Università degli Studi del Molise
Dipartimento di Scienze Economiche, Gestionali e Sociali (SEGeS)

ECONOMICS & STATISTICS DISCUSSION PAPER
No. 063/11

Panel-CADF Testing with R: Panel Unit Root Tests Made Easy

Claudio Lupi

Panel-CADF Testing with R: Panel Unit Root Tests Made Easy

Claudio Lupi
University of Molise

Abstract

This paper presents the R implementation of the panel covariate augmented Dickey-Fuller (panel-CADF) test proposed in Costantini and Lupi (2011), as well as the implementation of the tests advocated in Choi (2001) and Demetrescu, Hassler, and Tarcolea (2006). A panel-CADF extension of the test suggested in Hanck (2008) is also discussed and its size and power properties are investigated via Monte Carlo analysis. The simulation results show that the panel-CADF tests have interesting properties in terms of size and power. The R implementation illustrated here is part of the ongoing work on a new R package named **punitroots** (Kleiber and Lupi 2011).

Keywords: panel data, unit root, R.

1. Introduction

R is rapidly gaining the favour of econometricians and applied economists and recent dedicated packages are now available. To get an idea of the increasing interest surrounding R, it is sufficient to look at the CRAN Task Views "Econometrics", "Finance", "Social Sciences" and "Time Series" that include most of the econometrics-related packages.[1] Given the rising interest, in 2008 the *Journal of Statistical Software* dedicated a special volume to "Econometrics in R" (Zeileis and Koenker 2008) and recent economics and econometrics conferences often include special sessions with the same title.

Panel unit root testing is a rather specialized area of econometrics that has become increasingly popular among applied (macro-)economists. In fact, since the mid-nineties panel unit root tests have attracted considerable attention on the part of macroeconomists because of their good power properties and because they can help solving some interesting theoretical problems. At the time of writing (December 2011) a quick search of the key "panel unit root" using the Scopus bibliographical data base returned 434 titles over the period 1996–2011, with an increasing publication pace especially in the last few years (see Figure 1). Excellent surveys of the literature are contained in Choi (2006), Breitung and Pesaran (2008), and Banerjee and Wagner (2009).

Despite the increasing favour in the applied economic literature, panel unit root tests implementations are not widespread and R includes only a few of them. In fact, only two R packages, namely **plm** (Croissant and Millo 2008) and the concept package **punitroots** (Kleiber and Lupi 2011), implement panel unit root testing procedures in R: the functions discussed

[1]CRAN task views are visible at `http://cran.r-project.org/web/views/`.

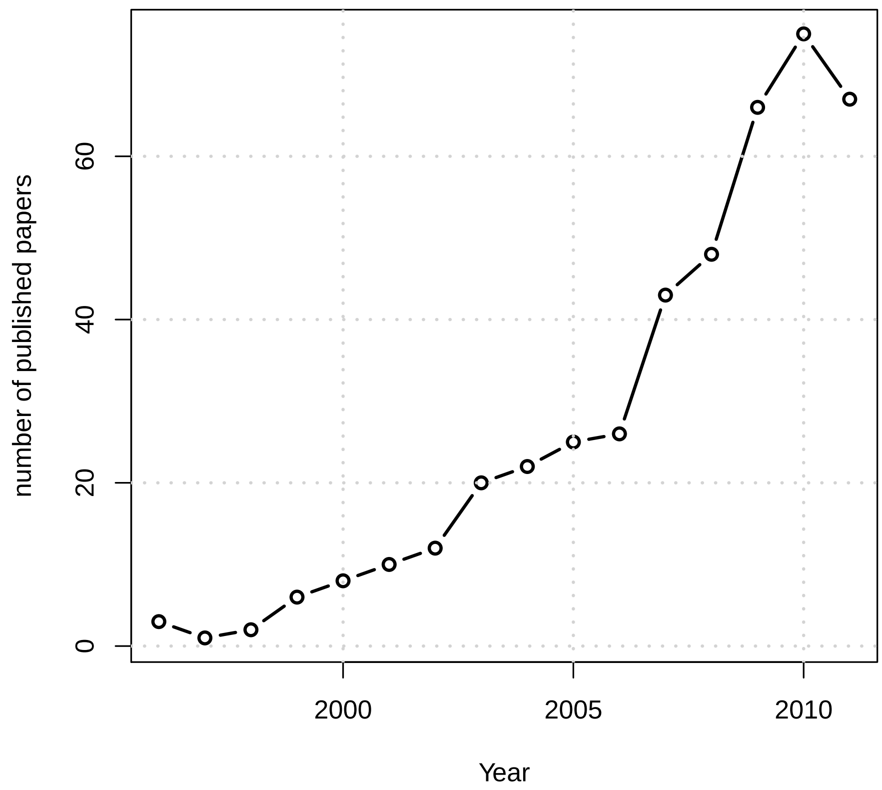

Figure 1: Published papers dealing with panel unit roots (data source: Scopus).

in this paper are all included in the **punitroots** package.[2]

This paper is mainly intended for an audience of practitioners and complements Costantini and Lupi (2011) in two main respects: first, by investigating the performance of a new panel covariate augmented Dickey-Fuller test based on Hansen (1995) and Hanck (2008) and second, by offering to practitioners a practical guide on how to use the functions included in **punitroots** to carry out panel covariate augmented Dickey-Fuller tests as well as other popular panel unit root tests (Choi 2001; Demetrescu *et al.* 2006).

The rest of the paper is organized as follows: in order for the paper to be self-contained, the next Section briefly introduces panel p value combination tests and Section 3 offers a concise review of covariate augmented Dickey-Fuller (CADF) tests. The following Section deals with a Monte Carlo analysis aiming at evaluating the finite sample properties of the panel-CADF tests. For the sake of illustration of the relevant R functions, an empirical application is carried out in Section 5. The final Section concludes.

[2]Development versions of **punitroots** are available from R-Forge at `https://r-forge.r-project.org/projects/punitroots/`.

2. p value combination tests

Maddala and Wu (1999) and Choi (2001) independently propose to solve the problem of panel unit root testing using p value combination tests. The null hypothesis is that all of the series in the panel are $I(1)$. The alternative hypothesis is that *at least one* of the series is $I(0)$. The tests proposed in Maddala and Wu (1999) and Choi (2001) are based on the idea that the p values from N independent augmented Dickey-Fuller (ADF, Said and Dickey 1984) tests can easily be combined to obtain a test on the joint hypothesis concerning all the $N > 1$ units. In both papers it is highlighted that under the null the p values p_i (with $i = 1, \ldots, N$) are independent $\mathrm{U}_{(0,1)}$ variables so that $-2 \log p_i \sim \chi^2(2)$. Therefore, for fixed N, as $T \to \infty$, under the null

$$P := -2 \sum_{i=1}^{N} \log p_i \xrightarrow{d} \chi^2(2N) \,. \tag{1}$$

Choi (2001) considers also different p values combination tests and suggests that the inverse-normal combination test based on the fact that under the null

$$Z := \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \Phi^{-1}(p_i) \xrightarrow{d} \mathrm{N}(0,1) \tag{2}$$

has the best overall performance, where convergence again holds for fixed N and $T \to \infty$.

