Testing for Unit Roots in Heterogeneous Panels under Cross Sectional Dependence

Lorenzo Trapani*
Cass Business School and Universita' di Bergamo

Preliminary Version: March 2004
This Version: 16 May 2004
Comments welcome - please do not cite without permission

*Comments welcome: Faculty of Finance, Cass Business School, 106 Bunhill Row, London EC1Y 8TZ Tel.+/44/(0)20/70408698; Fax.+/44/(0)20/70408885; e-mail: L.Trapani@city.ac.uk.

Testing for Unit Roots in Heterogeneous Panels under Cross Sectional Dependence

Abstract

In this contribution, the currently available tests for unit root in panels with cross sectional dependence are reviewed. Each test is described in its model assumptions, in its computational details and asymptotic and finite sample performances. To shed some further light on the validity of the tests when only small samples are available, a broad Monte Carlo analysis is performed that encompasses all the numerical evidence obtained so far. All routines are in Ox.

J.E.L. Classification Numbers: C15; C32, C33
Keywords: Panels with cross-sectional dependency; Unit root tests; Monte Carlo simulations.

1 Introduction

The issue of testing for a unit root in panels has received great attention since the seminal works by Levin and Lin (1992, 1993), who considered a wide class of heterogeneous panel data models allowing for serial correlation. As shown by several contributions, this issue has achieved a certain importance as far as empirical applications are concerned. It is a well established result that unit root tests for single equation models may suffer from lack of power with respect to local stationary alternatives, and the applied literature has suggested to employ unit root tests based on panel data representations to overcome this issue. Particularly, this seems a successful way to address the problems of testing for the purchasing power parity hypothesis - see for instance the early arguments in MacDonald (1996), but also the critiques in O'Connell (1998) and the more recent contributions by Choi (2001), Bornhorst (2002) and Moon and Perron (2003b)-, of studying convergence among economies in growth models or the behavior of firm-level data - see the contribution by Hall and Mairesse (2001) on this later issue.

Since Levin and Lin's (1992) contribution, several tests have stemmed that, though considering the existence of serial correlation in the idiosyncractic component, ruled out the possibility of any kind of spatial dependence among the panel units, thereby assuming that the individual time series in the panel were independently distributed; among the various contributions that carried out this research, one could for example consider Im, Pesaran and Shin (2003), Maddala and Wu (1999) and Choi (2001); early reviews of this approach are provided by Banerjee (1999) and Baltagi and Kao (2000). The assumption of incorrelated units proves to be a very restrictive one, and it may be argued that failing in taking account of the existence of cross sectional dependence could likely result in biased tests. Maddala and Wu (1999) provide simulation evidence that when there is cross-sectional dependence among the panel units, the performances of Levin and Lin (1992) and Im, Pesaran and Shin (2003) tests are poor, and this is especially the case when the degree of dependence is large; similar results are reported in Pesaran (2003).

Dealing with cross dependence in nonstationary panels has proved to be a nontrivial task. As shown by Chang (2003), Wald type statistics based on the usual least squares estimators depend on nuisance parameters defining correlation across units. Moreover, whenever dealing with a panel where the number of units N is larger than the time dimension T, the least squares SUR framework is incapable of providing consistent estimates of the cross correlation matrix - see for instance Breitung and Das (2003). The main challenge faced by the literature has hence been to develop nuisance parameters free tests. Several authors have elaborated on this point. Most of the contributions derived so far have relied upon an approximate linear factor model for cross dependence - see for instance Bai and Ng (2004), Moon and Perron (2003), Moon, Perron and Phillips (2003), Phillips and Sul (2003) and Pesaran (2003). Other approaches have been considered by Chang (2002), who proposes a nonlinear IV estimation and testing framework based on Park and Phillips' (1999, 2001) theory for nonlinear transformation of integrated variables; Chang (2003) proposes bootstrap to overcome the problem of the dependence on nuisance parameters; Breitung and Das (2003) use several least squares estimation techniques and derive a series of tests based on the t-statistics to address the issue.

From a practitioner's viewpoint, there are some characteristics common to all these tests that are worth taking account of:

- the capability of discriminating local alternatives (i.e. to discriminate the case of near integration from the integration one) and the possibility of handling deterministic components such as possibly heterogeneous linear trends (referred to by the literature as "incidental trends");
- the dynamics in the error term that is implicitly or explicitly allowed for;
- the panel's dimensions tests are designed for (i.e. the number of units N and time observations T and their relative size);
- the small sample performance detected via Monte Carlo simulations.

As far as the first issue is concerned, it was noted by Phillips and Ploberger (2003) and Moon, Perron and Phillips (2003) that factor model based tests have generally speaking no power against local alternative when incidental trends are present. This theoretical finding is also confirmed by Moon and Perron (2003), whose testing strategy is found to have no local power in the presence of heterogeneous linear trends. On the other hand, even when a test could be designed to cope with the presence of incidental trends, a weel established result is that its asymptotic power envelope is found to be of size $O(1/\left(TN^{1/4}\right))$ in contrast to the usual size $O(1/\left(TN^{1/2}\right))$ one gets when incidental trends are not present. It is to be noticed that when tests have a less standard form or are not explicitly designed to take account of this issue, obtaining analytical result for the asymptotic power envelope is more complicated - examples are Chang (2002) or Pesaran (2003), and an exception is the work by Breitung and Das (2003).

As to the second point, virtually all tests reviewed here have been specifically designed to take account of the presence of dynamics in the error term, even though Moon, Perron and Phillips (2003) neglect this feature to obtain closed form results on the local asymptotic power of their test. The literature has so far addressed the issue of embedding error autocorrelation in two non mutually exclusive ways. First, one could augment the regression by plugging in some lagged values of the dependent variables - this is the approach considered by Chang (2002), Phillips and Sul (2003), Pesaran (2003) and Breitung and Das (2003). Alternatively, in a factor model framework, one could allow for serial dependence and even non stationarity in the factors, as in Bai and Ng (2004).

Another crucial aspect is the dimensions of the panel that are hypothesized to obtain the asymptotic theory that underpins the tests. This is also important from the applied point of view, because the tools available for the analysis of panels where $N > T$ are different, and sometimes inapplicable, to the case when $T > N$. As far as the literature on unit root tests is concerned, it is worth emphasizing that the results obtained hold almost always, asymptotically, for large T and, sequentially, large N. From an applied perspective, this would imply that not only should both dimensions be large, but that T must be much larger than N. A partial exception to this is the paper by Chang (2002), which allows for finite N even though T must be infinite. On the contrary, Pesaran (2003) provides a totally different framework where N is required to be large and T can, in principle, be finite and as small as 3. This is a complete change in the usual Phillips and Moon's (1999) scheme where sequential limits are taken for $T \to \infty$ followed by $N \to \infty$.

Lastly, since all tests rely on asymptotic results, it is worth analyzing their finite sample performances in terms of size and power. It has been argued - see Breitung and Das (2003) - that the commonly employed factor representation is just an approximation of the true covariance structure, and therefore it could lead to substantial size distortion. Nonetheless, the results obtained so far are, according to simulation evidence, fairly positive; these are anyway heavily dependent on the degree of serial correlation and by the presence or not of linear (incidental) trends in the model. Very often tests are not designed to take account of a large amount of serial dependence, and numerical results show that under autocorrelation the tests performance is rather poor - for example, Phillips and Sul (2003) show that their test does not perform well when the errors have an MA(1) representation; in general, however, the Monte Carlo evidence provided in the articles reviewed is rather limited as far as this aspect is concerned. Also the presence of incidental trends makes the tests small sample performance either poorer (as it is the case in Phillips and Sul, 2003) or not powerful at all (as shown in Moon and

Perron, 2003). On the contrary, Pesaran's (2003) test seems to have good power and size for reasonably large T, and the same holds for the case of a linear trend present in the model. As far as the cross sectional and time dimensions N and T are concerned, they have, as expected, an influence on the tests performance. Even if in principle some tests are designed to hold for finite values of either N or T, evidence from simulation shows that they have severe size distortion or lack of power when either dimension is small - this is the case of Pesaran (2003), or Chang (2002). The performance gets even worse when factor estimation is called for, as shown by Moon and Perron (2003). In general, however, it ought to be emphasized that the Monte Carlo evidence provided in these articles is rather scanty, and most often some crucial features of the model that could affect small sample performance are overlooked - as an example, no analysis on the impact of error dynamics is conducted in Moon and Perron's (2003) contribution.

