Unit root tests for panel data – a survey and an application

Ana Paula Barreira*
Faculdade de Economia
Universidade do Algarve

Paulo Rodrigues
Faculdade de Economia
Universidade do Algarve

September 2005

Abstract

The importance of a priori check of the existence of unit roots in the panel data comes from the already known effect that the presence of unit roots in time series may cause a misinterpretation of estimated results. Adding the cross-section dimension to the time series dimension offers an advantage in testing for nonstationary and cointegration since cross-section increases the data set used in those tests, thus improving their power. However, the cross-section dimension also brings some new problems into question, namely the existence of cross-section dependency which can bias usual panel data unit root test results in small samples. This paper presents a survey of panel unit root tests, evidencing the most recent developments on the issue, including those that account for the presence of contemporaneous cross-correlation as well as for the presence of heterogeneous serial correlation. Parallel to the developments of panel unit root tests, great attention has also been given to cointegration tests. We briefly review the most widely referred cointegration tests.

We apply the reviewed panel unit root tests on an EU social variable which represents the population weight over than 65 years of age. We consider data running from 1970 to 2001. The panel unit root test results reveal to be sensitive to the prior assumptions regarding contemporaneous cross-correlation and heterogeneous serial correlation in small samples. The usual battery of panel unit root tests appear not to be adequate when a panel is composed by a mix of a stationary and nonstationary time series.

* Corresponding author: aprodrig@ualg.pt; , University of Algarve, Faculty of Economics, Campus de Gambelas, P–8005–139 Faro, PORTUGAL; Tel.: (+351) 289 800 900; Fax: (+351) 289 818 303.

1. Introduction

Testing for unit roots in time series is now common practice among empirical studies.[1] However, testing for unit roots in panels is quite recent, having the major developments in nonstationary panel models occurred since the middle of the 1990s. Panel data applications have shifted from micro panels with large N (number of cross sections) and small T (length of the time series) to macro panels with large N and large T.

The recent attention given to the problem in the econometrics of panel data emerges from numerous applications of time series procedures to panels, for which issues such as nonstationarity, spurious regressions and cointegration are becoming important.

Specifically, in the time-series framework when we consider two independent random vectors Y_{it} and X_{it} that are $I(1)$, or equivalently, nonstationary, and without a cointegrating relation between them, then if a time series regression for a certain i is performed, the regression coefficient will have a nondegenerative limit distribution and the regression is characterized as spurious.

Adding the cross-section dimension to the time series dimension offers an advantage when testing for nonstationarity and cointegration since cross-section increases the number of observations used in those tests, thus improving their power. Simultaneously, the cross-sectional dimension also brings some new problems into question, namely the existence of cross-section dependency which can bias panel unit root test results in small samples. These problems are only attenuated when the panel has large cross sectional and time series dimensions.

The paper presents a survey of the panel unit root tests, evidencing the most recent developments on the issue, including those that account for the presence of contemporaneous cross-correlation as well as for the presence of heterogeneous serial correlation. Baltagi and Kao (2000) and Cerrato (2002) also provide surveys on panel unit root tests. The former reviews panel unit root tests that assume cross section independence and the latter summarizes the main developments in panel unit root tests allowing cross section dependence.

In this paper we present an extension of these surveys including the main developments thereafter. Parallel to the panel unit root test developments, great attention has been given to cointegration tests. We also briefly review the most widely used cointegration tests.

Panel unit root tests are then performed on a social variable which represents population weight over than 65 years of age. The panel unit root tests results reveal to

[1] For literature reviews of unit root tests in the conventional time series framework see Stock (1994), Maddala and Kim (1998) and Phillips and Xiao (1998).

be sensitive to the prior assumptions regarding contemporaneous cross-correlation and heterogeneous serial correlation in small samples. The usual battery of panel unit root tests appear not be adequate when a panel is composed by a mix of stationary and nonstationary time series.

This paper is organised as follows. Some commonly used unit root panel data tests which rely on cross-section independence and their recent developments are briefly reviewed in section 2. In section 3 we describe most recent unit root panel data tests which relax the assumption of cross-section independence. Section 4 presents a review of the most widely referred cointegration tests. In section 5 we use population weight over 65 years of age data, collected for the former 15 EU member states, during 1970 to 2001, to perform the panel unit root tests reviewed. The main drawbacks related to the panel unit root test procedures presented are then pointed out in one example. Section 6 presents the concluding remarks.

2. Panel unit root tests with the assumption of cross-section independence

The assumption of independence across i is rather strong and relies on the argument by Quah (1994)[2] that modelling cross-sectional dependence is involved given that there is no natural ordering of the individual observations in a cross-section.

Levin and Lin (1992, 1993), Im, Pesaran and Shin (2003) and Maddala and Wu (1999) are the most important references of panel unit root tests that rely on cross sectional independence.

2.1. Panel unit root test to homogeneous cross sections

The Levin and Lin (1992, 1993) test ─ henceforth referred to as LL test─ treats panel data as being composed of homogeneous cross-sections, thus performing a test on a pooled data series. The LL test for unit roots in panel data is computed based on the following model:

$$y_{it} = \rho_i y_{i,t-1} + z_{it}^{'}\gamma + u_{it}, \tag{2.1}$$

where $i = 1,\ldots,N$ and $t = 1,\ldots,T$, z_{it} is the deterministic component and u_{it} is a stationary process. z_{it} can be zero, one, fixed effects - μ_i, and fixed effects and a time trend.

Under the homogeneity assumption, the LL test assumes that $\rho_i = \rho$ for all i and that $u_{it} \sim iid\left(0,\sigma_u^2\right)$.

[2] The test proposed by Quah (1994) is based on pooled OLS and considers the following simple dynamic panel $y_{it} = \rho y_{i,t-1} + u_{it}$, with $i = 1,\ldots,N$, $t = 1,\ldots,T$ and u_{it} are independently and identically distributed across i and t, with finite variance σ^2.