The advantages of the p value combination approach derive from its simplicity, the flexibility in specifying a different model for each panel unit, the ease in allowing the use of unbalanced panels, the possibility of using any unit root test, and the fact that the convergence results are proved using (fixed-N) T-asymptotics.

However, the assumption that the panel units are cross-sectionally independent is very restrictive. For this reason, building upon Hartung (1999), Demetrescu *et al.* (2006) propose a modification of Choi's inverse-normal combination test that can be used when the N p values are not independent. In particular, Hartung (1999) shows that if the *probits* $\Phi^{-1}(p_i)$ are correlated with common correlation ϱ, then under the null

$$Z_{\mathrm{H}} := \frac{1}{\sqrt{N\left(1 + \varrho(N-1)\right)}} \sum_{i=1}^{N} \Phi^{-1}(p_i) \sim \mathrm{N}(0,1) \,. \tag{3}$$

The actual modification proposed by Hartung (1999) and considered in Demetrescu *et al.* (2006) is slightly more complicated to allow for the fact that ϱ is unknown. A common practical implementation, used by Demetrescu *et al.* (2006) in their simulations, is[3]

$$\widehat{Z_{\mathrm{H}}} := \frac{\sum_{i=1}^{N} \Phi^{-1}(p_i)}{\left\{ N \left[1 + \left(\hat{\varrho}^* + 0.2 \sqrt{\frac{2}{N+1}} \left(1 - \hat{\varrho}^*\right) \right) (N-1) \right] \right\}^{\frac{1}{2}}} \tag{4}$$

where $\hat{\varrho}^*$ is a consistent estimator of ϱ such that $\hat{\varrho}^* = \max\{-1/(N-1), \hat{\varrho}\}$ with

$$\hat{\varrho} = 1 - (N-1)^{-1} \sum_{i=1}^{N} \left(\Phi^{-1}(p_i) - N^{-1} \sum_{i=1}^{N} \Phi^{-1}(p_i) \right)^2 \,. \tag{5}$$

[3] More general formulations can be applied that allow for unequal weighting of the p values and more accurate control of the significance level. See Hartung (1999, p. 851).

A rather different viewpoint has been suggested by Hanck (2008). In particular, Hanck (2008) observes that the problem of panel unit root testing can be recast in terms of a multiple testing problem, where the *complete null hypothesis* is again that all of the series are $I(1)$, against the alternative hypothesis that at least one series is $I(0)$. As is well known (see e.g. Shaffer 1995) the complete null cannot be rejected simply on the basis that $\min(p_i) < \alpha$ $(i = 1, \ldots, N)$ for a pre-specified level α, because such a procedure would result in a test having a size much larger than α. In fact, Simes (1986) shows that if a set of N hypotheses $\mathrm{H}_{0,1}, \ldots, \mathrm{H}_{0,N}$ are all true, and the associated test statistics are independent, then $\Pr\left(p_{(i)} > i\alpha/N\right) = 1 - \alpha$, where the $p_{(i)}$'s are the ordered p values such that $p_{(1)} \leqslant p_{(2)} \leqslant \ldots \leqslant p_{(N)}$. Furthermore, Sarkar and Chang (1997) show that Simes' equality holds also in the presence of positively dependent test statistics.[4] Therefore, Hanck (2008) suggests that the panel unit root null hypothesis can be tested easily using the intersection test presented in Simes (1986). The test is very simple to compute: denote by $p_{(i)}$ the ordered sequence of the N p values of each unit root test on each individual series. Given a pre-specified significance level α, the null is rejected if $p_{(i)} \leqslant i\alpha/N$ for any $i = 1, \ldots, N$.

As with Choi (2001) and Demetrescu *et al.* (2006), the original proposal contained in Hanck (2008) is to base the test on p values derived from individual standard ADF tests. However, in all of these procedures there is no limitation to the individual tests that can be employed: therefore here we will apply the same principles to p values obtained from individual covariate augmented Dickey-Fuller (CADF) tests of the form suggested in Hansen (1995), in this way obtaining the panel-CADF tests. While the CADF extension of Choi (2001) and Demetrescu *et al.* (2006) has already been analysed in detail in Costantini and Lupi (2011), the CADF extension of Hanck (2008) is new.

3. The panel covariate augmented Dickey-Fuller tests

The panel covariate Dickey-Fuller tests considered here are simple panel extensions, based on the p value combination methods outlined in Section 2, of the covariate augmented Dickey-Fuller (CADF) test advocated in Hansen (1995).

Hansen (1995) proves that significant power gains in unit root testing can be achieved if stationary covariates are included in otherwise conventional augmented Dickey-Fuller (ADF, Said and Dickey 1984) tests. The basic intuition behind Hansen's test is quite simple. If we want to test for the presence of a unit root in y_t and a stationary covariate Δx_t exists that is linearly related to Δy_t, then adding Δx_t to the ADF regression for y_t will increase the precision of the estimates and therefore the power of the test. Of course, although the intuition is simple, the details are much more complicated: in particular, Hansen (1995) proves that the resulting unit root test statistic under the null is no longer distributed according to a Dickey-Fuller distribution, but is instead distributed according to a weighted sum of a Dickey-Fuller and a standard normal distribution, where the weights are functions of a nuisance parameter.

To be more specific, consider

$$a(L)\Delta y_t = \delta y_{t-1} + v_t \tag{6}$$

$$v_t = b(L)\left(\Delta x_t - \mu_x\right) + e_t \tag{7}$$

where $a(L) := (1 - a_1 L - a_2 L^2 - \ldots - a_p L^p)$ is a polynomial in the lag operator L, $\Delta x_t \sim I(0)$,

[4] To be more precise, the property of total positivity of order 2 (TP2) must be satisfied.

$\mu_x := \mathsf{E}(\Delta x)$, $b(L) := (b_{q_2} L^{-q_2} + \ldots + b_{q_1} L^{q_1})$ is a polynomial where both leads and lags are allowed.[5] Furthermore, consider the long-run covariance matrix

$$\boldsymbol{\Omega} := \sum_{k=-\infty}^{\infty} \mathsf{E}\left[\begin{pmatrix} v_t \\ e_t \end{pmatrix} \begin{pmatrix} v_{t-k} & e_{t-k} \end{pmatrix}\right] = \begin{pmatrix} \omega_v^2 & \omega_{ve} \\ \omega_{ve} & \omega_e^2 \end{pmatrix} \tag{8}$$

and define the long-run squared correlation between v_t and e_t as

$$\rho^2 := \frac{\omega_{ve}^2}{\omega_v^2 \, \omega_e^2} \, . \tag{9}$$

The testing equation is very similar to the ordinary ADF equation:

$$a(L)\Delta y_t = \delta y_{t-1} + b(L)\Delta x_{t-1} + e_t \, . \tag{10}$$

As with the ADF test, also the CADF test can be augmented with a constant and a trend.