In the light of the review sketched above, it is our opinion that there are at least three key points that are missing within the unit root tests in panel data framework. First, from an applied econometrician's viewpoint it could be interesting to have a comparative analysis of the statistical rationale behind each test, so as to know when, given some characteristics of the panel at hand, using a certain test is recommended. Secondly, in the light of the scarce simulation evidence provided in the reviewed papers, a broad Monte Carlo analysis is required to understand which test one should select/rely on when dealing with finite dimension panels; we believe this issue to deserve a full treatment to make this tests applicable.

On the ground of these considerations, the plan of this work is threefold:

1. firstly, we will provide a detailed analysis of each test (Section 2);
2. secondly, a simulation study to fill some gaps left in the literature will be implemented. This is meant to provide some guidelines for practitioners and to shed some further light on size and power of tests under various circumstances;

The remainder of the paper is organised as follows. Section 2 reviews the existing approaches, highlighting their computational aspects and their performance under various circumstances. Section 3 reports evidence from a large set of Monte Carlo experiments. Section 4 concludes.

2 Background theory

In this Section, we will review the existing tests for unit root under the presence of cross sectional dependence. The aim of this section is both to refresh the existing literature embodying it in a general framework and to point out the main features of each test; more specifically, the following aspects will be considered for every test:

- the model representation, the specification of the error term dynamics and the way cross sectional dependence is modelled;
- the test statistics, with respect to both the computational viewpoint and the underpinning asymptotic theory;
- Monte Carlo evidence so far provided for the small sample behavior.

Before introducing the single tests, anyway, we will set out the general framework. Let y_{it} be the dependent variable, with $i = 1, ..., N$ and $t = 1, ..., T$, x_{it} a set of covariates including deterministic components such as fixed effects and linear trends and ε_{it} an idiosyncratic component whose properties will be discussed later on, but for which we do not assume that $E(\varepsilon_{it}\varepsilon_{js}) = 0$ for any $i \neq j$. The most general specification one can consider is

$$y_{it} = \rho_i y_{it-1} + f(x_{it}) + \varepsilon_{it} \tag{1}$$

where $f(\cdot)$ is a (linear) relationship. It is worth noticing that the hypothesis testing framework is always

$$\begin{cases} H_0 : \rho_i = 1 & \text{for } i = 1, ..., N \\ H_1 : \rho_i = 1 & \text{for } i = 1, ..., N_1 < N \quad \text{and } \rho_i < 1 \text{ for } i = N_1 + 1, .., N \end{cases} . \tag{2}$$

Therefore, whilst the null hypothesis means that all autoregressive roots in the panel are homogeneously equal to 1, rejecting it is in principle possible if only one of the ρ_is is less than unity, and heterogeneity is allowed for[1]. It is hence worth stressing that rejecting the null hypothesis does not mean that the panel is stationary, but only that there are some stationary units in it; all the tests considered here allow for the ρ_is to be heterogeneous under the alternative.

[1] Notice that early approaches, such as the ones due to Quah (1994) or Levin and Lin (1992, 1993) considered homogeneity of the ρ_is, threby not allowing for group specific effects under the alternative. The contribution of Im, Pesaran and Shin (2003) generalizes this framework letting the ρ_is to be heterogeneous under the alternative hypothesis.

It is also worth exploring the theoretical capability of tests to reject a sequence of local alternatives, to assess the tests power with respect to the case of near integration. To do this, the literature has considered the following specification for ρ_i under a (sequence of) local alternatives:

$$\rho_i = 1 - \frac{c_i}{TN^\eta},$$

where c_i is assumed to be a nonnegative random variable.[2] If one assumes, as in Moon and Perron (2003), that the c_is are iid with mean μ_c, then equation (2) can be conveniently rewritten as

$$\begin{cases} H_0 : c_i = 1 & \text{for all } i\text{s} \\ H_1 : c_i < 1 & \text{for some } i\text{s} \end{cases},$$

or, equivalently

$$\begin{cases} H_0 : \mu_c = 0 \\ H_1 : \mu_c > 0 \end{cases}.$$

The local-to-unity power is explored, when this is feasible, with respect to this set of assumptions and using the same tools as in Moon, Perron and Phillips (2003).

After this brief introduction to the null and alternative hypothesis specification, we will turn to the description of each test.

2.1 Chang (2002)

The model considered by Chang (2002) is the following specification of model (1):

$$y_{it} = \rho_i y_{it-1} + \varepsilon_{it}, \tag{3}$$

where $i = 1, .., N$ and $t = 1, ..., T_i$; hence, unbalanced panels are not ruled out. The hypothesis testing framework is the same as in (2) and an AR(p_i) specification is allowed for each error component term:

$$\varepsilon_{it} = \alpha_i (L) u_{it}, \tag{4}$$

u_{it} being a white noise. Assumptions made on the AR polynomial $\alpha_i (L)$ make it compatible with the class of invertible Gaussian ARMA models. Also, the presence of deterministic components is allowed, thereby giving the following alternative specification of (3):

$$z_{it} = \mu_i + \delta_i t + y_{it}. \tag{5}$$

[2] As already mentioned, η is found to be equal to 1/2 when no incidental trends appear in x_{it} and to 1/4 when these are considered among the explanatory variables.

It is to be noted that model (5) explicitly allows for incidental trends (since the δ_is are allowed to be heterogeneous across units).

Cross dependence is not explicitly modelled in this framework. The vector $\varepsilon_t = [\varepsilon_{1t}, ..., \varepsilon_{Nt}]'$ is assumed to have covariance matrix equal to $E(\varepsilon_t \varepsilon_t') = \Sigma$, but no specification is chosen to represent the off-diagonal elements of Σ. In this respect, Chang's approach can be viewed as a non parametric one, unlike the rest of the literature that specifies explicitly the cross dependence structure.

The regression in (3) is standardly augmented to take account of the autocorrelation structure in (4); under the null H_0, $\Delta y_{it} = u_{it}$, which leads to:

$$y_{it} = \rho_i y_{it-1} + \sum_{k=1}^{p_i} \beta_{ik} \Delta y_{i,t-k} + u_{it}.$$

To estimate this equation, the Author proposes an IV regression where the $\Delta y_{i,t-k}$s are used as instruments of themselves; to proxy y_{it-1}, one should use the new variable $F(y_{it-1})$. This is generated according to a regularly integrable function constrained not to be orthogonal to y_{it-1}; Chang proposes the following functional form

$$F(y_{it-1}) = y_{it-1} e^{-c_i |y_{it-1}|}$$

where $c_i = K T_i^{-1/2} s^{-1}(\Delta y_{it})$, with $s^2(\Delta y_{it}) = T_i^{-1} \sum_{t=1}^{T_i} (\Delta y_{it})^2$. The value of K is to be chosen, and even if the author claims that the tests performance is not very sensitive to such choice, this introduces a degree of arbitrariness in the generation of the IV $F(y_{it-1})$.[3] When the model employed is (5), the approach does not change, and the same IV regression approach is considered after demeaning and detrending the variable z_{it}. After estimating, the following standard t-test can be constructed for ρ_i for each unit:

$$t_{C,i} = \frac{\hat{\rho}_i - 1}{se(\hat{\rho}_i)}.$$

Chang proves that, when $T_i \to \infty$, $t_{C,i} \Longrightarrow N(0,1)$. Under stricter assumptions on the convergence of the T_is towards infinite, Chang - following the lines of Park and Phillips' (2001) theory - proves that the t_is are independent across i. So, even for finite N (and clearly for large N as well), the following statistics

$$S_{C,N} = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} t_i \tag{6}$$

[3] In her simulation exercise, Chang sets $K = 3$. The author also suggests to slightly increase this value as the time size T_i decreases, but no optimal rule is provided, and one should always do some sensitivity analysis to validate their choice of K.

is standard normal; $S_{C,N}$ is the statistics suggested for the purpose of testing the null H_0. Hence, the asymptotic theory underpinning Chang's approach to testing for unit roots does not require the number of units N to be large, but it does require all the time dimension of all units, T_i, to be large. If this is not the case, the t_is cannot be independent and the result established by equation (6) is no longer valid. It is worth noticing that even if computationally the test statistics $S_{C,N}$ is quite straightforward to implement,[4] the calculations needed to obtain the result in (6) are rather cumbersome. For this reason, no theoretical results on the asymptotic power envelope have been derived.

