Journal of Econometrics 175 (2013) 94–115

Contents lists available at SciVerse ScienceDirect

Journal of Econometrics

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

Panel unit root tests in the presence of a multifactor error structure☆

M. Hashem Pesaran[a,b], L. Vanessa Smith[c], Takashi Yamagata[d,*]

[a] University of Southern California, USA
[b] Trinity College, Cambridge, United Kingdom
[c] CFAP & CIMF, University of Cambridge, United Kingdom
[d] Department of Economics and Related Studies, University of York, United Kingdom

ARTICLE INFO

Article history:
Received 16 September 2009
Received in revised form
23 October 2012
Accepted 12 February 2013
Available online 18 February 2013

JEL classification:
C12
C15
C22
C23

Keywords:
Panel unit root tests
Cross section dependence
Multifactor error structure
Fisher inflation parity
Real equity prices

ABSTRACT

This paper extends the cross-sectionally augmented panel unit root test (*CIPS*) proposed by Pesaran (2007) to the case of a multifactor error structure, and proposes a new panel unit root test based on a simple average of cross-sectionally augmented Sargan–Bhargava statistics (*CSB*). The basic idea is to exploit information regarding the m unobserved factors that are shared by k observed time series in addition to the series under consideration. Initially, we develop the tests assuming that m^0, the true number of factors, is known and show that the limit distribution of the tests does not depend on any nuisance parameters, so long as $k \geq m^0 - 1$. Small sample properties of the tests are investigated by Monte Carlo experiments and are shown to be satisfactory. Particularly, the proposed *CIPS* and *CSB* tests have the correct size for *all* combinations of the cross section (N) and time series (T) dimensions considered. The power of both tests rises with N and T, although the *CSB* test performs better than the *CIPS* test for smaller sample sizes. The various testing procedures are illustrated with empirical applications to real interest rates and real equity prices across countries.

1. Introduction

There is now a sizeable literature on testing for unit roots in panels where both cross section (N) and time series (T) dimensions are relatively large. Reviews of this literature are provided in Banerjee (1999), Baltagi and Kao (2000), Choi (2004), and Breitung and Pesaran (2008). The so called first generation panel unit root tests pioneered by Levin et al. (2002) and Im et al. (2003) focused on panels where the errors were cross-sectionally uncorrelated. More recently, to deal with a number of applications such as testing for purchasing power parity or cross country output convergence, the second generation panel unit root tests have focused on the case where the errors are allowed to be cross-sectionally correlated.

Three main approaches have been proposed. The first, pioneered by Maddala and Wu (1999), and developed further by Chang (2004), Smith et al. (2004), Cerrato and Sarantis (2007) and Palm et al. (2011), apply bootstrap methods to panel unit root tests. The main idea of this approach is to approximate the distribution of the test statistic under cross section dependence by block bootstrap resampling to preserve the pattern of cross section dependence in the panel. This approach allows for general cross section dependence structures, however, it is mainly suited to panels with large T and relatively small N.

The second approach is due to Bai and Ng (2004, 2010) and proposes tests based on a decomposition of the observed series, y_{it}; $i = 1, 2, \ldots, N, t = 1, 2, \ldots, T$, into two unobserved components, common factors and idiosyncratic errors, and tests for unit roots in both of these components. It is also tested if the unobserved common factors are cointegrated. This is known as the PANIC (panel analysis of nonstationarity in idiosyncratic and common components) approach, and provides indirect tests of unit roots in the observed series. The factors are estimated from m^0 principal components (PC) of Δy_{it}. It is assumed that m^0, the true number of factors, is known or estimated from the observations. If it is found that the estimated factors contain unit roots and are not cointegrated it is then concluded that the N series are integrated of order 1. If the presence of a unit root in the factors is rejected, in the second stage the PANIC procedure

☆ We would like to thank Anindya Banerjee, Soren Johansen and Joachim Westerlund for useful comments. This version has also benefited greatly from constructive comments and suggestions by the Editor (Cheng Hsiao), an associate editor and three anonymous referees.

* Corresponding author. Tel.: +44 0 1904 323758.
E-mail address: ty509@york.ac.uk (T. Yamagata).

0304-4076/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jeconom.2013.02.001

applies panel unit root tests to the N idiosyncratic errors. Estimates of idiosyncratic errors are obtained as defactored observations, also known as PANIC residuals. Moon and Perron (2004) follow a similar approach in that they base their test on a principal components estimator of common factors. In particular, their test is based on defactored observations obtained by projecting the panel data onto the space orthogonal to the estimated factor loadings. Bai and Ng (2010) propose two panel unit root tests which are applied to the PANIC residuals. The first one is based on a pooled estimate of the autoregressive root fitted to the PANIC residuals, as in Moon and Perron (2004), and the second one employs a panel version of the modified Sargan–Bhargava test (*PMSB*).[1]

The third approach, proposed in Pesaran (2007), augments the individual Dickey–Fuller (DF) regressions of y_{it} with cross section averages, $\bar{y}_{t-1} = N^{-1}\Sigma_{j=1}^{N} y_{j,t-1}$ and $\Delta\bar{y}_t$, to take account of error cross section dependence. These cross-sectionally augmented DF regressions can be further augmented with lagged changes $\Delta y_{i,t-s}$, $\Delta\bar{y}_{t-s}$, for $s = 1, 2, \ldots$, to deal with possible serial correlation in the residuals. These doubly augmented DF regressions are referred to as CADF regressions. The panel unit root test statistic is then computed as the average of the CADF statistics. It is shown that the average statistic is free of nuisance parameters but, due to non-zero cross correlation of the individual CADF statistics, the average statistic has a non-normal limit distribution as N and $T \to \infty$. Monte Carlo experiments show that Pesaran's test has desirable small sample properties in the presence of a single unobserved common factor but show size distortions if the number of common factors exceeds unity.[2] A small sample comparison of some of these tests is provided in Gengenbach et al. (2009).[3]