Hansen (1995) shows that if some regularity conditions are satisfied, under the unit root null the t ratio for the coefficient δ in (10) is such that

$$\widehat{t(\delta)} \xrightarrow{w} \rho \frac{\int_0^1 W \, dW}{\left(\int_0^1 W^2\right)^{1/2}} + \left(1 - \rho^2\right)^{1/2} \mathrm{N}(0,1) \tag{11}$$

where W is a standard Wiener process and $\mathrm{N}(0,1)$ is a standard normal independent of W. When a model with a constant or with a constant and a linear trend is used, W is replaced by a demeaned or a detrended Wiener process, respectively (see Hansen 1995, for details).

Once the problem of computing the p values from (11) is solved, then it becomes easy to apply p value combination methods to derive a panel covariate augmented Dickey-Fuller test. In fact, this is the idea followed by Costantini and Lupi (2011) who propose a panel covariate augmented Dickey-Fuller test (that they label pCADF) and also present a method to compute the asymptotic p values.[6] The p value combination suggested by Costantini and Lupi (2011) follows Choi (2001) when no cross-dependence is detected and Demetrescu *et al.* (2006) in the presence of cross-dependence. As far as the choice of the covariates is concerned, Costantini and Lupi (2011) suggest using observed series that can be theoretically related to the series of interest as well as covariates derived on the basis of statistical reasoning. In particular, (see Costantini and Lupi 2011) suggest using as the stationary covariate for each variable to be tested the average of the first difference of the other series in the panel; as an alternative, the difference of the first principal component among the series under investigation can be used as well. The latter procedure aims at extracting an underlying nonstationary common factor among the observed series, and use its first differences as the stationary covariate. Of course, in this case the panel CADF test refers explicitly to cross-dependent time series. However, in general given that different stationary covariates can be selected for each series, the method can be applied also to panels made of independent units.

The use of Hansen's CADF test, instead of the conventional ADF test, ensures that the panel test has better power properties. Furthermore, in Costantini and Lupi (2011), contrary to

[5] More general formulations allow for deterministic terms and vector valued covariates Δx_t.

[6] The R implementation of the computation of the p values of the asymptotic distribution (11) is described in detail in Lupi (2009).

Demetrescu *et al.* (2006), Hartung's procedure for cross-correlation correction is applied only when the p value of the cross-correlation test advocated by Pesaran (2004) is lower than a pre-specified threshold whose default value is set to 0.10.

In addition to the panel covariate augmented Dickey-Fuller test advocated in Costantini and Lupi (2011), here we consider also the extension to the CADF tests of the ADF-based test suggested in Hanck (2008). Given its relation to Simes' procedure, here we label the latter test as sCADF.

4. Monte Carlo analysis

In this Section we carry out a small Monte Carlo analysis to compare the size and power properties of the the different unit root tests implemented (as of today) in **punitroots**. We use a slightly simplified version of a data generating process (DGP) that was originally proposed in Costantini and Lupi (2011). This DGP encompasses other DGPs commonly used in panel unit root Monte Carlo analyses. For brevity we will consider only three paradigmatic experiments. In the first, all the series are cross-sectionally independent. Given that cross-section independence can hardly be seen as a realistic feature of macro-panels, in the other two experiments the time series are generate more realistically in such a way that they are cross-sectionally dependent.

The behaviour of four tests is compared in each experiment, namely Choi (2001) (Choi), Demetrescu *et al.* (2006) (DHT), three versions of the pCADF test advocated in Costantini and Lupi (2011),[7] and four versions of the Simes-based test suggested in Hanck (2008).[8] All these tests share the same null and alternative hypotheses.

4.1. The Data Generating Process

The Data Generating Process (DGP) used in the simulations can be represented as

$$\Delta \boldsymbol{y}_t = \boldsymbol{D}\boldsymbol{y}_{t-1} + \boldsymbol{u}_t \tag{12}$$

$$\begin{pmatrix} \boldsymbol{u}_t \\ \xi_t \end{pmatrix} = \begin{pmatrix} \boldsymbol{B} & \boldsymbol{\gamma} \\ \boldsymbol{0}' & \lambda \end{pmatrix} \begin{pmatrix} \boldsymbol{u}_{t-1} \\ \xi_{t-1} \end{pmatrix} + \begin{pmatrix} \boldsymbol{\eta}_t \\ \varepsilon_t \end{pmatrix} \tag{13}$$

$$\begin{pmatrix} \boldsymbol{\eta}_t \\ \varepsilon_t \end{pmatrix} \sim \mathrm{N}\left(\boldsymbol{0}, \boldsymbol{\Sigma}\right) = \mathrm{N}\left[\begin{pmatrix} \boldsymbol{0} \\ 0 \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\sigma}_{12} \\ \boldsymbol{\sigma}'_{12} & \sigma_{22} \end{pmatrix}\right] \tag{14}$$

where Δ is the usual difference operator, $\boldsymbol{y}_t = (y_{1t}, \ldots, y_{Nt})'$, $\boldsymbol{u}_t = (u_{1t}, \ldots, u_{Nt})'$, $\boldsymbol{D} = \mathrm{diag}(\delta_1, \ldots, \delta_N)$, $\boldsymbol{B} = \mathrm{diag}(\beta_1, \ldots, \beta_N)$, $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_N)'$ and $\boldsymbol{\eta}_t = (\eta_{1t}, \ldots, \eta_{Nt})'$. Under the unit root null we have $\delta_i = 0\ \forall i$, while under the alternative $\delta_i < 0$ for some or for all the i's.[9]

[7] The three version differ in the way the stationary covariate is selected. The first (pCADF) consider the case where the "correct" stationary covariate (or a good proxy of it) is used. The second (pCADF.PC) assumes that the panel is balanced (it is transformed in a balanced panel in the case it is not) and utilises the differenced first principal component of the N series as the stationary covariate. The last (pCADF.DY) is again valid for a balanced panel and for each series takes the difference of the average of the other series as the stationary covariate.

[8] The four variants differ in the test they are based upon. The first (sADF) is based on the p values of standard ADF tests, as in Hanck (2008). The others (sCADF, sCADF.PC, and sCADF.DY) are based on the p values of CADF tests, with the stationary covariates selected as above, and are suggested here for the first time.

[9] A detailed discussion of the properties of this DGP can be found in Costantini and Lupi (2011).

In the Monte Carlo we use $(\boldsymbol{\eta}_t', \epsilon_t)' \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma})$, with

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\sigma}_{12} \\ \boldsymbol{\sigma}_{12}' & \sigma_{22} \end{pmatrix} \tag{15}$$

and $\mathrm{diag}(\boldsymbol{\Sigma}) = \boldsymbol{\imath}$. $\lambda = 0.2$ and $\beta_i \sim \mathrm{U}_{[0.2,0.4]}$ in all the experiments.