Small sample performance is analyzed via a set of simulations; the author considers the behavior of $S_{C,N}$ under AR(1) errors and various (N, T) couples. [5] The test is shown to suffer from very mild size distortion even when T is small - i.e. equal to 25 - while such distortion is almost completely absorbed when T increases; these results hold for all cross sectional dimension. As far as power is concerned, this is shown to be quite good, and better than the one of Im, Pesaran and Shin's (2003) statistics. These findings are contradicted by Im and Pesaran (2003), who show, theoretically and by simulation, that when N increases the test becomes oversized; particularly, they find that Chang's test is grossly over-sized even for mild levels of cross sectional dependence even for relatively small values of N. Such evidence indicates that Chang's test is fit for the long panel case (i.e. large T), but should not be used when N is large.[6]

2.2 Breitung and Das (2003)

Breitung and Das (2003) consider, following the same lines as Levin and Lin (1992, 1993), a homogeneous version of equation (3)

$$y_{it} = \rho y_{it-1} + \varepsilon_{it},$$

or, in a Dickey-Fuller form

$$\begin{aligned} \Delta y_{it} &= (1-\rho)\, y_{it-1} + \varepsilon_{it} = \\ &= \phi y_{it-1} + \varepsilon_{it}. \end{aligned} \tag{7}$$

[4] An attempt to use this test to check the presence of a unit root in a panel of firms' profit rate was made by Ioannidis *et al.* (2003).

[5] These are the combinations from the sets: $N = \{5, 15, 25, 50, 100\}$ and $T = \{25, 50, 100\}$.

[6] It is interesting to notice that such distorsion occurs even when N is equal to 5 for $T = 100$, thereby suggesting that the favourable simulation results provided in Chang (2003) are due to her particular choice of the simulation parameters.

Unbalanced panels are not allowed for, so the time dimension, which will be referred to as T, is the same across all units. Also the presence of deterministic components is analyzed, allowing for fixed effects (but not for trends)

$$z_{it} = \mu_i + y_{it}. \tag{8}$$

In this cases, the null and alternative hypotheses simplify to

$$\begin{cases} H_0 : \rho = 1 \\ H_1 : \rho < 1 \end{cases}.$$

Like in Chang (2002), no parametric specification for cross dependence is imposed, the only assumption being that the vector $\varepsilon_t = [\varepsilon_{1t}, ..., \varepsilon_{Nt}]'$ is assumed to have covariance matrix equal to $E(\varepsilon_t \varepsilon_t') = \Sigma$, without constraining the off-diagonal elements in Σ to be zero.

In order to take account of serial correlation, the data generating process for Δy_{it} is augmented to make it an AR(p) process, where p is again homogeneous across units:

$$\Delta y_{it} = (1 - \rho) y_{it-1} + \sum_{j=1}^{p} \gamma_{ij} \Delta y_{it-j} + u_{it}. \tag{9}$$

Notice that since the γ_{ij}s can differ among units, heterogeneity is not ruled out in the dynamics specification, and it is confined to ρ.

Stacking Δy_{it} and y_{it-1} in the N-dimensional vectors $\Delta y_t = [\Delta y_{1t}, ..., \Delta y_{Nt}]'$ and $y_{t-1} = [y_{1t-1}, ..., y_{Nt-1}]'$, and defining $\phi \equiv (1 - \rho)$, one gets the SUR version of equation (7):

$$\Delta y_t = \phi y_{t-1} + \varepsilon_t. \tag{10}$$

If any dynamics is present, the authors suggest to pre-whiten the observations, i.e. to plug in equation (10) the variables $y_{it}^* = y_{it} - \sum_{j=1}^{p} \hat{\gamma}_{ij} y_{it-j}$, where $\hat{\gamma}_{ij}$ are consistent estimates of the γ_{ij}s. To take account of the presence of deterministic components, a consistent demeaning scheme is applied, and the μ_is are proxied via the initial values y_{i0}, thereby replacing y_{it-1} with $(y_{it-1} - y_{i0})$; this follows from Phillips and Schmidt's (1992) argument. In both cases, the results described below for equation (7) keep holding. To estimate ϕ and test for the unit root hypothesis, the authors consider an OLS framework that uses Arellano's (1987) robust estimation technique. The proposed statistics is the following t-test

$$t_{BD} = \frac{\sum_{t=1}^{T} y_{t-1}' \Delta y_t}{\sqrt{\sum_{t=1}^{T} y_{t-1}' \hat{\Sigma} y_{t-1}}}, \tag{11}$$

where the estimator $\hat{\Sigma}$ is standardly computed as

$$\hat{\Sigma} = T^{-1} \sum_{t=1}^{T} \left(\Delta y_t - \hat{\phi} y_{t-1} \right) \left(\Delta y_t - \hat{\phi} y_{t-1} \right)',$$

$\hat{\phi}$ being the OLS estimate of ϕ in (10).

Another alternative put forward by the authors is the GLS based test; given the transformation $\tilde{y}_t = \hat{\Sigma}^{-1/2} y_t$, this is defined as

$$t_{gls} = \frac{\sum_{t=1}^{T} \tilde{y}_{t-1}' \Delta \tilde{y}_t}{\sqrt{\sum_{t=1}^{T} \tilde{y}_{t-1}' \tilde{y}_{t-1}}}. \tag{12}$$

Computationally, it ought to be noticed that this testing framework would result to be infeasible when $N > T$, which is a well known issue in SUR estimation. For such reason, asymptotic theory is derived under the standard sequential limit assumption; following the same lines as in Phillips and Moon (1999), Breitung and Das prove that under the null hypothesis, $t_{BD} \Longrightarrow N(0,1)$.[7] Requiring that N and T increase sequentially means, from a practitioner's viewpoint, that (though computationally easy) this test cannot be implemented for "short" panels, where the number of units exceeds the number of time observations. The asymptotic power envelope derived by the authors is proved to have the form $O\left(1/\left(TN^{1/2}\right)\right)$, which is a standard result when incidental trends are not considered.

As far as small sample performance is concerned, the authors consider a Monte Carlo exercise where in the simulated panel ρ is assumed to be homogeneous. Results show that the robust OLS test performs quite well in terms of power, always outperforming Chang's (2002) IV based method; nonetheless, it does not attain its nominal size, which is claimed to be achieved only when T is very large with respect to N.

2.3 Phillips and Sul (2003)

Within a comprehensive analysis on estimation and testing in heterogeneous panels with cross sectional dependence, Phillips and Sul (2003) propose a general framework that also allows for testing the null hypothesis of a unit

[7] Under the more restrictive assumption that $N/T \to 0$, a GLS estimation procedure could be considered. Since this is more efficient than OLS, one would expect a more powerful test statistic. Asymptotically, the t-test based on GLS would have the same standard normal distribution. All the conclusions drawn with respect to the robust OLS procedure sketched here hold for this procedure as well.

root in a panel data model. The process y_{it} is modelled as in equation (3),[8] where the error term is modelled according to a common factor representation:

$$\varepsilon_{it} = \lambda_i \eta_t + u_{it}, \tag{13}$$

where the u_{it}s are contemporaneously incorrelated and $E\left(u_{it}^2\right) = \omega_i^2$. The presence of deterministic components can be taken into account by rewriting equation (3) as equation (5); even within this framework, fixed effects and incidental trends are accounted for.

Equation (13) explicitly models cross sectional dependence. The authors choose a single common factor representation which can be treated according to a standard of assumptions,[9] and under the standard assumption of η_t iid over t as a standard normal, the degree of contemporaneous correlation is given by $E\left[\varepsilon_{it}\varepsilon_{jt}\right] = \lambda_i \lambda_j$ for $i \neq j$, and is therefore controlled for by the factor loadings. Notice that, given the vector $\lambda = [\lambda_1, ..., \lambda_N]$ and the matrix $\Omega = diag\left\{\omega_1^2, ..., \omega_N^2\right\}$, it holds that $E\left[\varepsilon_t \varepsilon_t'\right] \equiv \Sigma = \Omega + \lambda\lambda'$.

When serial correlation is present, equation (5) is augmented exactly as in Chang (2002), allowing for heterogeneity in the lag lengths p_i and in the dynamics parameters.