In this sense, the LL test is defined as $H_0\!:\rho=1$ against the alternative hypothesis that $H_a:|\rho|<1$. The test procedure is then designed to evaluate the null hypothesis that each individual in the panel has unit root properties versus the alternative hypothesis that all cross section series in the panel are stationary.

Levin, Lin and Chu's (2002), referred to as LLC hereafter, suggest some adjustments to the unit root test described above.[3]

Using also a pooling approach, the unit root test is implemented by a three-step procedure. In step 1 Augmented Dickey-Fulley (ADF) regressions are estimated on each cross-section in the panel and residuals computed.

Using an ADF type regression as

$$\Delta y_{it}=\rho_i y_{i,t-1}+\sum_{j=1}^{p_i}\varphi_{ij}\Delta y_{it-j}+z_{it}^{'}\gamma+\varepsilon_{it}\,, \tag{2.2}$$

the residuals are obtained from the two following auxiliary regressions: $\hat{e}_{it}=\Delta y_{it}-\sum_{j=1}^{p_i}\hat{\pi}_{ij}\,\Delta y_{it-j}-\hat{z}_{it}^{'}\gamma$ and $\hat{v}_{i,t-1}=y_{i,t-1}-\sum_{j=1}^{p_i}\tilde{\pi}_{ij}\,\Delta y_{it-j}-\tilde{z}_{it}^{'}\gamma$.

The residuals are then weighted by the regression standard error of expression (2.2) to control for heterogeneity across cross sections, becoming $\bar{e}_{it}$ and $\bar{v}_{i,t-1}$.

In step 2, the ratio of long-run to short-run standard deviations is estimated, for each cross section, which is then used to adjust the mean of the t -bar statistic found in step 3, when the model includes either fixed effects or both fixed and time effects. The LLC test is the outcome of pooling all cross sectional and time series to estimate $\bar{e}_{it}=\delta\bar{v}_{i,t-1}+\bar{\varepsilon}_{it}$. The null hypothesis is now described as $H_0\!:\delta=0$, and the t-statistic defined as usual.

Given the small time dimension of most panels, the emphasis has been put on models with homogeneous dynamics. However, Pesaran and Shin (1995) and subsequent developments have shown the inconsistency of pooled estimators under dynamic heterogeneous panels.

2.2. Panel unit root test for heterogeneous cross sections

The homogeneity hypothesis can be however too restrictive since panel data can be composed by several cross-sections with different autoregressive coefficients. The main argument is that under the alternative hypothesis the same convergence rate across countries can bias panel unit root tests. Imposing homogeneity when coefficient heterogeneity is present in cross-section data can result in misleading conclusions.

[3] Bayoumi and MacDonald (1999) present some applications of the LLC unit root tests.

The Im, Pesaran and Shin (2003) panel unit root test ─ hereafter referred to as IPS test ─ presents an alternative to overcome this restriction.

An average of the ADF tests to contemplate the case when u_{it} is serially correlated and the correlation properties vary across cross sections is suggested by IPS. Relaxing the *iid* assumption of the u_{it}, it can be observed that when $u_{it} = \sum_{j=1}^{p_i} \varphi_{ij} u_{it-j} + \varepsilon_{it}$, the following regression model needs to be considered to test for the existence of unit roots in panel data:

$$y_{it} = \rho_i y_{i,t-1} + \sum_{j=1}^{p_i} \varphi_{ij} \Delta y_{it-j} + z_{it}^{'} \gamma + \varepsilon_{it} \,. \tag{2.3}$$

The null hypothesis is defined as $H_0 : \rho_i = 1$ for all i, whereas now the alternative hypothesis is given as $H_a : |\rho_i| < 1$, for at least one i.

This test relies on the autoregressive properties of each cross section, being the final result of the IPS test based on an average of the individual ADF statistics.

The IPS t-bar statistic is the average of the individual ADF statistics, *i.e.* $\bar{t} = \frac{1}{N} \sum_{i=1}^{N} t_{\rho_i}$, where t_{ρ_i} is the individual t-statistic for testing H_0 in (2.3).

The order of augmentation used for the ADF test in each cross-section can be chosen based on a information criteria such as the *Akaike Information Criterion* (AIC) or the *Schwarz Information Criterion* (BIC).[4]

In the limit, the IPS test as $T \to \infty$ followed by $N \to \infty$,[5] converges to $t_{IPS} = \frac{\sqrt{N}\left(\bar{t} - E\left[t_{iT} | \rho_i = 1\right]\right)}{\sqrt{Var\left[t_{iT} | \rho_i = 1\right]}} \Rightarrow N(0,1)$.

IPS assume that t_{iT} are *iid* and have finite mean as well as finite heterogeneous variances, σ_i^2. The values of $E\left[t_{iT} | \rho_i = 1\right]$ and $Var\left[t_{iT} | \rho_i = 1\right]$ have been computed by IPS via simulation for different values of T and $p_i^{'}$s. [6]

[4] Ng and Perron (2005) emphasize the importance of choosing a correct lag specification for the ADF regressions in finite samples and provide a guide to robust model selection. In Ng and Perron's (2001) article, the authors propose a *Modified Information Criteria* that seems to perform better than usual information criteria such as the AIC and the BIC, for samples with $T > 100$.

[5] Phillips and Moon (1999) clearly defined the implications of the way the cross sectional and time series dimension approach infinity for the definition of asymptotic properties of estimators and tests proposed for nonstationary panels, imposing cross-sectional independence.

[6] A 1997's version of the IPS article includes two tests, one corresponds to the t-bar statistic already described and the other relies on the average of Lagrange Multiplier (LM) tests. The LM-statistic allows for heterogeneity in the value of ρ_i and lets the errors ε_{it} be serially autocorrelated with different serial correlation properties across units.

The Maddala and Wu (1999) panel unit root test─ henceforth referred to as MW test ─ is inspired in a Fisher type test that combines P-values from unit root tests for each cross-section i.