The key design parameters are $\boldsymbol{D}$, $\boldsymbol{\gamma}$, $\boldsymbol{\Sigma}_{11}$ and $\boldsymbol{\sigma}_{12}$. $\boldsymbol{D}$ defines which of the series are $I(1)$ or $I(0)$; under the null $\boldsymbol{D} = \mathrm{diag}(\mathbf{0})$, while under the alternative $\delta_i \sim \mathrm{U}_{[-0.09,-0.01]}$ (which means an autoregressive parameter equal to 0.95 on average). In particular, the power of the tests is examined when the fraction of stationary series (on the total number of series in the panel) is equal to 0.2 and 0.5. $\boldsymbol{\gamma}$, $\boldsymbol{\Sigma}_{11}$ and $\boldsymbol{\sigma}_{12}$ determine the presence and the strength of cross-correlations among the panel units.

The design parameters used in the three experiments are summarized in Table 1.

Experiment	$\boldsymbol{\gamma}$	$\boldsymbol{\sigma}_{12}$	$\boldsymbol{\Sigma}_{11}$
1	$\mathbf{0}$	$\mathbf{0}$	$\boldsymbol{I}$
2	$\mathbf{0}$	$(\boldsymbol{\sigma}_{12})_i \sim \mathrm{U}_{[0.7,0.9]}$	$(\boldsymbol{\Sigma}_{11})_{ij} = (\boldsymbol{\sigma}_{12})_i \times (\boldsymbol{\sigma}_{12})_j$
3	$\gamma_i \sim \mathrm{U}_{[0.7,0.9]}$	$(\boldsymbol{\sigma}_{12})_i \sim \mathrm{U}_{[0.7,0.9]}$	$(\boldsymbol{\Sigma}_{11})_{ij} = (\boldsymbol{\sigma}_{12})_i \times (\boldsymbol{\sigma}_{12})_j$

Table 1: Values of the key design parameters in the three experiments.

In experiment 1 all the units are cross-sectionally independent. This is the setting under which the "first generation tests" proposed in Maddala and Wu (1999); Choi (2001); Levin, Lin, and Chu (2002); Im, Pesaran, and Shin (2003) have been developed. In experiment 2, innovations are cross-correlated and the y's are no longer cross-sectionally independent. Furthermore, although $\boldsymbol{\gamma} = \mathbf{0}$, the correlation between $\boldsymbol{\eta}_t$ and ϵ_t breaks down weak exogeneity of ξ_t (Engle, Hendry, and Richard 1983). In experiment 3, besides the innovations being cross-correlated, the units share a stationary common factor ξ_t which increases cross-dependence among the series of interest. Imposing $(\boldsymbol{\Sigma}_{11})_{ij} = (\boldsymbol{\sigma}_{12})_i \times (\boldsymbol{\sigma}_{12})_j$ $(i \neq j)$ in experiments 2 and 3 ensures that $\boldsymbol{\Sigma}_{11}$ and $\boldsymbol{\sigma}_{12}$ are logically consistent and that $\boldsymbol{\Sigma}$ in (15) is positive definite (see Costantini and Lupi 2011).

In order to use values that can be easily encountered in practical applications, in each experiment $N = 10$ time series units are generated over $T = 100$ time periods. All experiments are carried out using 2,000 replications. Lag selection in the ADF and CADF models is operated via the BIC.

4.2. Simulation results

Comparing the size and the power of different tests can sometimes be difficult because of the simulation errors associated with estimating rejection probabilities and because of the difficulty of comparing power across tests having different actual size. In order to take into account both these aspects and to allow for direct comparison of power across the tests, in this paper we compare the tests on the basis of their intrinsic power, as proposed in Lloyd (2005).

Simulation results are reported graphically in Figures 2 to 4. For ease of comparison, all the graphs are reported on the same scale.

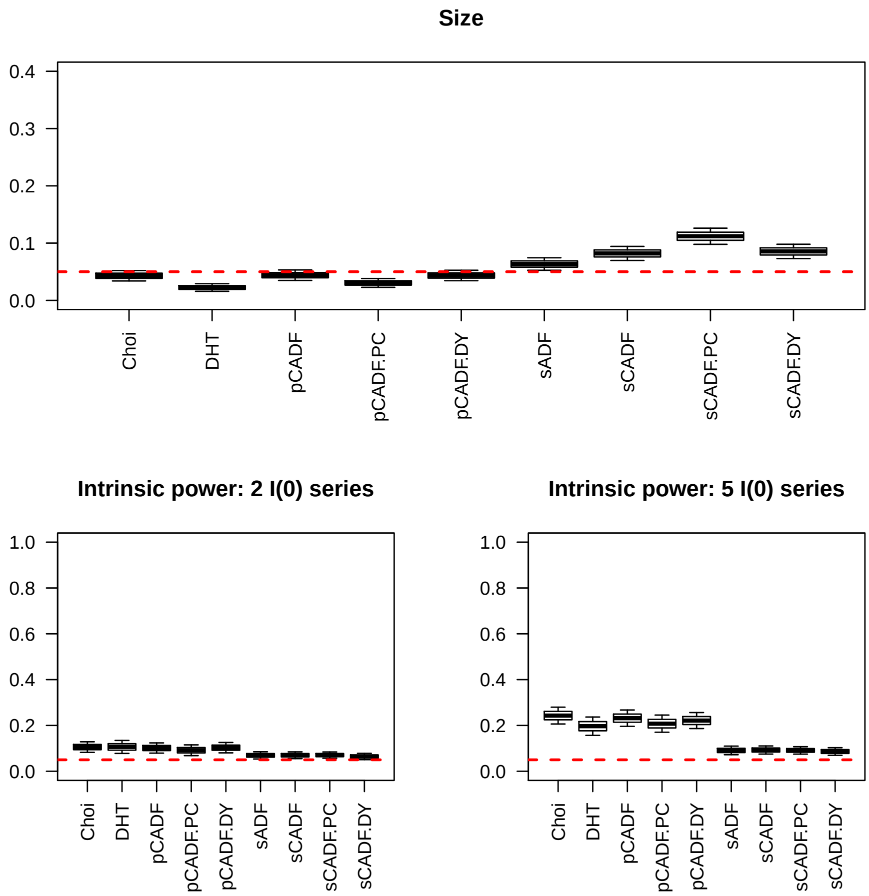

Figure 2: Simulation results: estimated rejection probabilities in experiment 1. The dashed horizontal line represents the 5% nominal size. The central line in each box is the estimated rejection probability, the box describes the $\pm\sigma$ interval around the estimated value and the whiskers extend to $\pm 2\sigma$.

The tests have approximately correct size under experiment 1, with the exception the CADF tests based on Simes (1986), that tend to reject too often. Demetrescu *et al.*'s test corrects for dependence when this is not present, with a small negative effect on the size. In experiment 1 there is no cross-dependence among the panel units and the time series of interest are independent of the stationary covariate. In such circumstances Choi's test and the pCADF test are asymptotically equivalent (see Costantini and Lupi 2011). Therefore, it is not surprising that the simulation results for the two tests in this case are very similar. The intrinsic power of the tests is in general fairly low, especially for the Simes-based tests whose power remains very low irrespective of the number of stationary series in the panel.