The approach considered by Phillips and Sul consists of getting rid of cross sectional dependence by making the y_{it}s, or the z_{it}s if the correct specification is (5), orthogonal across units. When one gets processes that are cross sectionally incorrelated, say y_{it}^*, then the tests proposed by the early literature for heterogeneous panels can be applied; therefore, the one proposed by Phillips and Sul could be viewed as an algorithm to make the y_{it} contemporaneously incorrelated. Such orthogonalization procedure is implemented as follows with respect to the case of no serial correlation:

1. **Step 1**: construct the moment matrix of (demeaned and detrended) $\hat{\Sigma} = T^{-1}\sum_{t=1}^{T} \hat{\varepsilon}_t \hat{\varepsilon}_t'$. For large T and fixed N, this is a consistent estimator of Σ;

2. **Step 2**: estimate Ω and λ solving iteratively the following optimization problem (solutions are denoted as $\hat{\Omega}$ and $\hat{\lambda}$):

$$\left(\hat{\Omega}, \hat{\lambda}\right) = \arg\min_{\Omega,\lambda} tr\left\{\left[\hat{\Sigma} - \left(\Omega + \lambda\lambda'\right)\right]\left[\hat{\Sigma} - \left(\Omega + \lambda\lambda'\right)\right]'\right\}.$$

[8] Phillips and Sul (2003) consider a homogeneous panel model, where $\rho_i = \rho$ for all is. Nonetheless, the unit root tests they propose are valid also under heterogeneity, as implicitly suggested in Moon and Perron (2003).

[9] See for instance the comprehensive analysis by Bai and Ng (2004) or Moon and Perron (2003).

The result is a consistent estimator for Ω and λ[10], again for large T and fixed N ;

3. **Step 3**: construct the orthogonal complement $N \times (N-1)$ matrix $\hat{\lambda}_{\perp}$ and the $N \times (N-1)$ matrix $\hat{F}_{\lambda} = \left[\hat{\lambda}_{\perp}' \hat{\Omega} \hat{\lambda}_{\perp}\right]^{-1/2} \hat{\lambda}_{\perp}'$;

4. **Step 4**: orthogonalize the (demeaned and detrended) vector $y_t = [y_{1t}, ..., y_{Nt}]'$ getting the new, $N-1$-dimensional $y_t^* = \hat{F}_{\lambda} y_t$.

This procedure removes consistently cross sectional dependence[11]; it is anyway worth emphasizing that the time dimension T is required to be substantially larger than the cross sectional one N in order for the algorithm to yield consistency. Several testing approaches can be employed after making the data orthogonal; particularly, Phillips and Sul study the meta-analysis approach analyzed by Maddala and Wu (1999) and considered in a similar context by Choi (2001). Such testing approach can be described as follows: let π_i be the p-value of a unit root test applied to the i-th unit of the panel. Then the following statistics can be defined:

$$P = -2 \sum_{i=1}^{N-1} \ln(\pi_i) \tag{14}$$

$$Z = \frac{1}{\sqrt{N}} \sum_{i=1}^{N-1} \Phi^{-1}(\pi_i), \tag{15}$$

where $\Phi^{-1}(\cdot)$ is the inverse of the Gaussian cdf. Following Choi (2001), all the three statistics are standard normal under the sequential limit scheme.[12]

Small sample evidence is reported for both P and Z statistics. These are found to always outperform the Im, Pesaran and Shin (2003) test, which is not surprising also in the light of the previous findings of Maddala and Wu (1999). When there is no dynamics in the error term, or when this is restricted to the AR(1) case, these tests perform very well even when the alternative hypothesis is close to the unit root case; size is almost unbiased and constant when N increases. Nonetheless, size distortion appear for both tests under more MA(1) dynamics, showing that augmenting the regression

[10] The details of this moment based optimization procedure, and the steps required to compute $\hat{\Omega}$ and $\hat{\lambda}$ are illustrated in Philips and Sul's work.

[11] When there is serial correlation, this procedure is still valid, provided that one uses the residuals from the augmented regression.

[12] Other testing approaches are possible within this framework; a complete treatment is in Phillips and Sul.

may not be enough to get rid of serial dependence. It is also worth noticing that results have been obtained for small values of N compared to T and generating ε_{it} assuming that the single factor structure is the true one. It is unclear what could happen to size and power when either N s large or when the factor structure is to be estimated/misspecified. Moreover, no numerical study has been undertaken to compare size and power of the tests stemming from the orthogonalization procedure sketched above and other testing approaches that explicitly take account of cross sectional dependence.

2.4 Moon and Perron (2003)

Moon and Perron's (2003) study is grounded on a similar approach to that of Phillips and Sul (2003). The model considered by the authors for y_{it} is the same as in equation (3), and the error component ε_{it} is assumed to follow a factor model

$$\varepsilon_{it} = \lambda_i^{0\prime}\eta_t^0 + u_{it}. \tag{16}$$

Two sources of autocorrelation are considered: an AR(∞) structure is allowed for u_{it}, whose DGP is expressed as

$$u_{it} = \sum_{j=0}^{\infty} d_{ij} w_{t-j},$$

and factors η_t^0 are assumed to be independent of each other but not serially incorrelated. The cases of fixed effects and incidental trends are embodied in this framework, leading to the same structure as in equation (5). Moreover, the testing procedure introduced by Moon and Perron does not require any augmentation of either (3) or (5); consistent estimates of the (short and long run) variances of u_{it} suffice to make the testing procedure feasible.

This approach to modelling cross sectional dependence via a parsimonious, approximate representation in terms of common factors is in principle the same as in Phillips and Sul (2003). However, now k factors are allowed within this representation, and hence both λ_i^0 and η_t^0 are k-dimensional vectors. The extent of autocorrelation is given by $E\left[\varepsilon_{it}\varepsilon_{jt}\right] = \lambda_i^{0\prime} E\left[\eta_t^0 \eta_t^{0\prime}\right] \lambda_j^0$.[13]

From the computational viewpoint, Moon and Perron derive a procedure where the number of common factors is not specified a priori, and is to be estimated. To construct the test statistics, consider equation (5) without incidental trends - i.e. the model $z_{it} = \mu_i + y_{it}$. Letting Z be a $T \times N$ matrix

[13] Assumptions on the factors η_t^0 in terms of moments existence and dynamics are the ones usually dealt with in the literature. Therefore they will not be reported here - see however Bai and Ng (2004) for a comprehensive treatment of this topic.

such that $\{Z'\}_{it} = z_{it}$, and $Z_{-1} = LZ$, L being the lag operator, and defining the error term matrix u and the factor components Λ^0 and Ξ^0 in a similar fashion, one can rewrite the model as

$$Z = \rho Z_{-1} + \Xi^0 \Lambda^{0\prime} + u. \tag{17}$$

Consider now the pooled estimate of ρ is given by

$$\hat{\rho}_{pool} = \frac{tr\left(Z'_{-1}Z\right)}{tr\left(Z'_{-1}Z_{-1}\right)}.$$

Moon and Perron estimation and testing procedure can be written down as the following algorithm:

1. **Step 1**: construct the $T \times N$ matrix of residuals $\hat{\varepsilon} = Z - \hat{\rho}^{pool} Z_{-1}$;

2. **Step 2**: obtain a consistent estimate of the number of factors k, say $\hat{k}$. This issue was dealt with by Bai and Ng (2002), and it can be summarized as follows:

   (a) select an arbitrary, $N \times r$ matrix Λ_r;

   (b) minimize, with respect to r, a criterion function $CF(\cdot)$ of the form

   $$CF(r) = \varphi\left[W_{NT}(\Lambda_r, r)\right] + rG_{NT}$$

   where

   $$W_{NT}(\Lambda_r, r) = \frac{tr\left[\hat{\varepsilon}P_\Lambda \hat{\varepsilon}'\right]}{NT},$$

   P_Λ is the projection matrix $I_N - \Lambda_r\left(\Lambda_r \Lambda'_r\right)\Lambda'_r$ and G_{NT} a penalty function;[14]

3. **Step 3**: given $\hat{k}$, estimate consistently Ξ^0 and Λ^0. This is performed by minimizing

   $$V_{NT} = \frac{tr\left[\left(\hat{\varepsilon} - \Xi\Lambda\right)\left(\hat{\varepsilon} - \Xi\Lambda\right)'\right]}{NT}$$

   with respect to Ξ and Λ under either $\Lambda\Lambda'/N = I_{\hat{k}}$ or $\Xi\Xi'/T = I_{\hat{k}}$, denoting the solutions as $\tilde{\Xi}$ and $\tilde{\Lambda}$. Consider the rescaled estimate of Λ

   $$\check{\Lambda} = \tilde{\Lambda}\left(\frac{\tilde{\Lambda}'\tilde{\Lambda}}{N}\right)^{1/2},$$

   and the projection matrix $\check{P}_\Lambda = I_N - \check{\Lambda}\left(\check{\Lambda}\check{\Lambda}'\right)\check{\Lambda}'$;

[14]Bai and Ng (2002) propose different criteria and specifications of the penalty function, each of which is optimal for different combinations of (N, T).

