way the null and the alternative hypotheses are treated, which is not usually present in the univariate time series (or cross-section) models. This is because the same null hypothesis is imposed across all i but the specification of the alternative hypothesis is allowed to vary with i. This asymmetry is assumed away in homogeneous panels. However, as demonstrated in Pesaran and Smith (1995) neglected heterogeneity (even if purely random) can lead to spurious results in dynamic panels. Therefore, in cross-country analysis where slope heterogeneity is a norm, the asymmetry of the null and the alternative hypotheses has to be taken into account. The appropriate response critically depends on the relative size of N and T. In large N heterogeneous panel data models with T small (say around 15) it is only possible to devise sufficiently powerful unit root tests which are informative in some average sense, namely whether the null of a unit root can be rejected in the case of a significant fraction of the countries in the panel.[1] 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. But if T is large enough for reliable country-specific inferences to be made, then there seems little rationale in pooling countries into a panel.

References

Breitung, J. and M. H. Pesaran (2008). Unit roots and cointegration in panels. In L. Matyas and P. Sevestre (Eds.), *The Econometrics of Panel Data* (Third ed.). Springer-Verlag.

Chang, Y. (2002). Nonlinear iv unit root tests in panels with cross-sectional dependency. *Jounal of Econometrics 110*, 261–292.

[1]Some of these difficulties can be circumvented if slope heterogeneity can be modelled in a sensible and parsimonious manner.

Chang, Y. (2004). Bootstrap unit root test in panels with cross-sectional dependency. *Journal of Econometrics 120*, 263–293.

Chortareas, G. and G. Kapetanios (2009). Getting PPP right: Identifying mean-reverting real exchange rates in panels. *Journal of Banking & Finance 33*, 390–404.

Hanck, C. (2009). For which countries did PPP hold? A multiple testing approach. *Empirical Economics 37*, 93–103.

Harris, R. D. F. and H. E. Tzavalis (1999). Inference for unit roots in dynamic panels where the time dimension is fixed. *Jounal of Econometrics 91*, 201–226.

Im, K. S., M. H. Pesaran, and Y. Shin (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics 115*, 53–74.

Kapetanios, G. (2003). Determining the poolability properties of individual series in panel datasets. Queen Mary, University of London Working Paper No. 499.

Levin, A., C. Lin, and C. Chu (2002). Unit root tests in panel data: Asymptotic and finite-sample properties. *Journal of Econometrics 108*, 1–24.

Moon, H. R. and B. Perron (2010). Beyond panel unit root tests: Using multiple testing to determine the non stationarity properties of individual series in a panel. Department of Economics working paper 2010-04, University of Montreal.

Ng, S. (2008). A simple test for nonstationarity in mixed panels. *Journal of Business & Economic Statistics 26*, 113–127.

Pesaran, M. H. (2007). A pair-wise approach to testing for output and growth convergence. *Jounal of Econometrics 138*, 312–355.

Pesaran, M. H. and R. Smith (1995). Estimation of long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics 68*, 79–113.

Smeeks, S. (2010). Bootstrap sequential tests to determine the stationary units in a panel. Maastricht University.

Westerlund, J. and J. Breitung (2009). Myths and facts about panel unit root tests. Mimeo, University of Gothenburb.