The effect of the presence of cross-dependence on the size of the tests is clear from Figure 3, with Choi's and the Simes-CADF tests rejecting too often. This is an expected outcome as far as Choi's test is concerned, but it is somewhat unanticipated for the tests based on Simes (1986) that should be able to cope with (positive) cross-dependence. The pCADF test tends to over-reject slightly when the true covariate is used. However, the pCADF has larger

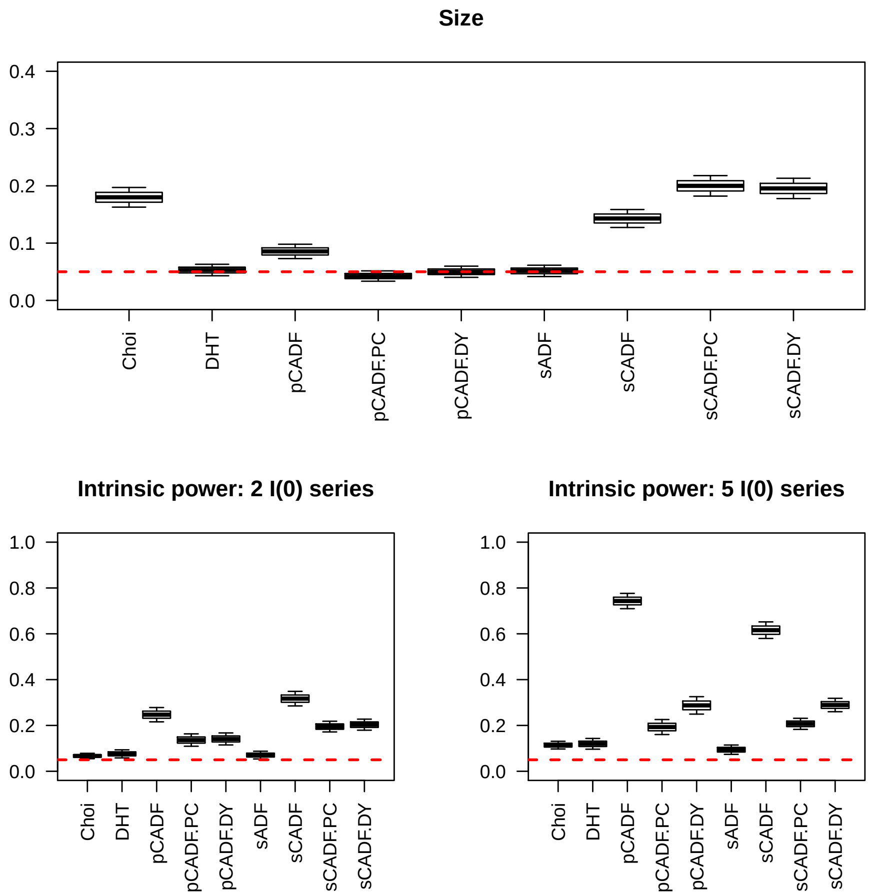

Figure 3: Simulation results: estimated rejection probabilities in experiment 2. The dashed horizontal line represents the 5% nominal size. The central line in each box is the estimated rejection probability, the box describes the $\pm\sigma$ interval around the estimated value and the whiskers extend to $\pm 2\sigma$.

intrinsic power than the other tests, due to the fact that the individual CADF tests use the information embodied in the stationary covariate. The version of the pCADF test based on the difference of the average of the other series is slightly better than the version based on the differenced first principal component: indeed, in experiment 2 there is no common factor that can be captured by the principal component. The Simes-based CADF tests also have impressive intrinsic power, despite their being oversized.

In experiment 3 more dependence is present, with ξ_t being a common factor. Choi's test is obviously significantly over-sized, while sCADF test has nearly correct size. The intrinsic power of all the CADF-based tests is very high.

All in all, a few clear indications to practitioners seem to emerge from these experiments. First, panel unit root tests that do not allow for cross-section dependence should be avoided, unless there are strong reasons to believe that the units are indeed cross-sectionally independent. In any instance, the use of a test for cross-dependence (see e.g. Pesaran 2004) is strongly suggested. This is not a new result, of course: early criticisms in this direction were already

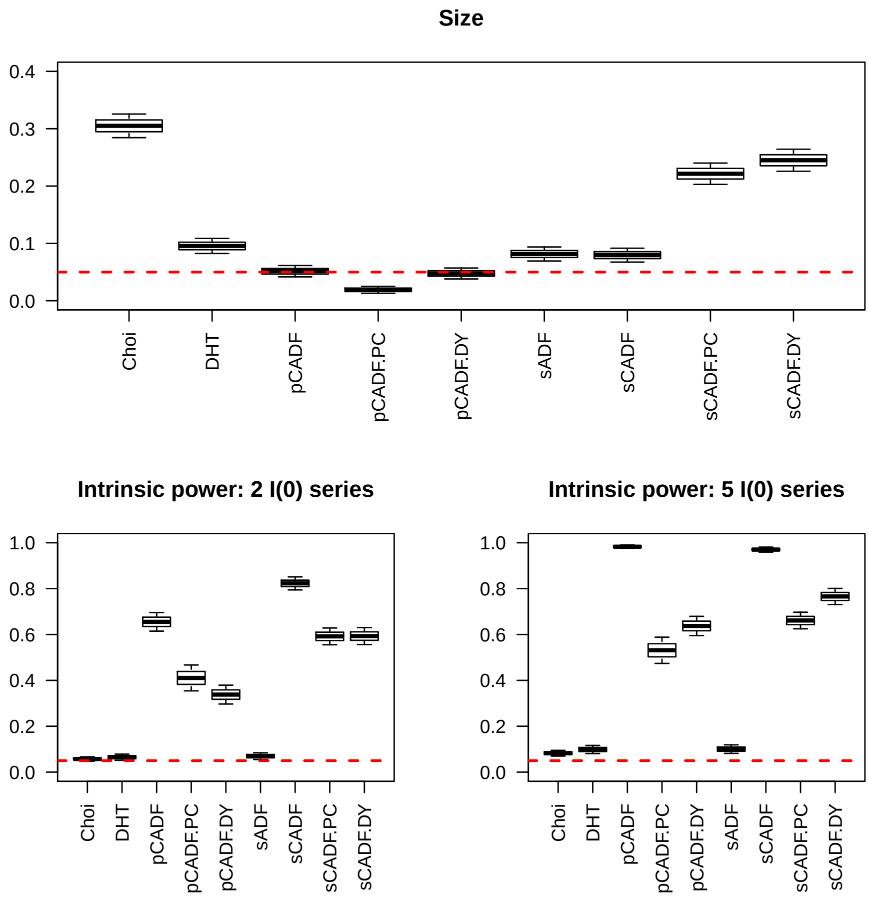

Figure 4: Simulation results: estimated rejection probabilities in experiment 3. The dashed horizontal line represents the 5% nominal size. The central line in each box is the estimated rejection probability, the box describes the $\pm\sigma$ interval around the estimated value and the whiskers extend to $\pm 2\sigma$.

present in O'Connell (1998) and Maddala and Wu (1999), among others. Second, the test proposed by Demetrescu *et al.* (2006) can cope with the existence of moderate cross-dependence, but has low intrinsic power. Simes-type based on ADF tests have in general fairly good size properties, but poor intrinsic power: on the contrary, using CADF tests gives good intrinsic power but some size distortions. Last but not least, the use of stationary covariates in panel unit root tests based on p value combination offers a simple way to obtain panel unit root tests with good size and large power gains.