4. **Step 4**: obtain consistent estimates[15] for $\omega_u^2 = \lim \frac{1}{N} \sum_{i=1}^{N} \left(\sum_{j=0}^{\infty} d_{ij} \right)^2$, $\phi_u^4 = \lim \frac{1}{N} \sum_{i=1}^{N} \left(\sum_{j=0}^{\infty} d_{ij} \right)^4$ and $\lambda_u = \lim \frac{1}{N} \sum_{i=1}^{N} \left(\sum_{l=1}^{\infty} \sum_{j=0}^{\infty} d_{ij} d_{ij+l} \right)$, namely $\hat{\omega}_u^2$, $\hat{\phi}_u^4$ and $\hat{\lambda}_u$;[16]

5. **Step 5**: compute the pooled OLS estimate for ρ with respect to the orthogonalized version of (17):

$$\hat{\rho}_{pool}^{+} = \frac{tr\left(Z_{-1}' \check{P}_\Lambda Z\right) - NT\hat{\lambda}_u}{tr\left(Z_{-1}' \check{P}_\Lambda Z_{-1}\right)}. \qquad (18)$$

This algorithm to estimate ρ is similar to that in Phillips and Sul (2003), but now the number of factors is no longer constrained to be equal to 1. It is also worth noticing that all the intermediate estimators obtained in the algorithm are consistent even with respect to the joint limit $(N, T) \to \infty$. The authors show that

$$\begin{aligned} t_{MP}^{a} &= \frac{\sqrt{N}T\left(\hat{\rho}_{pool}^{+} - 1\right)}{\sqrt{2\hat{\phi}_u^4 / \hat{\omega}_u^4}} \Longrightarrow N(0,1) \\ t_{MP}^{b} &= \sqrt{N}T\left(\hat{\rho}_{pool}^{+} - 1\right)\sqrt{\frac{1}{NT^2} tr\left(Z_{-1}' \check{P}_\Lambda Z_{-1}\right)\frac{\hat{\omega}_u^4}{\hat{\phi}_u^4}} \Longrightarrow N(0,1) \end{aligned}$$

and either statistic can be employed. Their asymptotic power envelope is straightforwardly shown to be of order $O\left(1/\left(TN^{1/2}\right)\right)$.

When heterogeneous trends are accounted for, test statistics are the same after a slight modification of (18). Anyway, no asymptotic power is present under local alternatives.

Monte Carlo evidence shows that almost in all cases t_{MP}^{b} outperforms t_{MP}^{a}; when k estimation is called for, according to Bai and Ng (2002), N should be at least equal to 20 to achieve a good estimate $\hat{k}$. When N is smaller, usually k is overestimated, which affects the size of Moon and Perron's test. The test performance is usually good, even in presence of cointegrated factors or linear heterogeneous trends when the true value of the ρ_is is not too close to 1. However, all experiments are conducted for $T = 100$ at least, and N never larger than 20, and it may be interesting to check the performance when both dimensions are large but comparable; moreover, the number of common factors considered in this set of experiments is at most equal to 2, and again it could be worth exploring the case of a larger factor model.

[15] Moon and Perron consider kernel estimates; in principle, however, any other consistent estimator would work.

[16] Limits are taken for $N \to \infty$.

2.5 A set of locally powerful tests

The set of tests proposed by Ploberger and Phillips (2002), Moon and Phillips (2003) and Moon, Perron and Phillips (2003) address the issue of deriving locally powerful tests to check the presence of a unit root in panels where no time or cross-sectional dependences are present.

The model considered in the three cases is the same as in equations (3) and (5), which could be written more conveniently in vector form as

$$\begin{aligned} Z &= \beta G' + Y \\ Y &= \bar{\rho} Y_{-1} + E \end{aligned}$$

where Z, Y, Y_{-1} and E have the same meaning as in (17), $\bar{\rho} = diag\{\rho_1, ..., \rho_N\}$, $\beta = (\beta_1', \beta_2')$, $\beta_1 = [\mu_1, ..., \mu_N]'$, $\beta_2 = [\theta_1, ..., \theta_N]'$ and $G = (i_T, t_T') = (g_1, ..., g_T)'$ with i_T a T-dimensional row vector of ones and $t_T = (1, ..., T)$. Therefore, incidental trends are allowed for. The error components ε_{it} are assumed to be iid and Gaussian across i and t with variance σ^2.

The computational aspects of the three tests can now be introduced.

2.5.1 Ploberger and Phillips's (2002) test

The statistics suggested by Ploberger and Phillips is constructed according to the following algorithm:

1. **Step 1**: estimate β as $\hat{\beta} = [\Delta Z \Delta G][\Delta G' \Delta G]^{-1}$, where $\Delta Z = (z_0, \Delta z_1, ..., \Delta z_T)$ and $\Delta G = (g_0, \Delta g_1, ..., \Delta g_T)$

2. **Step 2**: detrend Z obtaining the GLS residual $\hat{Y} = Z - \hat{\beta} G'$;

3. **Step 3**: derive a consistent estimate for σ^2,[17] namely $\hat{\sigma}^2$, and define $\omega_T^{PP} \equiv T^{-2} \sum_{t=1}^{T} t(1 - tT^{-1})$,

4. **Step 4**: compute the test statistics V_{NT}^{PP} as

$$V_{NT}^{PP} = \frac{\sqrt{N}}{\hat{\sigma}^2} \left[\frac{1}{NT^2} tr\left(\hat{Y} \hat{Y}' \right) - \omega_T^{PP} \hat{\sigma}^2 \right]. \tag{19}$$

[17] An example, considered by Moon and Phillips (2003), could be

$$\hat{\sigma}^2 = (NT)^{-1} tr\left(\hat{E}' \hat{E} \right)$$

where $\hat{E}$ is the residual from the regression of $\hat{Y}$ on $\hat{Y}_{-1}$.

The test rejects the null of a unit root for small values of V_{NT}^{PP}; when both N and T converge to infinite,[18] $V_{NT}^{PP} \Longrightarrow N(0, 1/45)$.

2.5.2 Moon and Phillips (2003)

Moon and Phillips propose a similar algorithm to the one reported above, using OLS instead of GLS to detrend the data. Let $P_G = I_T - G(G'G)^{-1} G'$ and $D_T = diag\{1, T^{-1}\}$; then:

1. **Step 1**: derive a consistent estimate $\hat{\sigma}^2$ of σ^2 and compute $\omega_T^{MP} = T^{-2}\left[\sum_{t=1}^{T} t - \sum_{t=1}^{T}\sum_{s=1}^{T} \min(t,s)\, h(t,s)\right]$, where

$$h(t,s) = g_t' (G'G)^{-1} g_s;$$

2. **Step 2**: compute the test statistics V_{NT}^{MP} as

$$V_{NT}^{MP} = \frac{\sqrt{N}}{\hat{\sigma}^2}\left[\frac{1}{NT^2} tr\left(Z P_G Z'\right) - \omega_T^{MP}\hat{\sigma}^2\right]. \tag{20}$$

Even in this case, the test rejects the null of a unit root for small values of V_{NT}^{MP}. Under the same assumptions for asymptotic theory, $V_{NT}^{MP} \Longrightarrow N(0, 11/6300)$.

2.5.3 Moon, Perron and Phillips (2003)

This test is explicitly designed for local alternatives of the form

$$\rho_i = \rho = 1 - \frac{c}{TN^{1/2}}. \tag{21}$$

with $c < 0$. The parameter c must be specified by the user, but numerical evidence shows that the test is only mildly sensitive to the choice of c; the authors suggests to set it equal to values between 1 and 2.

Define now the modified difference operator $\Delta_c = diag(1 - \rho L)$; then the algorithm to obtain Moon, Perron and Phillips statistic is:

1. **Step 1**: define

$$L(c) \equiv \min_{\beta} L_{NT}(c, \beta) = tr\left[\left(\Delta_c Z - \beta \Delta_c G'\right)\left(\Delta_c Z - \beta \Delta_c G'\right)'\right];$$

[18] The limit theory here differs from Philips and Moon's (1999) scheme, and it requires that, together with increasing, $N^{3/4}/T \to 0$. Practically, this means that T cannot be "too smaller" than N.