Being in contradiction with LLC's alternative hypothesis that imposes a homogeneous ρ_i across cross sections, this test also allows for different autoregressive coefficients across i.

The MW unit root test is defined as $P = -2\sum_{i=1}^{N} \ln p_i$, with P being distributed as χ^2 with $2N$ degrees of freedom as $T_i \to \infty$ for all N.

This test presents an advantage over the IPS test since it does not require a balanced panel, however, the test presents also a significant disadvantage associated to the fact that the p-values must be derived through Monte Carlo simulation.

2.3. Comparison of the performance of the panel unit root tests

Comparing the previous three unit root tests is not appropriate since these procedures rely on different null hypothesis. Although, Breitung (2000) investigates the local power of the LL and IPS tests and concludes that the LL test is very sensitive to the lag augmentation. The author also found that the power of the LL and IPS tests is very sensitive to the specification of the deterministic terms.

Karlsson and Lothgren (1999) compare small sample power properties of the LL and the IPS tests and show that panel unit root tests can have high power when a small fraction of the series are stationary and low power when a large fraction is stationary.

MW also point out that under the presence of cross correlation, situation that neither of the three tests described above contemplates, the LL test is the one with the worst power performance.

Cumulatively, when the assumption of independent error terms is violated, the derived distributions of IPS and MW test statistics are no longer valid. In the former case the t-bar statistic does not have the stated variance and consequently the asymptotic normal distribution and in the second case the test does not have a χ^2 distribution. MW propose a bootstrap method that allows for a reduction of the size distortions of the test under cross sectional correlation, although it does not eliminate them.

Banerjee *et al.* (2001-b) compared the LL test, with the IPS LM-bar test, the IPS t-bar and with the MW test and found that under the presence of cross-unit cointegration relations, which imply the existence of weak exogeneity, the MW test presents the worst size distortions and the IPS-t-bar test it is the one which performs better.

Hadri (2000) also proposes a residual based LM test for the null hypothesis that time series for each i are stationary around a deterministic level or around a deterministic trend against the alternative of a unit root in panel data. Hadri considers the following two models $y_{it} = r_{it} + \varepsilon_{it}$ and $y_{it} = r_{it} + \beta_i t + \varepsilon_{it}$, with $r_{it} = r_{i,t-1} + u_{it}$. The null hypothesis is defined as $H_0 : \lambda = 0$ against $H_a : \lambda > 0$, where $\lambda = \frac{\sigma_u^2}{\sigma_\varepsilon^2}$. This test can be used as a complement to the tests that consider nonstationarity as the null hypothesis, since in this case, the null hypothesis corresponds to stationarity.[7]

Choi (2000) has explored this approach called confirmatory analysis. The author argues that combining a test under the null hypothesis of stationarity with a test under the null of unit root in panel data can improve the reliability of test inferences over using either test alone, when the two tests corroborate each other. Further, if under different null hypothesis the two tests reject their respective nulls simultaneously, this disagreement is a sign that the panel data under study has a mixed structure, where unit root time series coexist with stationary time series.

Banerjee *et al.* (2001-a) argue that the assumption of a tie absence between time series of a panel data is very often violated when analysing macroeconomic time series across countries. Further, the authors prove that the presence of cross sectional correlation adds a problem, since when it is not adequately considered as in the common cointegration analysis on panel data, it can lead to the finding of “spurious” cointegration in cross-section regressions.[8] In this sense, the common used unit root tests and cointegration methods can only be employed successfully in a restricted number of cases.

3. Panel unit root tests without the assumption of cross-section independence

3.1. Testing for cross-section independence

In order to evaluate if the time series that compose a panel are in fact correlated or not, Granger (1969) causality analysis is used. This test allows approaching the question of how much time series y Granger cause x and vice-versa. Note, however, that one does not imply the other.

After the selection of a reasonable lag length, the Granger causality analysis is based on two regressions, generally defined as

[7] Following this argument, Bac and Pen (2002) study the unit root properties of the health care expenditures and per capita GDP panel data, using both the IPS and the Hadry approaches.

[8] Larsson and Lyhagen (1999 and 2000) and Larsson, Lyhagen and Lothgren (2001) develop a likelihood-based panel test of cointegrating rank in heterogeneous panel models based on the average of the individual rank trace statistics as presented by Johansen (1995). However, the authors found that the test requires a large time series dimension to be implemented. Further, the model proposed allows cointegration within units but rules out cointegration relationships across the units and simultaneously assumes that the cointegration rank is the same for each unit, conditions which are often violated as shown by Banerjee *et al.* (2001-a).

$$\begin{aligned} y_t &= \alpha_0 + \alpha_1 y_{t-1} + ... + \alpha_l y_{t-l} + \beta_1 x_{t-1} + ... + \beta_l x_{t-l} + \varepsilon_t \\ x_t &= \alpha_0 + \alpha_1 x_{t-1} + ... + \alpha_l x_{t-l} + \beta_1 y_{t-1} + ... + \beta_l y_{t-l} + e_t \end{aligned},$$

for all possible pair of (x, y) time series in the panel and where l is the selected lag length.

The Granger causality is reported through an F-statistic that corresponds to a Wald-statistic for the joint null hypothesis that $\beta_1 = ... = \beta_l = 0$. The null hypothesis is therefore that *x* does not Granger-cause *y* in the first regression and that *y* does not Granger-cause *x* in the second regression.

Through the analysis of the relations that are established across cross sectional series, it is possible to evaluate the existence of cross section correlation and, consequently, the adequacy of panel unit root tests that rely on the assumption of cross sectional regressions independence.[9]

Unit root tests in panel data that take into account the existence of cross sectional correlation have recently being developed and follow two orientations. One solves the problem using a non-linear regression approach and the other relies on a Seemingly Unrelated Regression (SUR) procedure.

In what follows panel unit root tests that allow for the existence of dependencies across units are described, which configures a second branch of the literature in this research area.