5. An application to real data

We briefly illustrate how to use **punitroots** to perform panel covariate augmented Dickey-Fuller tests in practice. In particular, we investigate the presence of a unit root in OECD unemployment rates time series. The presence of unit roots in unemployment series is considered as an important stylized fact in terms of unemployment hysteresis (a seminal paper

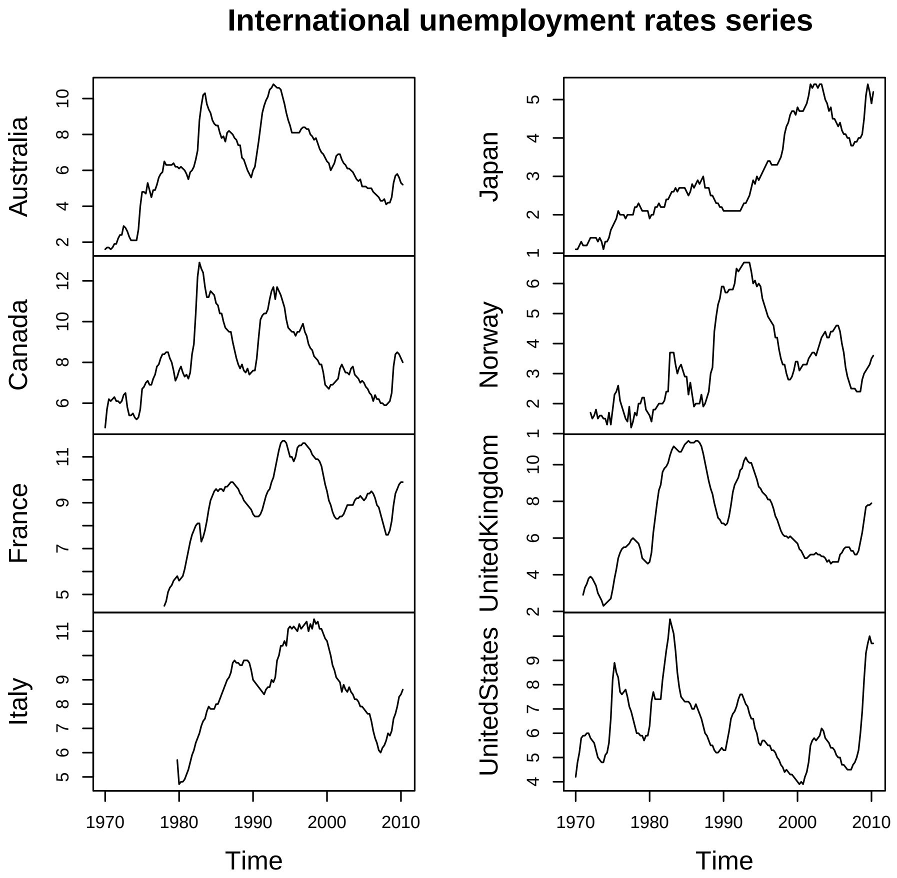

Figure 5: unemployment rates series. Source: OECD.StatExtracts.

in this area is Blanchard and Summers 1986). Of course, here we do not pursue the goal of obtaining support to firm theoretical conclusions on the subject.

The package **punitroots** can be downloaded from R-Forge, but the easiest way to install it is to use the command[10]

```
R> install.packages("punitroots", repos="http://R-Forge.R-project.org")
```

Assuming that the package has been correctly installed, we start by loading the package and the data that are readily available in **punitroots**:

```
R> library("punitroots")
R> data("OECDunemp")
```

A plot of the series is offered in Figure 5. The object `OECDunemp` is a quarterly multiple time series that spans the period 1970q1–2010q2, although some individual series are shorter. The original data source is OECD (2010).

[10]As usual, text following "`R>`" and written `using this font` should be intended as an input to be given in the R console by the user. The symbol "`+`" at the beginning of a line is a continuation line symbol, and should not be typed by the user. Other lines written `using this font` and not headed by "`R>`" report the output of the given commands.

Even if there is no strict contradiction between the unit root hypothesis and the fact that the unemployment rates series are bounded by construction (see e.g. Brunello, Lupi, and Ordine 2000), we nevertheless transform the data in order to avoid the bounding problem altogether. In particular, we transform the unemployment rates as $u := \log(U/(100 - U))$, with U the original unemployment rate series. Furthermore, in order to make the empirical example more interesting, we exclude the US unemployment rate from our sample.

```
R> u <- log(OECDunemp[,-8]/(100-OECDunemp[,-8]))
R> colnames(u) <- colnames(OECDunemp)[-8]
```

The command that performs the pCADF test, as well as Choi's and Demetrescu *et al.*'s, is `pCADFtest()` whose usage is fairly simple:

```
pCADFtest(Y, X=NULL, covariates=NULL, crosscorr=0.10, type="trend",
          max.lag.y=1, min.lag.X=0, max.lag.X=0,
          criterion=c("none", "BIC", "AIC", "HQC", "MAIC"), ...)
```

Here `Y` may be either a matrix or (preferably) a multiple time series. `X` and `covariates` are necessary if one wants to specify the covariates to be used in a proper panel CADF test (we will return on this aspect later). `crosscorr` specifies the significance level of a cross-dependence test (Pesaran 2004) used to decide if Hartung's correction for cross-section dependence should be applied: `crosscorr = 0` implies that no correction is used, while `crosscorr = 1` is specified if the user wants to force the use of Hartung's correction. `type` denotes the deterministic kernel of the models and can be specified as `"none"` (for models with no deterministic components), `"drift"` (for models with a constant term), and `"trend"` (for models with a constant and a linear trend): different deterministic components can be used for different panel units, if desired. `max.lag.y` is the maximum lag to be used for the dependent variable. `max.lag.X` and `min.lag.X` are the maximum and the minimum lag to be used for the covariates (if present): `min.lag.X` can be negative and in this case it represents the maximum lead. `criterion` specifies the lag selection criterion: various criteria can be used, including Akaike's Schwarz's and Hannan and Quinn's. It is worth noting that `pCADF()` does not require that all the series have the same length, and that we do not need to adjust manually the length of each series. Instead, the whole length of each series is used by default.