2. **Step 2**: construct ω_{1T}^{MPP} and ω_{2T}^{MPP}, defined as

$$\omega_{1T}^{MPP} = -T^2 \sum_{t=1}^{T} (t-1) + 2T^{-3} \sum_{t=1}^{T} (t-1)^2 - \frac{1}{3}$$

and

$$\omega_{2T}^{MPP} = T^{-5} \sum_{t=1}^{T} \sum_{s=1}^{T} (t-1)(s-1) \min [(t-1), (s-1)] - \frac{2}{3} T^{-3} \sum_{t=1}^{T} (t-1)^2 - \frac{1}{9},$$

and derive a consistent estimate $\hat{\sigma}^2$ of σ^2;

3. **Step 3**: compute the test statistic as

$$V_{NT}^{MPP} = \frac{1}{\hat{\sigma}^2} [L(c) - L(0)] + cN^{3/4} + c^2 \omega_{1T}^{MPP} N^{1/2} + c^4 \omega_{2T}^{MPP}. \quad (22)$$

Here too, the test rejects the null of a unit root for small values of V_{NT}^{MPP}. Under the same assumptions for asymptotic theory, $V_{NT}^{MP} \Longrightarrow N(0, c^4/45)$.

Moon, Perron and Phillips (2003) study the asymptotic power envelope for these tests. They show that not withstanding the presence of incidental trends in the model, all the three tests have nontrivial power against local alternatives; not surprisingly, the test derived by Moon and Phillips (2003) is dominated by the other two, that share the same power.

Monte Carlo evidence shows that power is anyway reduced by the presence of trends; theoretically, this corresponds to the well established result that the asymptotic power envelope when trends are present is shrunk towards $O\left(1/\left(TN^{1/4}\right)\right)$ from $O\left(1/\left(TN^{1/2}\right)\right)$. Also, Ploberger and Phillips' (2002) and Moon and Phillips' (2003) tests tend to underreject the null. It is also worth noticing that experiments are conducted for small values of N - the largest being 30. Even if this is consistent with the asymptotic theory derived, it could be worth exploring the case where N is larger.

A computational note. Even though this testing approach may seem a contradiction with the tests discussed so far, since it rules out the presence of cross sectional dependence, the statistical framework illustrated here can be made fit the case of panels with serial and cross sectional dependence. In principle, filtering out consistently these two characteristics should suffice. Therefore, these tests could be applied to a panel data model, provided that this is orthogonalized according to one of the procedures detailed above.

2.6 Bai and Ng (2004)

Bai and Ng (2004) provide a complete set of tools to test the degree of integration of each series y_{it}. The model employed to represent the panel is

$$\begin{aligned} y_{it} &= \lambda_i^{0\prime}\eta_t^0 + \varepsilon_{it} \\ z_{it} &= \mu_i + \delta_i t + y_{it} \\ (I-L)\,\eta_t^0 &= C(L)\,w_t \\ (1-\rho_i L)\,\varepsilon_{it} &= D_i(L)\,u_{it.} \end{aligned}$$

the number of factors is equal to k, and k_1 factors are assumed to be nonstationary, i.e. the rank of $C(1) = k_1$. Notice that heterogeneous linear trends and autocorrelation are explicitly modelled in this unified framework.

The issue of testing for unit roots here is addressed in a slightly different way than in the other papers: unit roots may be present in both the factor specification and in the error term ε_{it}. The two cases are conceptually different: if non stationarity is due to a common factor, this means that there is a nonstationary component that is specific to all the series in the panel. On the contrary, if nonstationarity arises from the idiosyncratic component ε_{it}, this can be regarded as a series specific source of integratedness.

To illustrate the testing procedure, consider first the case where the model contains only the intercept μ_i:

$$z_{it} = \mu_i + y_{it}.$$

Taking the first difference Δz_{it}, one can rewrite the model as

$$\Delta z_{it} = \lambda_i^{0\prime}\Delta\eta_t^0 + \Delta\varepsilon_{it} \tag{23}$$

and λ_i^0 and $\Delta\eta_t^0$ can be estimated using Bai and Ng's (2002) method sketched above, obtaining $\hat{\lambda}_i$ and $\Delta\hat{\eta}_t$. Let now $\Delta\hat{z}_{it} = \Delta z_{it} - \hat{\lambda}_i^{\prime}\Delta\hat{\eta}_t$, and define

$$\begin{aligned} \hat{e}_{it} &= \sum_{j=2}^{t}\Delta\hat{z}_{ij} \\ \hat{\eta}_t &= \sum_{j=2}^{t}\Delta\hat{\eta}_j. \end{aligned}$$

To test for a unit root in $\hat{e}_{it}$, Bai and Ng suggest using a standard ADF test; the associated statistics will be denoted as $ADF_e^c(i)$ for unit i. As far as detecting a unit root in $\hat{\eta}_t$ is concerned, this exercise is the same as estimating k_1. Two cases can be considered:

- either $k=1$, in which case one should apply ADF test to $\hat{\eta}_t$, obtaining ADF_η. This will tell whether $k_1 = 0$ or 1;
- or $k > 1$. In this case, after demeaning $\hat{\eta}_t$ to get $\hat{\eta}_t^c$ the following iterative algorithm can be employed:

  1. **Step 1**: set $m = k$;
  2. **Step 2**: compute the eigenvector space associated with the m largest eigenvalues of $T^{-2}\sum_{t=2}^{T}\hat{\eta}_t^c\hat{\eta}_t^{c\prime}$, say $\hat{\lambda}_\perp$;
  3. **Step 3**: test for $H_0 : k_1 = m$ using the variable $\hat{x}_t^c \equiv \hat{\lambda}_\perp'\hat{\eta}_t^c$. Two testing approaches can be considered:
     (a) **Step 3.1**:
        i. consider the Bartlett kernel $K(j) = 1 - j/(J+1)$, $j = 1, ..., J$, and let $\hat{\xi}_t^c$ be the residuals from estimating a first order VAR in $\hat{x}_t^c$;
        ii. construct the non parametric estimate $\hat{\Sigma}_c = \sum_{j=1}^{J} K(j)\left[T^{-1}\sum_{t=2}^{T}\hat{\xi}_{t-j}^c\hat{\xi}_t^{c\prime}\right]$;
        iii. compute

$$\hat{\Phi}_c^c(m) = \frac{1}{2}\left[\sum_{t=2}^{T}\left(\hat{x}_t^c\hat{x}_{t-1}^{c\prime} + \hat{x}_{t-1}^c\hat{x}_t^{c\prime}\right) - T\left(\hat{\Sigma}_c + \hat{\Sigma}_c'\right)\right]\left[\sum_{t=2}^{T}\hat{x}_{t-1}^c\hat{x}_{t-1}^{c\prime}\right]^{-1}$$

        and let $\hat{\nu}_c^c(m)$ be its smallest eigenvalue;
        iv. define the test statistic $MQ_c^c(m) = T\left[\hat{\nu}_c^c(m) - 1\right]$.
     (b) or, alternatively, **Step 3.2**:
        i. estimate a VAR of order p - to be user specified[19] - in $\hat{x}_t^c$ and construct the polynomial $\hat{\Pi}(L) = I_m - \sum_{k=1}^{p}\hat{\Pi}_k L^k$;
        ii. get the filtered $\hat{x}_t^{c*}$ as $\hat{x}_t^{c*} = \hat{\Pi}(L)\hat{x}_t^c$;
        iii. compute

$$\hat{\Phi}_c^f(m) = \frac{1}{2}\left[\sum_{t=2}^{T}\left(\hat{x}_t^{c*}\hat{x}_{t-1}^{c*\prime} + \hat{x}_{t-1}^{c*}\hat{x}_t^{c*\prime}\right)\right]\left[\sum_{t=2}^{T}\hat{x}_{t-1}^{c*}\hat{x}_{t-1}^{c*\prime}\right]^{-1}$$

[19] In order for the asymptotic theory to hold, the authors require that $\left[p^3/\min(N,T)\right] \to 0$. Practically, this means that one should choose p so that p^3 is smaller (of at least one order of magnitude) than $\min(N,T)$. A possible choice, suggested in the simulation exercise by the authors, could be

$$p = 4\left[\frac{\min(N,T)}{100}\right]^{1/4}.$$

and let $\hat{\nu}_c^f(m)$ be its smallest eigenvalue;

iv. define the test statistic $MQ_c^f(m) = T\left[\hat{\nu}_c^f(m) - 1\right]$.

4. **Step 4**: if H_0 is accepted, then $\hat{k}_1 = m$. Otherwise, set $m = k - 1$ and go back to Step 2.

If a linear trend is present, one should consider the demeaned version of Δz_{it}, $\Delta\eta_t^0$ and $\Delta\varepsilon_{it}$, namely Δz_{it}^*, $\Delta\eta_t^{0*}$ and $\Delta\varepsilon_{it}^*$ respectively. Then the same passages as before can be applied; only, in this case one should explicitly take account of the presence of a linear trend when using ADF or the corresponding algorithm to test for a rank reduction in $C(1)$.