3.2. The use of instrumental variables in panel unit root tests

To deal with the presence of cross-sectional dependency, the first model described is the panel unit root test proposed by Chang (2002), based on non-linear IV estimation of the usual ADF type regression for each cross-sectional unit, using as instruments non-linear transformations of the lagged levels. The test statistic is defined as an average of individual IV t-ratios, which is asymptotically normal, and does not require the tabulation of critical values.[10]

The IV test is performed on the following autoregressive model:

$$y_{it} = \rho_i y_{i,t-1} + \sum_{j=1}^{p_i} \varphi_{ij} \Delta y_{it-j} + \varepsilon_{it} \tag{3.1}$$

and the null hypothesis is given as $H_0 : \rho_i = 1$, against the alternative $H_a : |\rho_i| < 1$ for some i. The rejection of the null does not imply that the entire panel is stationary, while

[9] The cross sectional dependence can appear from cross-sectional correlation or from cross-sectional cointegration or from both, as pointed out by Bornhorst (2002). However, given the inability to deal with cointegration across cross-sections through a Johansen (1995) cointegration test given our *N* length, our panel unit root tests will be mainly concerned with the former form of cross-sectional dependence.

[10] In previous work, Chang (2004) deals with the presence of cross-sectional dependency using a bootstrap procedure, which leads to limiting distributions of the unit root tests in panels that are non-standard.

the unit root non-rejection means that all y_{it} 's have unit roots.[11] The cross-sectional dependency is present in the innovations ε_{it}.[12]

Expression (3.1) is estimated using a non-linear function $F(y_{i,t-1})$ for the lagged level $y_{i,t-1}$. For the lagged differences ($\Delta y_{i,t-1},\dots,\Delta y_{i,t-p_i}$), the variables themselves are used as instruments. The autoregressive order for each cross-sectional unit is selected using the *Schwartz Information Criterion* (BIC).

The use of the instrument $F(y_{i,t-1})$ requires that it be correlated with the regressor $y_{i,t-1}$. Chang (2002) presents some regularly integrable functions that might be used as instruments in the IV estimation. From them we chose the following function: $xe^{-|x|}$ as instrument for the IV panel unit root tests used in our application in section 5.

From the autoregressive coefficient of expression (3.1), estimated for each cross section, the IV t-ratio statistics are obtained by $Z_i = \dfrac{\hat{\rho}_i - 1}{s\left(\hat{\rho}_i\right)}$, where $s\left(\hat{\rho}_i\right)$ is the standard error of the IV estimator $\hat{\rho}_i$. The IV t-ratio Z_i follows a standard normal distribution and the cross-sectional independence is ensured by the integrable function used as instrument in individual estimations, since, as is shown by Chang (2002), the non-linear instruments $F(y_{i,t-1})$ and $F(y_{j,t-1})$ are asymptotically uncorrelated, even when $y_{i,t-1}$ and $y_{j,t-1}$ are correlated.

Chang's panel unit root test is based on an average of the cross-sectional t-ratios statistics. The average IV t-ratio statistic is thus defined as $S_N = \dfrac{1}{N}\sum_{i=1}^{N} Z_i$.

The limit theory of this panel unit root test is developed for models with no deterministic components, although it can easily be implemented to models with nonzero means and deterministic trends by replacing the lagged level $y_{i,t-1}$ in expression (3.1) with the adaptive demeaned $y^{\mu}_{i,t-1}$ or with the adaptive detrended $y^{\tau}_{i,t-1}$, respectively.

---

[11] Chang and Song (2002) using also as instruments non-linear transformations of the lagged levels, present a panel unit root test that accounts for cross-sectional dependencies of innovations and for the presence of cointegration across cross-sections. The unit root tests are performed at individual levels adding to expression (3.1) the component $\sum_{j=1}^{Q_i} \beta'_{ik} w_{i,t-k}$, where w_{it} are the covariates added to the ADF regression for the i-th cross-sectional unit.

[12] Choi's (2002) work introduces cross sectional dependency through error-components (more precisely through the time effect variable) and not through innovation terms of autoregressive processes. The panel unit root test is performed by combining p-values from the ADF test applied to each cross-section, whose nonstochastic trend components and cross-sectional correlations are eliminated following an approach suggested by Elliot, Rothenberg and Stock's (1996) for conventional unit root tests. However, a t-statistic computed from quasi-differenced data also suffers from a Nickell (1981) type bias so that a bias correction is required to obtain a reasonable test procedure.

Chang's test for the presence of a unit root is now based on the test regression $y^{\mu}_{it} = \rho_i y^{\mu}_{i,t-1} + \sum_{j=1}^{p_i} \varphi_{ij} \Delta y^{\mu}_{it-j} + \varepsilon_{it}$ or $y^{\tau}_{it} = \rho_i y^{\tau}_{i,t-1} + \sum_{j=1}^{p_i} \varphi_{ij} \Delta y^{\tau}_{it-j} + \varepsilon_{it}$, respectively. Knowing that the stochastic component y_{it} comes from a time series z_{it}, with a nonzero mean, as in the former case, given as $z_{it} = \mu_i + y_{it}$, or with more general deterministic time trend, as in the latter case, given as $z_{it} = \mu_i + \delta_i t + y_{it}$.

The transformations are called recursive because data is used up to period $t-1$ instead of using the full sample.

In this sense, the respective transformed time series are obtained as,

$$y^{\mu}_{it} = z_{it} - \frac{1}{t-1} \sum_{k=1}^{t-1} z_{ik}$$

$$y^{\mu}_{i,t-1} = z_{i,t-1} - \frac{1}{t-1} \sum_{k=1}^{t-1} z_{ik}$$

$\Delta y^{\mu}_{i,t-k} = \Delta z_{i,t-k}, \quad k = 1,\ldots, p_i$, for the adaptive demeaning and

$$y^{\tau}_{it} = z_{it} + \frac{2}{t-1} \sum_{k=1}^{t-1} z_{ik} - \frac{6}{t(t-1)} \sum_{k=1}^{t-1} k z_{ik} - \frac{1}{T_i} z_{iT_i}$$

$$y^{\tau}_{i,t-1} = z_{i,t-1} + \frac{2}{t-1} \sum_{k=1}^{t-1} z_{ik} - \frac{6}{t(t-1)} \sum_{k=1}^{t-1} k z_{ik}$$

$\Delta y^{\tau}_{i,t-k} = \Delta z_{i,t-k} - \frac{1}{T_i} z_{iT_i}, \quad k = 1,\ldots, p_i$, where the term z_{iT_i} / T_i is the total sample mean of Δz_{it}, for the adaptive detrending.