Let's use `pCADFtest()` to perform first a Choi test based on individual ADF tests, without controlling for the presence of possible cross-dependence. We start by defining models with a constant and at most 5 lags; the number of lags will be selected automatically by the AIC criterion:

```
R> Choi <- pCADFtest(Y=u, type = "drift", max.lag.y = 5, criterion = "AIC",
+ crosscorr = 0)
R> print(Choi)


        Panel-ADF test

data:
test statistic = -2.8315, mean.rho2 = NA, p-value = 0.002316
```

where `mean.rho2` refer to the average value of the estimated nuisance parameter ρ^2 in (9) which is not present in ADF models. The test rejects the null with a p value of 0.002.

Now we use the data to decide how to deal with cross-correlation. To do this, it is sufficient to eliminate "`cross.corr = 0`" from the command above:[11]

```
R> pADF <- pCADFtest(Y=u, type = "drift", max.lag.y = 5, criterion = "AIC")
R> print(pADF)

        Panel-ADF test

data:
test statistic.Ht = -1.2518, mean.rho2 = NA, p-value = 0.1053
```

It is clear that, once we allow for for cross-dependence, the panel test does not reject the null at the 10% significance level. However, more details can be gathered by using `summary()` instead of `print()`:

```
R> summary(pADF)

Panel Augmented DF test
Correction for cross-correlation: TRUE

                 p.value p
Australia     0.06614352 4
Canada        0.12748163 1
France        0.04573385 4
Italy         0.05087018 3
Japan         0.47135661 3
Norway        0.23291544 4
UnitedKingdom 0.23493749 5
                          Panel-ADF test
test statistic:               -1.2518101
p-value:                       0.1053195
```

The test summary reports the result of the panel test as well as the outcomes of the individual ADF tests. The column `p` reports the selected lags for each panel unit. The line "`Correction for cross-correlation: TRUE`" indicates that cross-dependence has been detected and Hartung's correction has been used in the combination of the p values as suggested in Demetrescu *et al.* (2006).

We might also decide to use a model with a constant and a linear trend for Japan, and models with only a constant for all the other countries:

```
R> deterministics <- rep("drift", ncol(u))
R> deterministics[5] <- "trend"
R> pADF2 <- pCADFtest(Y=u, type = deterministics, max.lag.y = 5, criterion = "AIC")
R> summary(pADF2)
```

[11]Note that we put the results in the object `pADF` (panel-ADF), since this is no longer a Choi test.

```
Panel Augmented DF test
Correction for cross-correlation: TRUE

                p.value p
Australia     0.06614352 4
Canada        0.12748163 1
France        0.04573385 4
Italy         0.05087018 3
Japan         0.21452460 3
Norway        0.23291544 4
UnitedKingdom 0.23493749 5
                         Panel-ADF test
test statistic:              -1.2675068
p-value:                      0.1024871
```

The result in this case is not very different.

We run now a proper pCADF test using the first difference of the first principal component of the y's as the stationary covariate. Here we imagine that there is a common factor that can be well approximated by the first principal component of the observed series. The task is easily accomplished by the command

```
R> pCADF.PC <- pCADFtest(Y=u, covariates = "PC", max.lag.y = 5, max.lag.X = 5,
+ type="drift", criterion = "AIC")
```

where `covariates = "PC"` is used to specify that the differenced first principal component is used as the covariate for all the panel units.

```
R> summary(pCADF.PC)

Panel Covariate Augmented DF test
Correction for cross-correlation: TRUE

                p.value       rho2 p q1 q2
Australia     0.78561539 0.07263228 4  2  0
Canada        0.05983713 0.18829530 4  0  0
France        0.73196380 0.89048975 1  2  0
Italy         0.55292228 0.92249686 3  2  0
Japan         0.69249008 0.71214085 3  5  0
Norway        0.16293142 0.74505825 3  1  0
UnitedKingdom 0.04078047 0.34522178 2  2  0
                         Panel-CADF test
test statistic:               -0.9326297
average estimated rho^2:       0.5537622
p-value:                       0.1755056
```

The panel test is not significant, despite the presence of a couple of seemingly significant individual p values. `rho2` is the estimated value of ρ^2, while `p`, `q1`, and `q2` indicate the selected orders of the lag polynomials.

We might also suspect that, if Δy_i is correlated with some or all the Δy_j (with $i \neq j$), then a potentially useful covariate for Δy_i is the average of Δy_j. This can be done easily:

```
R> pCADF.DY <- pCADFtest(Y=u, covariates = "DY", max.lag.y = 5, max.lag.X = 5,
+ type="drift", criterion = "AIC")
R> summary(pCADF.DY)

Panel Covariate Augmented DF test
Correction for cross-correlation: FALSE

                  p.value      rho2 p q1 q2
Australia     0.453165020 0.4401330 4  2  0
Canada        0.003856746 0.4627678 1  1  0
France        0.273749081 0.9627261 1  2  0
Italy         0.165666322 0.9102227 3  0  0
Japan         0.496594223 0.6072464 3  4  0
Norway        0.086129152 0.9480290 3  1  0
UnitedKingdom 0.061046011 0.8357018 2  0  0
                          Panel-CADF test
test statistic:              -2.749520647
average estimated rho^2:      0.738118126
p-value:                      0.002984125
```

It is interesting to note that using Δy_j makes the test to reject and has an influence on the decision on cross-dependence correction.

Another option is to carry out the pCADF test using stationary covariates chosen on the basis of economic reasoning. Here we use the first difference of the logs of the GDP of each country as the stationary covariates in the CADF regressions. In this case `X = X.GDP` indicates that the variables contained in the object `X.GDP` will be used as the stationary covariates. In order to specify that the first column of `X.GDP` is the covariate for the first time series in `u`, the second for the second and so forth, we use the argument `covariates = 1:ncol(u)` that indicates the column of `X.GDP` corresponding to each column of `u`:[12]

```
R> data("OECDgdp")
R> X.GDP <- diff(log(OECDgdp[,-8]))
R> pCADF.X <- pCADFtest(Y=u, X=X.GDP, covariates=1:ncol(u),
+ type="drift", max.lag.y = 5, max.lag.X = 5, criterion = "AIC")
R> summary(pCADF.X)

Panel Covariate Augmented DF test
Correction for cross-correlation: TRUE

                  p.value      rho2 p q1 q2
Australia     0.052083262 0.4846805 4  2  0
```

[12] It is also possible to have more than just one covariate for each time series. See the package manual for details.

```
Canada        0.076602092 0.4141025 1  1  0
France        0.038472251 0.2238134 1  3  0
Italy         0.367854054 0.9338642 3  0  0
Japan         0.004873835 0.2473101 0  2  0
Norway        0.362539946 0.7784478 3  4  0
UnitedKingdom 0.204977459 0.2967127 2  2  0
                          Panel-CADF test
test statistic:               -1.82123086
average estimated rho^2:       0.48270445
p-value:                       0.03428589
```

If we examine the results, it is easy to see that the cross-dependence correction has been used but the panel unit root null is now rejected.