The data generating mechanisms underlying the PANIC approach differ in one important respect from the ones considered by Moon and Perron (2004) and Pesaran (2007). The latter studies assume that under the null of unit roots the common factor components have the same order of integration as the idiosyncratic components, whilst the PANIC approach allows the order of integration of the factors to differ from that of the idiosyncratic components. However, if the primary objective of the exercise is to test for unit roots in the observed series, y_{it}, the distinction between the common and idiosyncratic components of y_{it} is not essential. The distinction will become relevant if the unit root null hypothesis of y_{it} is not rejected. In that case it would indeed be of interest to investigate further whether the source of the non-stationarity lies with the common factors, the idiosyncratic components, or both.

The present paper extends Pesaran's *CIPS* test to the case of a multifactor error structure. This is a non-trivial yet important extension which is much more broadly applicable. It has also the advantage of being intuitive and simple to implement. Following Bai and Ng (2010) we also consider a panel unit root test based on simple averages of cross-sectionally augmented Sargan–Bhargava type statistics, which we denote by *CSB*. The presence of multiple unobserved factors poses a number of additional challenges. In order to deal with a multifactor structure, we propose to utilise the information contained in a number of k additional variables, $\mathbf{x}_{it}$, that together are assumed to share the common factors of the series of interest, y_{it}. The ADF regression for y_{it} is then augmented with cross-sectional averages of y_{it} and $\mathbf{x}_{it}$.[4]

The requirement of finding such additional variables seems quite plausible in the case of panel data sets from economics and finance where economic agents often face common economic environments. Most macroeconomic theories postulate the presence of the same unobserved common factors (such as shocks to technology, tastes and fiscal policy), and it is therefore natural to expect that many macroeconomic variables, such as interest rates, inflation and output share the same factors. If anything, it would be difficult to find macroeconomic time series that do not share one or more common factors. For example, in testing for unit roots in a panel of real outputs one would expect the unobserved common shocks to output (that originate from technology) to also manifest themselves in employment, consumption and investment. In the case of testing for unit roots in inflation across countries, one would expect the unobserved common factors that correlate inflation rates across countries to also affect short-term and long-term interest rates across markets and economies. The fundamental issue is to ascertain the nature of dependence and persistence that is observed across markets and over time. The present paper can, therefore, be viewed as a first step in the process of developing a coherent framework for the analysis of unit roots and multiple cointegration in large panels.

Initially we develop the tests assuming that m^0, the true number of factors is known, and show that the limit distribution of *CIPS* and *CSB* tests does not depend on any nuisance parameters, so long as $k \geq m^0 - 1$. But, in practice m^0 is rarely known. Most existing methods of estimating m^0, such as the information criteria of Bai and Ng (2002), assume that the unobserved factors are strong, in the sense discussed in Chudik et al. (2011). However, in many empirical applications we may not be sure that all the factors are strong. Bailey et al. (2012, BKP) show that the strength of the factors is determined by the nature of the factor loadings, and depends on the exponent of the cross-sectional dependence, α, defined as $\ln(n)/\ln(N)$, where n is the number of non-zero factor loadings. The value $\alpha = 1$ corresponds to the case of a strong factor, while $\alpha < 1$ gives rise to a large set of practically plausible values ranging from semi-strong to weaker factors. BKP find that for many macroeconomic and financial series of interest, the value of the exponent is less than one. This result casts some doubt on the practical justification of panel unit root tests based on estimated factors by principal components, which are discussed above. The solution offered in this paper deals with the uncertainty surrounding the true number of factors by assuming that there exists a sufficient number of k additional regressors that *together* share at least $m^0 - 1$ of the factors in the model that influence the variable under consideration. This approach does not require all the factors to be strong. This way, by selecting $k = m_{\max} - 1$, where $m_{\max}$ is the assumed maximum number of factors, the estimation of m^0 will not be needed.

For our tests, following the discussions in Im et al. (2003), we propose to use critical values which depend on the values of k, N, T, and lag-augmentation order, p, as they are expected to provide much better finite sample approximations. In empirical applications it is important that the tests being considered have the correct size, otherwise their use could result in misleading conclusions.

1 Westerlund and Larsson (2009) provide further theoretical results on the asymptotic validity of the pooled versions of the PANIC procedure.

2 The cross section augmentation procedure is also employed by Hadri and Kurozumi (2011) in their work on testing the null of stationarity in panels.

3 Other panel unit root tests have also been proposed by Chang (2002), who employs a non-linear IV method, Choi and Chue (2007) who use a subsampling method to account for cross-section correlation, and Phillips and Sul (2003) who use an orthogonalisation procedure to deal with error cross dependence in the case of a single common factor.

4 The idea of augmenting ADF regressions with other covariates has been investigated in the unit root literature by Hansen (1995) and Elliott and Jansson (2003). These authors consider the additional covariates in order to gain power when testing the unit root hypothesis in the case of a single time series. In this paper we augment ADF regressions with cross section averages to eliminate the effects of unobserved common factors in the case of panel unit root tests.