3.3. The use of seemingly unrelated regression (SUR) in panel unit root tests

The second approach considered that allows for the presence of contemporaneous cross-correlation and heterogeneous serial correlation of the regression residuals was suggested by Breuer, McNown and Wallace (1999), hereafter BNW.[13]

Returning to the IPS and the MW panel unit root tests, it can be said that these tests are joint hypothesis tests in the sense that they assume that under the null hypothesis all units of a panel contain a unit root. When the joint null hypothesis is rejected it is possible that one or a few time series in the panel contribute to this finding. Cumulatively, given that these tests allow for the autoregressive parameter to differ across cross sections under the alternative, then the rejection of the null hypothesis means that not all units of the panel contain a unit root. Effectively, a mixture of stationary and nonstationary time series can cohabit in the same panel data set.[14]

[13] A recent empirical application of the BNW panel unit root test is presented by Wagner (2003), who uses the test on budget stabilization funds in order to separate pool estimates, considering, on the one hand, American states that present stationary data and, on the other hand, those with nonstationary data.

[14] Kónya (2001) compares the LL, IPS, MW and BNW unit root tests performance on a panel composed by the logarithm of real GDP across OECD countries. The author emphasizes that the major

In this sense and having in mind the limitations associated with the previous panel unit root tests, the main advantage of BNW test is to be able of determining which cross sectional series rejects the null hypothesis of a unit root and which does not.

Taking into account this argument, the main incentive now is to test for each panel unit the null and the alternative hypothesis using a SUR framework, which exploits the information in the error covariances to produce efficient estimators and potentially more powerful test statistics. The structure of hypothesis follows the ADF specification used in the IPS test procedure. Given the use of the SUR frame and of the ADF type test regression we define this procedure as SURADF.

The panel specification that is used in the *SURADF* estimation is

$$\Delta y_{1,t} = \rho_1 y_{1,t-1} + \sum_{j=1}^{p_i} \varphi_{1,j} \Delta y_{1,t-j} + z_{1,t}^{'} \gamma_1 + \varepsilon_{1,t}$$

.

.

.

$$\Delta y_{N,t} = \rho_N y_{N,t-1} + \sum_{j=1}^{p_i} \varphi_{N,j} \Delta y_{N,t-j} + z_{N,t}^{'} \gamma_N + \varepsilon_{N,t}$$

where the null hypothesis is $H_0 : \rho_i = 0$ for each time series of the panel.

In general the *SURADF* is a more powerful test than the ADF test. For the $I(0)$ time series, BNW show, based on median rejection rates, that *SURADF* has twice the power or even more then a single equation ADF to reject the null hypothesis when the autoregressive coefficient on each $I(0)$ time series is 0.90. However, these power gains vanish for an autoregressive coefficient between 0.95 and 0.99.

The BNW test has however the disadvantage of converging to non standard distribution, implying the need for simulation of the necessary critical values. To compute these critical values it is necessary to consider the estimated covariance matrix for the system under analysis, the sample size and the number of panel units. This means that each study has its own critical values.

4. Cointegration tests in panel data

Parallel to the panel unit root test developments, great attention has been given to cointegration tests and estimation within regression models in panel data.

The most widely referred cointegration tests are the ones introduced by McCoskey and Kao (1998), Kao (1999) and Pedroni (1999). McCoskey and Kao (1998) derive a residual-based LM test for the null of cointegration in panel data, allowing for varying

advantage of BNW test is to permit the identification of those time series that are stationary and those that are nonstationary.

slopes and intercepts. The test evidences good power performance for panels where $T > 50$, being indicated for panels with a similar number of cross-sections ($N > 50$). Kao (1999) develops a framework for understanding the behaviour of spurious panel regression using a fixed-effects model to estimate panel regression when the dependent and independent variables are $I(1)$ processes. Kao also presents two types of cointegration tests for panel data; Dickey-Fulley (DF) and the Augmented DF type tests and derives the asymptotic distributions for each case. Pedroni (1999) derives asymptotic distributions and critical values for several residual-based tests with multiple regressors of the null of no cointegration in panels. The model includes regressions with fixed and time effects and allows heterogeneity across units resulting from the presence of cointegrating vectors and from the dynamics of the error process.

McKoskey and Kao (1999) compare the three panel data tests for cointegration. The authors found that in those cases where economic theory predicts a long run steady state relationship, the null of cointegration rather than the null of no cointegration seemed to be more appropriate.

In what concerns the estimation of cointegrating relations between variables, the principal references are the dynamic OLS (DOLS) approach proposed by Kao and Chiang (2000) and the fully modified OLS (FMOLS) procedure developed by Pedroni (2000).[15]

The DOLS estimation method describes a system of cointegrating regressions between y_{it} and x_{it}, where the disturbances are stationary. However the estimation procedure assumes too restrictive assumptions like independence across i of the error terms and no cointegration of the $I(1)$ regressors.

The FMOLS also assumes the same restrictions as DOLS but allows the associated serial correlation properties of the error processes to vary across individual units of the panel. Pedroni proposes, in this sense, a group mean t-statistic to test the null hypothesis that the cointegrating vector β between y_{it} and x_{it} is equal across sections: $H_0{:}\beta_i = \beta_0$ versus the alternative hypothesis that $H_a{:}\beta_i \neq \beta_0$ for all i, so that the values of β are not necessarily constrained to be homogeneous across units.