In order to test for a unit root in the idiosyncratic component, Bai and Ng discuss the limit distributions (under the hypothesis that both T and N are large) for $ADF_e^c(i)$, ADF_η, and the statistics one could employ in Step 3 of the algorithm; results are obtained for the case where also a trend is present.

To obtain a test for the presence of a unit root in the idiosyncratic component, as is consistent with the spirit of the other contributions, Bai and Ng propose, on the basis of Choi (2001) and similarly to Phillips and Sul (2003), a meta-analytical approach. Let the p-values of $ADF_e^c(i)$ be denoted as π_i^c; then the test statistics

$$P_e^c = \frac{-2\sum_{i=1}^N \log\left(\pi_i^c\right) - 2N}{2\sqrt{N}},$$

can be employed to test for H_0 in (2). This statistic is found to be standard normally distributed.

Three issues are worth discussing as far as the asymptotic theory needed here is concerned. Firstly, both N and T are required to be large; this approach would not work for the short panel case (i.e. when T is finite), or for a panel with a limited number of units. Secondly, the way N and T increase is slightly different from the sequential/joint scheme. In this case, provided that both dimensions are large, any combination of order of magnitudes of them in principle would lead to the results sketched above. Last, it is worth emphasizing that neither idiosyncratic components nor factors are observed. To test for their stationarity, one has to employ their estimates; Bai and Ng show that these are consistent, and that therefore proxying unobserved variables with their estimated counterparts is a valid exercise.

Monte Carlo evidence provides some further insights. Firstly, $N = 20$ suffices to ensure consistency of the estimated factors. This result is due to the particular estimation method suggested here, that considers first differenced data rather than levels; in this latter case, consistency could not be ensured when the idiosyncratic component is $I(1)$. As far as tests are concerned, simulations show that tests have good characteristics when N is at least equal to 40, and also when multiple factors are considered. However, three relevant issues are not tackled in the set of exercises presented by Bai and Ng. First, it should be explored the size distortion and the power reduction when either N or T (or both) are small, to have an assessment of the finite sample performance of the test. Second, some further analysis ought to be done to check the tests performances when the number of factors k is large and comparable with $\min[N, T]$. Last, local power is not fully analyzed (only some results are present), and it could be worth exploring the tests capability of discriminating between integratedness and near integratedness.

2.7 Pesaran (2003)

Pesaran (2003) considers an approach that completely differs from the other ones in terms of asymptotic theory. The model considered is the same as in equations (3) and (13). Therefore, cross dependence in modelled according to a common factor representation, and only one factor is present in the model. Serial correlation is explicitly accounted for, writing the error term DGP as

$$\varepsilon_{it} = \sum_{j=1}^{p} \gamma_{ij} \varepsilon_{it-j} + \lambda_i \eta_t + u_{it}.$$

Like Bai and Ng (2004), dynamics is assumed in the factor structure, but integratedness is ruled out by assumption.

To present the computational details of Pesaran's statistics, appeal can be made to a result derived by Pesaran (2002), according to which the common factor η_t can be proxied by the cross sectional mean $\bar{y}_t = N^{-1} \sum_{i=1}^{N} y_{it}$ and its lagged values, provided that N is large and that $\bar{\lambda} \equiv N^{-1} \sum_{i=1}^{N} \lambda_i \neq 0$ for any N. Hence, model (3)-(13) can be rewritten in first differenced form as

$$\Delta y_{it} = a_i + b_i y_{it-1} + c_i \bar{y}_t + d_i \Delta \bar{y}_t + e_{it}. \tag{24}$$

The presence of unit root can be tested for, in this model where all the elements on the RHS are observable, using a standard t-statistics for the coefficient b_i, namely $t_i(N, T)$.[20] In case of autocorrelation, equation (24) can be augmented, plugging in terms like Δy_{it-j} and $\Delta \bar{y}_{t-j}$ to take account of serial dependence in ε_{it} and in the factor respectively. Similarly to a Dickey-Fuller framework, the presence of fixed effects and trends can be accounted for by demeaning and detrending the variable y_{it}.

[20] The only slight modification suggested by Pesaran is to estimate the variance of the residual using $(T - k + 1)$ in the denominator, k being the number of regressors.

Under the null hypothesis, Pesaran derives the asymptotic distribution for $t_i(N,T)$ when $N \to \infty$, for all the considered cases (with or without deterministic components, with or without serial correlation). For both the large[21] and finite T cases, $t_i(N,T)$ is shown not to depend on nuisance parameters, and even though its distribution is non standard, its critical values can be computed once and for all. Therefore, this statistical framework fits the case when N is large, irrespectively of the dimension of T that can be as small as 4. The statistic $t_i(N,T)$ is the building block needed to test for H_0 in (2). Even if in principle the $t_i(N,T)$s could be employed to construct meta-analysis statistic, Pesaran focuses on developing a framework similar to that in Im, Pesaran and Shin (2003); the proposed statistics is

$$D_{NT} = \frac{1}{N}\sum_{i=1}^{N} t_i(N,T)$$

where $\bar{t}$ is the stochastic limit for $t_i(N,T)$ when both N and T are large. Again, the distribution is non standard but does not depend on nuisance parameters; Pesaran provides a procedure to compute its critical values.

Monte Carlo evidence for small sample performance shows that the test seems to have no size or power problems when $T > 20$, while its power becomes very low when T is smaller than 10, regardless of the value of N. However, it should be stressed that the simulation exercises are restricted to the case when a single factor model is the correct specification; it could be worth exploring the test's size and power when there is more than one factor in the true DGP.

3 Monte Carlo evidence

This Section considers a broad extension of the simulation results already existing in the literature. As stated in the previous section, some features of the tests are still to be explored, and from the practitioner's viewpoint it is not always clear which testing procedure should be employed given the data at hand. In this section, we will first comment the design of our Monte Carlo experiment, and subsequently we will report the results, providing some guidelines to choose among tests.

[21] The limit theory employed here is both the joint limit theory and the sequential one, where N is allowed to increase before T.

3.1 The design of the Monte Carlo experiment

The DGP that we will employ for our simulation exercise follows directly from the literature:

$$\begin{aligned} y_{it} &= \mu_i + \theta_i t + \rho_i y_{it-1} + \varepsilon_{it} && (25) \\ (1-\alpha_i L)\,\varepsilon_{it} &= \lambda_i' \eta_t + (1+\beta_i L)\,u_{it} && (26) \\ \eta_t &= \Phi \eta_{t-1} + w_t, && (27) \end{aligned}$$

where η_t is a k-dimensional vector, $\Phi = diag\,\{\phi_i\}$, $i = 1,..,N$, $t = 1,...,T$ and as far as the model's parameters are concerned, the following assumptions will be made:

1. μ_i and θ_i represent the deterministic components in each unit of the panel. While the presence of μ_i is not proved to have a strong impact on the test performance (which is affected by it only via the demeaning scheme employed), θ_i is well known to be of importance when one needs to evaluate the local power of a unit root test. Following the same approach as in Moon and Perron (2003), both parameters, when present, will be generated as iid draws from a standard normal distribution:

   $$\begin{aligned} \mu_i &\sim iidU\,(-a, a) \\ \theta_i &\sim iidU\,(-b, b)\,; \end{aligned}$$

   where a and b will be chosen so as to control for the parameters heterogeneity;

2. the autoregressive root ρ_i will be simulated according to two different schemes:

   (a) as far from unity heterogeneous parameters, to assess size and power in the smal sample case without considering the issue of local power. Exploring this case is necessary since resulst are still missing under many combinations of (N, T), especially when the former is "large". In such case, following the same approach as in the literature, ρ_i under the alternative will be randomly drawn from a uniform distribution:

   $$\rho_i \sim iidU\,[0.5, 0.9]\,;$$

   (b) to assess size and especially power when the alternative is local to unity. This is necessary to rank the various tests when both

(N,T) are large, since most of the reviewed paper do not tackle this issue, and to do the same exercise in the finite sample case. Following the comprehensive study by Moon, Perron and Phillips (2003), this set of simulation exercise will be worked out with respect to the following representation

$$\rho_i = 1 - \frac{c_i}{TN^\eta}$$

where $\eta = 1/2$ when no incidental trends are present and $\eta = 1/4$ when they are considered; the c_is will be drawn from the following distributions[22]:

$$\begin{aligned} c_i &\sim iidU\left[0,2\right] \\ c_i &\sim iidU\left[0,4\right] \\ c_i &\sim iidU\left[0,8\right] \\ c_i &\sim iid\chi^2\left(1\right) \\ c_i &\sim iid\chi^2\left(2\right) \\ c_i &\sim iid\chi^2\left(4\right) \\ c_i &= c \sim iidU\left[0,2\right] \\ c_i &= c \sim iid\chi^2\left(1\right); \end{aligned}$$

3. as far as equation (26) is concerned, two issues are of importance:

    (a) the error's dynamics is shown to play an important role in assessing the tests power, as proved by Phillips and Sul (2003). Therefore, the ARMA structure we allow for will be simulated according to three scenarios:

    - firstly, we will consider a low and high autocorrelated AR(1) structure with no MA effects

    $$\begin{aligned} \alpha_i &\sim iidU\left[0,0.4\right] \\ \alpha_i &\sim iidU\left[0.5,0.9\right]; \end{aligned}$$

    - secondly, we will consider a richer autoregressive structure representing it via the MA polynomial $(1-\beta_i L)$

    $$\begin{aligned} \beta_i &\sim iidU\left[0,0.4\right] \\ \beta_i &\sim iidU\left[0.5,0.9\right]; \end{aligned}$$

[22] The rationale behind this set of exercises is exposed in Moon, Perron and Phillips (2003).