Another possibility is to use the ADF-based test proposed by Hanck (2008) as well as the new CADF-based version of the same test proposed in this paper. This can be done by invoking the command `Simes()` in **punitroots**:

```
Simes(pCADFtest.results, alpha = 0.05)
```

where `pCADFtest.results` is an object of class `pCADFtest` (an object where the results of a previous *p*CADF test have been saved) and `alpha` is the desired significance level. For example:

```
R> Simes(Choi)

[1] TRUE
```

is the ADF-based test in the original form proposed by Hanck (2008), while

```
R> Simes(pCADF.X)

[1] FALSE
```

is the test in the new form that uses the more powerful individual CADF tests. The answer of this test is simply `TRUE` if the test does not reject the null and `FALSE` if the null is rejected. Consistently with the previous results, the panel unit root null is rejected only using individual CADF tests with proper covariates. The test can also be carried out using different significance levels simultaneously, as in

```
R> Simes(pCADF.X, alpha=c(0.01, 0.05, 0.10))

[1]  TRUE FALSE FALSE
```

so that in our example it is immediately evident that the test is significant at the 5% but not at the 1% level.

6. Summary

R offers the possibility of carrying out panel unit root testing in a friendly environment. Here we show that panel covariate augmented Dickey-Fuller tests as well as other p value combination tests can be carried out easily, using the functions `pCADFtest()` and `Simes()` in package **punitroots** (Kleiber and Lupi 2011). A small Monte Carlo analysis shows the performance of the tests under different conditions. An empirical example is carried out to illustrate the use and the flexibility of the commands.

punitroots is a concept package mainly devoted to the analysis of unit roots in panel of economic time series, with and without cross-dependence.

The present version of this paper has been prepared using R version 2.14.1 (2011-12-22) (R Development Core Team 2011) and package **punitroots** 0.0-1 (Kleiber and Lupi 2011).

Acknowledgements

This paper has been prepared for presentation at the 5th International conference on Computational and Financial Econometrics (CFE'11), Senate House, University of London, London, UK, 17-19 December 2011. I would like to thank Mauro Costantini, Giovanni Millo, Christian Kleiber, and Achim Zeileis for comments and discussion. The usual disclaimer applies.

References

Banerjee A, Wagner M (2009). "Panel Methods to Test for Unit Roots and Cointegration." In TC Mills, K Patterson (eds.), *Applied Econometrics*, volume 2 of *Palgrave Hanbook of Econometrics*, chapter 13, pp. 632–726. Palgrave MacMillan, Basingstoke.

Blanchard OJ, Summers LH (1986). "Hysteresis and the European Unemployment Problem." *NBER Macroeconomic Annual*, **1**, 15–78.

Breitung J, Pesaran MH (2008). "Unit Roots and Cointegration in Panels." In L Mátyás, P Sevestre (eds.), *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*, volume 46 of *Advanced Studies in Theoretical and Applied Econometrics*, 3rd edition, chapter 9, pp. 279–322. Springer-Verlag, Berlin.

Brunello G, Lupi C, Ordine P (2000). "Regional Disparities and the Italian NAIRU." *Oxford Economic Papers*, **52**(1), 146–177. ISSN 00307653.

Choi I (2001). "Unit Root Tests for Panel Data." *Journal of International Money and Finance*, **20**(2), 249–272.

Choi I (2006). "Nonstationary Panels." In TC Mills, K Patterson (eds.), *Econometric Theory*, volume 1 of *Palgrave Handbook of Econometrics*, chapter 13, pp. 511–539. Palgrave MacMillan, Basingstoke.

Costantini M, Lupi C (2011). "A Simple Panel-CADF Test for Unit Roots." *Oxford Bulletin of Economics & Statistics*. Forthcoming.

Croissant Y, Millo G (2008). "Panel Data Econometrics in R: The plm Package." *Journal of Statistical Software*, **27**(2), 1–43. URL http://www.jstatsoft.org/v27/i02/.

Demetrescu M, Hassler U, Tarcolea AI (2006). "Combining Significance of Correlated Statistics with Application to Panel Data." *Oxford Bulletin of Economics and Statistics*, **68**(5), 647–663.

Engle RF, Hendry DF, Richard JF (1983). "Exogeneity." *Econometrica*, **51**(2), 277–304.

Hanck C (2008). "An Intersection Test for Panel Unit Roots." *mimeo*, Technische Universität Dortmund.

Hansen BE (1995). "Rethinking the Univariate Approach to Unit Root Testing: Using Covariates to Increase Power." *Econometric Theory*, **11**(5), 1148–1171.

Hartung J (1999). "A Note on Combining Dependent Tests of Significance." *Biometrical Journal*, **41**(7), 849–855.

Im KS, Pesaran MH, Shin Y (2003). "Testing for Unit Roots in Heterogeneous Panels." *Journal of Econometrics*, **115**(1), 53–74.

Kleiber C, Lupi C (2011). ***punitroots**: Tests for Unit Roots in Panels of (Economic) Time Series, With and Without Cross-sectional Dependence*. R package version 0.0-1, URL https://r-forge.r-project.org/projects/punitroots/.

Levin A, Lin CF, Chu CSJ (2002). "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties." *Journal of Econometrics*, **108**(1), 1–24.

Lloyd CJ (2005). "Estimating Test Power Adjusted for Size." *Journal of Statistical Computation and Simulation*, **75**(11), 921–934.

Lupi C (2009). "Unit Root CADF Testing with R." *Journal of Statistical Software*, **32**(2), 1–19.

Maddala GS, Wu S (1999). "A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test." *Oxford Bulletin of Economics and Statistics*, **61**(Supplement 1), 631–652.

O'Connell PGJ (1998). "The Overvaluation of Purchasing Power Parity." *Journal of International Economics*, **44**(1), 1–19.

OECD (2010). "OECD.Stat Extracts." URL http://stats.oecd.org/Index.aspx.

Pesaran MH (2004). "General Diagnostic Tests for Cross Section Dependence in Panels." *Cambridge Working Papers in Economics 0435*, Department of Applied Economics, University of Cambridge.

R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL http://R-project.org.

Said SE, Dickey DA (1984). "Test for Unit Roots in Autoregressive-Moving Average Models of Unknown Order." *Biometrika*, **71**(3), 599–607.

Sarkar SK, Chang CK (1997). "The Simes Method for Multiple Hypothesis Testing With Positively Dependent Test Statistics." *Journal of the American Statistical Association*, **92**(440), 1601–1608.

Shaffer JP (1995). "Multiple Hypothesis Testing." *Annual Review of Psichology*, **46**, 561–584.

Simes RJ (1986). "An Improved Bonferroni Procedure for Multiple Tests of Significance." *Biometrika*, **73**(3), 751–754.

Zeileis A, Koenker R (2008). "Econometrics in R: Past, Present, and Future." *Journal of Statistical Software*, **27**(1), 1–5.

Affiliation:

Claudio Lupi
Department of Economics, Management, and Social Sciences (SEGeS)
University of Molise
Via De Sanctis I-86100 Campobasso, Italy
E-mail: `lupi@unimol.it`