5. The implementation of panel unit root tests

In this section we analyse panel unit root characteristics using series of population weight over 65 years of age. We identify the panel data series as POP65 and our application includes data for the former fifteen EU member countries. The sample covers a period of 32 years. Published data are available on a year-to-year basis from

[15] Kao, Chiang and Chen (1999) present an application of these two estimation methods under panel cointegration.

1970 to 2001. Data sources are Eurostat (2001) until 1999 and CIA-World Factbook afterwards.

The first step consists of checking for the presence of a unit root in the variable. In this sense, both the IPS and the Chang panel unit root tests are applied.

Performing a standard ADF test on each individual time series with an intercept and with both an intercept and a time trend, the IPS test is implemented, exploiting the panel dimension of the data set. The number of lags p_i included in individual regressions to eliminate residual autocorrelation is chosen according to the *BIC Criterion* with a maximum length of 8 periods. To confirm absence of residual serial correlation as required for the implementation of IPS' t-bar test, a LM test is performed on the residuals with a lag span of three.

IPS panel unit root test results show that the null hypothesis of a unit root cannot be rejected for the panel, when the test only includes an intercept. The panel unit root tests performed on the entire country set and on the set of countries included in the application are reported in table 1 in the appendix.

As already referred to, the IPS test assumes cross-country independence. Nevertheless being the panel unit root test, which performs better comparatively with other tests with the same assumption, as already pointed out in sub-section 2.2., other panel unit root test techniques are more adequate when cross-section dependence is present.

The existence of cross-section dependencies is shown using a Granger Causality Test with 2 lags, the results of which are presented in Table 2 of the appendix. This Table reports the results of the F-statistic that are statistically significant, meaning that time series expressed in line Granger cause the time series expressed in column. As can be observed, the panel series present a causality relation across some of the countries, thus violating the independence assumption required by the IPS panel unit root test.

In this sense, Chang's non-linear IV unit root tests are performed, in each cross-section. The number of lags p_i included in individual regressions is once again chosen according to the *BIC Criterion* with a maximum length of 8 periods. An analysis of the residual serial correlation LM test is also performed with a lag span of three. As can be observed in table 1, the unit root null of Chang's test can be rejected when considering the panel as a whole.

Finally the *SURADF* unit root test is performed. This procedure makes use of the panel data setting and uses seemingly unrelated regressions, but performs separate unit root tests in each panel unit, as previously described. This procedure has the advantage of identifying which cross-sections are stationary and which cross-sections have unit roots. This analysis is interesting in the sense that in a panel the rejection of a unit root does not mean that all units are stationary.

After performing the *SURADF* unit root test it can be observed that the panel data are composed of a mix between stationary and nonstationary time series, where the bold value tests indicate the cases for which the null hypothesis are not rejected.

The present results (vide Table 1) indicate the presence of a unit root puzzle in the cross sections that compose the panel. As indicated, the rejection of the null of a panel unit root does not imply that all time series are stationary. More, the rule appears to be a mix between stationary and non-stationary time series when the presence of cross sectional dependence is taken into account.

Further, since cointegration tests are derived under the assumption of cross sectional independence across individual units, which seem not to be appropriate here as well as for the most macroeconomic data. Indeed cross-sectional dependency seems quite apparent for many economic panel data we encounter. As such, assuming the independence across cross-sectional units is quite restrictive.

6. Concluding remarks

Testing for unit roots in time series is common practice among empirical studies. However, testing for unit roots in panels is a recent process, with major developments in nonstationary panel models originating in mid-1990s. Recent attention has been given to panel data issues arising from numerous time series procedures applied to panels.

Levin and Lin (1992, 1993), Im, Pesaran and Shin (2003) and Maddala and Wu (1999), are authoritative references of panel unit root tests that rely on cross sectional independence.

To deal with the presence of cross-sectional dependency, Chang (2002) proposed a panel unit root test based on non-linear instrumental variable (IV) estimation of the usual Augmented Dickey-Fuller (ADF) type regression for each cross-sectional unit, using as instruments non-linear transformations of the lagged levels. The test statistic is defined as an average of individual IV t-ratios, which is asymptotically normal, and does not require the tabulation of critical values.

In this paper, we apply the Im, Pesaran and Shin test, and the Chang panel unit root tests. The panel unit root tests that we employ are joint hypothesis tests in the sense that all units of a panel contain a unit root. When the joint null hypothesis is rejected it is possible that one or a few time series in the panel contribute to this finding. Cumulatively, given that these tests allow the autoregressive parameter to differ across cross sections under the alternative, then the rejection of the null hypothesis means that not all units of the panel contain a unit root. Effectively, a mixture of stationary and nonstationary time series can cohabit in the same panel data.

Given the limitations associated with the previous panel unit root tests, we use an alternative test that allows for the presence of contemporaneous cross-correlation and heterogeneous serial correlation of the regression residuals as suggested by Breuer, McNown and Wallace (1999).

The main advantage of the BNW test is that it allows us to determine which cross sectional series reject the null hypothesis of a unit root and which do not. The BNW test uses a SUR framework for testing each panel unit under the null and the alternative hypothesis, exploiting the information in the error covariances to produce efficient estimators and potentially more powerful test statistics.

We perform panel unit root test on a panel data series of 15 EU member countries using data relative to population weight over 65 years of age for a 32 years span, comprising the period from 1970 to 2001. Our panel unit root test results indicate the rejection of the null of a panel unit root by IPS or Chang's tests. Such results do not imply however that all time series are stationary. Moreover, when we look at the BNW test the rule appears to be a mix between stationary and non-stationary time series when the presence of cross sectional dependence is taken in to account.

Our analysis confirm that the power of panel unit root tests based on individual unit root tests such as IPS and Chang's tests can be distorted under small samples, influencing cointegration test and panel data estimations results.

References

Bac, Catherine and Yannick le Pen (2002), "An international comparison of health care expenditure determinants", Paper presented at the *10th International Conference on Panel Data*, Berlin, July 5-6

Baltagi, Badi H. and Chihwa Kao (2000), "Nonstationary panels, cointegration in panels and dynamic panels: A survey", Center for Policy Research, *Working Paper No16*

Banerjee, Anindya, Massimiliano Marcellino and Chiara Osbat (2001-a), "Some cautions on the use of panel methods for integrated series of macro-economic data", European University Institute, *Discussion Paper No00/20*

Banerjee, Anindya, Massimiliano Marcellino and Chiara Osbat (2001-b), "Testing for PPP: Should we use panel methods?", Paper presented at *The Institute for Fiscal Studies Annual Conference*, 12 July

Bayoumi, Tamim and Ronald MacDonald (1999), "Deviations of exchange rates from purchasing power parity: A story featuring two monetary unions", *IMF Staff Papers*, Vol.46, Issue1, pp.89-102

Bornhorst, Fabian (2002), "A note on the use of Panel Unit Root Tests on sectional dependent data: An application to PPP", Paper presented at the *VIIth Spring Meeting of Young Economists*, Université Paris 1, Paris, April 18-20

Breitung, Jorg (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, pp.161-178, JAI: Amsterdam

Breuer, Janice Boucher, Robert McNown and Myles Wallace (1999), "Series-specific tests for a unit root in a panel setting with an application to real exchange rates", Center for Economic Analysis, Department of Economics, University of Colorado at Boulder, *Working Paper No99-09*

Cerrato, Mario (2002), "The cross sectional dependence puzzle", Paper presented at the *Royal Econometric Society Annual Conference*

Chang, Yoosoon (2002), "Nonlinear IV unit root tests in panels with cross-sectional dependency", *Journal of Econometrics*, Vol.110, No2, pp. 261-292

Chang, Yoosoon (2004), "Bootstrap unit root tests in panels with cross-sectional dependency", *Journal of Econometrics*, Vol. 120, Issue 2, pp. 263-293

Chang, Yoosoon and Wonho Song (2002), "Panel unit root tests in the presence of cross-sectional dependency and heterogeneity", Paper presented at *the 10th International Conference on Panel Data*, Berlin, July 5-6

Choi, Chi-Young (2000), "Confirmatory analysis in panel unit root tests", Paper presented at the *MEG Conference*, University of Chicago, 9 October

Choi, In (2002), "Combination unit root tests for cross-sectionally correlated panels", *Mimeo*, Available at: http://ihome.ust.hk/~inchoi/InChoiPhilVol.pdf

Elliot, G., T. Rothenberg and J. Stock (1996), "Efficient tests for an autoregressive unit root", *Econometrica*, Vol.64, pp.813-836

Granger, Clive W.J., (1969), "Investigating Causal Relations by Econometric Models and Cross Spectral Methods", *Econometrica*, Vol.37, pp.424-438.

Hadri, Haddour (2000), "Testing for stationarity in heterogeneous panel data", *Econometrics Journal*, Vol.3, pp.148-161

Im, Kyung So, M. Hashem Pesaran and Yongcheol Shin (1997), "Testing for unit roots in heterogeneous panels", Department of Applied Economics, University of Cambridge, *Working Paper*

Im, Kyung So, M. Hashem Pesaran and Yongcheol Shin (2003), "Testing for unit roots in heterogeneous panels", *Journal of Econometrics*, Vol.115, No1, pp.53-74

Johansen, S. (1995), *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press, Oxford and New York

Kao, Chihwa (1999), "Spurious regression and residual-based tests for cointegration in panel data", *Journal of Econometrics*, Vol.90, pp.1-44

Kao, Chihwa and Min-Hsien Chiang (2000), “On the estimation and inference of a cointegrated regression in panel data”, in B. Baltagi and C. D. Kao (eds), *Nonstationary Panels, Panel Cointegration and Dynamic Panels*, Advances in Econometrics, Vol.15, Elsevier Science, New York, pp.179-222

Kao, Chihwa, Min-Hsien Chiang and Bangtian Chen (1999), “International R&D spillovers: An application of estimation and inference in panel cointegration”, *Oxford Bulletin of Economics and Statistics*, Vol.61, Issue4, pp.693-711

Karlsson, Sune and Mickael Lothgren (1999), “On the power and interpretation of panel unit root tests”, *Economics Letters,* Vol.66, pp.249-255

Kónya, László (2001), “Panel data unit root tests with an application”, Central European University*, Working Paper No2/2001*

Larsson, Rolf and Johan Lyhagen (1999), “Likelihood-based inference in multivariate panel cointegration models”, Stockholm School of Economics, *Working Paper Series in Economics and Finance No331*

Larsson, Rolf and Johan Lyhagen (2000), “Testing for common cointegrating rank in dynamic panels”, Stockholm School of Economics, *Working Paper Series in Economics and Finance No378*

Larsson, Rolf, Johan Lyhagen and Mickael Lothgren (2001), “Likelihood-based cointegration tests in heterogeneous panels”, *Econometrics Journal,* Vol.4, pp.109-142

Levin, Andrew and Chien-Fu Lin (1992), “Unit root tests in panel data: Asymptotic and finite-sample properties”, Department of Economics, University of California - San Diego, *Discussion Paper No92-23*

Levin, Andrew and Chien-Fu Lin (1993), “Unit root tests in panel data: New results”, Department of Economics, University of California - San Diego, *Discussion Paper No93-56*

Levin, Andrew, Chien-Fu Lin and Chia-Shang James Chu (2002), “Unit root tests in panel data: Asymptotic and finite-sample properties”, *Journal of Econometrics*, Vol.108, Issue1, pp.1-24

Maddala, G. S. and In-Moo Kim (1998), *Unit roots, cointegration and structural change,* Cambridge University Press

Maddala, G. S. and Shaowen Wu (1999), “A comparative study of unit root tests with panel data and a new simple test”, *Oxford Bulletin of Economics & Statistics,* Vol.61, pp.631-652

McCoskey, Suzanne and Chihwa Kao (1998), “A residual-based test of the null of cointegration in panel data”, *Econometric Reviews,* Vol.17, pp.57-84

McCoskey, Suzanne and Chihwa Kao (1999), "A Monte Carlo comparison of tests for cointegration in panel data", Maxwell School of Citizenship and Public Affairs, Syracuse University, Center for Policy Research, *Working Paper No3*

Ng, Serena and Pierre Perron (2001), "Lag length selection and the construction of unit root tests with good size and power", *Econometrica*, Vol.69, pp.1519-1554

Ng, Serena and Pierre Perron (2005), "A note on the selection of time series models", *Oxford Bulletin of Economics & Statistics*, Vol.67, Issue1, pp.115-134

Nickell, Stephen (1981), "Biases in dynamic models with fixed effects", *Econometrica*, Vol.49, pp.1399-1413

Pedroni, Peter (1999), "Critical values for cointegration tests in heterogeneous panels with multiple regressors", *Oxford Bulletin of Economics & Statistics*, Vol.61, pp.653-670

Pedroni, Peter (2000), "Fully modified OLS for heterogeneous cointegrated panels" in B. Baltagi and C. D. Kao (eds), *Nonstationary Panels, Panel Cointegration and Dynamic Panels*, Advances in Econometrics, Vol.15, Elsevier Science, New York, pp.93-120

Pesaran, M. Hashem, and Ron P. Smith (1995), "Estimating long-run relationships from dynamic heterogeneous panels", *Journal of Econometrics*, Vol.68, pp.79-113

Phillips, Peter C. B. and Hyungsik R. Moon (1999), "Linear regression limit theory for nonstationary panel data", *Econometrica*, Vol.67, pp.1057-1111

Phillips, Peter C. B. and Zhijie Xiao (1998), "A primer on unit root testing", *Journal of Economic Surveys*, Vol.12, Issue5, pp.423-470

Quah, Danny (1994), "Exploiting cross-section variations for unit root inference in dynamic data", *Economics Letters*, Vol.44, pp.9-19

Stock, James H. (1994), "Unit roots, structural breaks and trends" in R. F. Engle and D. L. McFadden (eds), *Handbook of Econometrics*, Vol.4, Elsevier Science

Wagner, Gary A. (2003), "Are state budget stabilization funds only the illusion of savings?", *The Quarterly Review of Economics and Finance*, Vol.43, Issue2, pp.213-238

Appendix

Table 1 - Estimated Panel Data Unit Root Tests for POP65 (1970-2001)

	IPS PANEL TEST with intercept			IPS PANEL TEST with intercept and trend			CHANG PANEL TEST			SURADF(15)		
	ADF test	Lags	LM TEST(3)	ADF test	Lags	LM TEST(3)	Zi	Lags	LM TEST(3)	Crit.V.(5%)	Test	Ho Validation
AUS	-3,3272	1	1,1433	-2,9282	2	2,7646	-2,7053	1	2,1216	-4,6775	-5,0960	Rejects Ho
BEL	0,3127	1	1,6770	-1,4226	1	1,9490	-1,0839	3	1,8653	-5,1541	**-0,5743**	**Accepts Ho**
DNK	-2,4937	2	3,1606	-2,2517	3	2,7859	-1,1377	2	2,7017	-2,7767	-4,5342	Rejects Ho
FIN	-1,9189	2	2,2980	-4,5134	4	5,4346	-0,9445	2	1,3896	-4,2463	-5,4989	Rejects Ho
FRA	-1,1296	1	2,2104	-2,5040	1	1,0580	-1,2560	3	1,4544	-4,7936	**-1,9909**	**Accepts Ho**
GER	-1,1240	1	4,3644	-2,2043	1	4,7453	-1,6204	1	3,6280	-4,2416	**-1,0978**	**Accepts Ho**
GRE	2,0585	8	5,0464	-3,4269	5	2,6565	-3,1464	5	1,1555	-2,8834	**1,9628**	**Accepts Ho**
IRE	-0,5640	0	4,7539	-3,2784	4	2,3701	-1,1157	0	1,4924	-3,8281	**-0,1448**	**Accepts Ho**
ITA	1,0919	1	3,4319	-1,0916	1	2,6432	-0,8571	1	1,6674	-4,7692	**0,4972**	**Accepts Ho**
LUX	-1,4819	1	2,1310	-2,9387	2	2,0884	-0,4403	1	1,3308	-4,1037	**-2,4374**	**Accepts Ho**
NLD	-1,9911	1	1,5122	-0,5666	1	1,8359	-0,4859	2	2,3334	-4,1595	-4,2113	Rejects Ho
POR	0,8640	6	4,2388	-2,4912	6	4,1622	-1,8751	6	5,0376	-4,0610	**-0,2818**	**Accepts Ho**
SPA	4,4199	1	1,4073	-0,7711	1	2,1748	0,0346	0	2,0444	-3,3105	**4,7699**	**Accepts Ho**
SWE	-2,4295	2	1,2292	-0,6576	1	2,8675	-1,1276	2	1,3630	-2,9126	-4,0257	Rejects Ho
UK	-2,5747	1	1,2772	-1,2834	1	1,6508	-0,2371	2	1,4456	-4,1952	-4,5171	Rejects Ho
PANEL(15)	12,7724	**Accepts Ho**		-10,9899	**Rejects Ho**		-4,6472	**Rejects Ho**				
Crit. Value(5%)	-1,9			-2,53			-2,5760					

Table 2 - Granger Causality (Lag=2)

	POP65FIN	POP65GRE	POP65DNK	POP65ITA	POP65FRA	POP65GER	POP65SWE
POP65AUS	6,1848						
POP65BEL		24,4922					
POP65IRE		18,0716	6,98005				25,2543
POP65FRA	8,8372	6,52867		22,1279			
POP65GER	7,97855			7,86161			
POP65GRE	6,53625						
POP65SPA				7,98417	7,4907	6,01195	
POP65NLD						6,46071	