- lastly, we will consider the case of negative MA coefficients[23]

$$\begin{aligned}\beta_i &\sim iidU\,[0, -0.4] \\ \beta_i &\sim iidU\,[-0.5, -0.9]\,;\end{aligned}$$

(b) also, the structure allowed for the factors η_t is important. Firstly, several tests consider the presence of only one factor, and it could be worth exlporing up to which point considering two factors, one of which "less important" (i.e. with a smaller loading vector λ_i) can invalidate the results in terms of power and size. Secondly, some evidence on the factors dynamics will be provided, generating the ϕ_is as

$$\begin{aligned}\phi_i &\sim iidU\,[0.4, 0.8] \\ \phi_i &\sim iidU\,[0.9, 0.95]\,.\end{aligned}$$

This exercise is also needed in the light of the conclusions drawn by Bai and Ng (2004), who show how the variance of the process that generates the factors has an impact on the test's size;

4. last, since the literature lacks evidence on the short panel case, some simulations will be devoted to this issue as well. The couples (N, T) that we plan to study are: $(N, T) \in \{5, 10, 20, 50, 100, 200\} \times \{5, 10, 20, 50, 100, 200\}$.

4 Simulation results

Here we report some of the outcomes of our Monte Carlo exercises; these are grouped in the following classes:

1. studies on the local power of unit root tests;

2. studies on asymptotic size and power;

3. studies on small sample size and power.

The tests we compare in our simulations are:

1. Maddala and Wu's (1999) Fisher's type tests; we consider both the version in equation (14) and the one in equation (15);

[23]Following the findings in Phillips and Sul (2003).

2. Im, Pesaran and Shin (2003) test under the hypothesis of no cross sectional dependence;

3. Pesaran's (2004) test;

4. Breitung and Das's (2004) statistics - both t_{BD} and t_{gls};

5. Chang's (2002) IV based test;

6. three tests based on Philips and Sul's (2003) orthogonalization procedure: we first apply Im, Pesaran and Shin's (2003) test on the orthogonalised data and then the two Fisher's type tests reportee in equations (14) and (15);

7. Moon and Perron's (2003) tests;

8. the three tests by Phillips and Ploberger (2002), Moon and Phillips (2003) and Moon, Perron and Phillips (2003);

9. Bai and Ng's (2004) test.

In total, we compare 16 different tests.

4.0.1 The local power of unit root tests

As the literature has pointed out, a difficulty of testing for unit roots in panels is related with the possible absence of local to unity power when incidental (i.e. heterogeneous) trends are present. In such case, test often don't have power and tend to accept the null of a unit root when instead only near integration is present. In this section, we studied two separate issues:

1. the local to unity power when no incidental trends are present (as a benchmark for the subsequent case);

2. the local to unity power when there are heterogeneous trends.

Local to unity power when no trends are present

Local to unity power with incidental trends The results reported in Tables 1.1.a-1.4.b are derived under four different scenarios:

- the case where ε_{it} is white noise;

- the case where a positive AR root is present, setting $\alpha_i \sim iidU\,[0.5, 0.9]$ in (26);

- the case where a negative MA root is present,setting $\beta_i \sim iidU[-0.7,-0.9]$ in (26);

- the ARMA case, with $\alpha_i \sim iidU[0.5,0.9]$ and $\beta_i \sim iidU[0.5,0.9]$.

No dynamic structure is considered for the factors in this set of simulations, and two different levels of cross sectional dependence are modelled: λ_i in equation (26) is simulated as a uniformly distributed random variable whose support is chosen equal to $[0,0.2]$ to model small cross section dependence and $[-1,3]$ to represent a large amount of cross section dependence. The main results can be summarised as follows:

- meta analytical tests provide excellent results in term of power, and seem to be almost perfectly able to discriminate the presence of a nearly unit root. This should suggest that their power envelope is probably less shrunk than the one of other techniques, which is at most $O\left(TN^{-1/4}\right)$;

- tests that are not explicitly designed to take account of possible near integration, such as Im, Pesaran and Shin's (2003) or Pesaran's (2004) statistics have virtually no power against local alternatives. This also happens when cross sectional dependence is removed according to Phillips and Sul's (2002) orthogonalisation scheme, and it happens irrespectively of the presence and degree of cross sectional dependence;

- our results confirm the findings of Moon and Perron (2003), whose tests are shown to have no power against local to unity alternatives when heterogeneous trends are present;

- the three locally powerful procedures presented here - namely Phillips and Ploberger's, Moon and Perron's, Moon, Perron and Phillips' - seem to have good power when ε_{it} is white noise. A remarkable exception is Moon, Perron and Phillips's testing scheme, which has virtually no power. This apparently counterintuitive result, which can also be proved not to hold analytically, shows that the dependence of the results of this test on the choice of c is rather strong, and therefore attention should be paid when calibrating the test;

- Bai and Ng test performs quite well, especially when the cross sectional dimension N increases, and irrespectively of the error term dynamics;

- it is worth noticing that the choice of (N,T) does not affect results very much. This is because the choice of the values for the local to unity roots ρ_i are tailored on the basis of the values of N and T, in order to make them invariant of either dimension;

- the impact of AR and ARMA dynamics is particularly noticeable on Phillips and Ploberger's, Moon and Perron's, Moon, Perron and Phillips' tests. These are not designed for error terms ε_{it} that are not white noise, and when such hypothesis is violated there is a substantial loss of power. On the contrary, tests that are explicitly designed to take account of serial dependence and to adjust the data for it are almost not affected by the presence of dynamics in ε_{it}. Also, it is worth stressing that non parametric tests based on Fisher's transformation are not affected by ARMA components;

- negative MA coefficients seem to have a striking impact on unit root tests. Even if results from meta analysis confirm that this does not depend on the error dynamics, other tests that have no local power become very powerful in this case - this is for example the case of Chang's (2002) IV based test and Pesaran's (2004) test, which needs a large number of time observations (larger than 20), to achieve non trivial power.

References

[1] Arellano, M. (1987), "Computing Robust Standard Errors for Within-Groups Estimators", *Oxford Bulletin of Economics and Statistics*, 49, 431-434.

[2] Bai, J. and Ng, S. (2002), "Determining the Number of Factors in Approximate Factor Models", *Econometrica*, 70, 191-221.

[3] Bai, J. and Ng, S. (2004), "A PANIC Attack on Unit Roots and Cointegration", forthcoming in *Econometrica*.

[4] Baltagi, B.H. and Kao, C. (2000), "Nonstationary Panels, Cointegration in Panels and Dynamic Panels: a Survey", *Advances in Econometrics*, 15, 7-51.

[5] Banerjee, A. (1999), "Panel Data Unit Roots and Cointegration: an Overview", *Oxford Bulletin of Economics and Statistics*, 61, 607-629.

[6] Bornhorst, F. (2002), "A Note on the Use of Panel Unit Root Tests on Sectional Dependent Data: An Application to PPP", manuscript, European University Institute.

[7] Breitung, J and Das, S. (2003), "Panel Unit Root Tests Under Cross Sectional Dependence", manuscript, University of Bonn.

[8] Chang, Y. (2002), "Nonlinear IV Unit Root Tests in Panels with Cross-Section Dependency", *Journal of Econometrics*, 110, 261-292.

[9] Chang, Y. (2003), "Bootstrap Unit Root Tests in Panels with Cross-Section Dependency", forthcoming in *Journal of Econometrics*.

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

[11] Hall, B.H., Mairesse, J. (2001), "Testing for Unit Roots in Panel Data: An Exploration Using Real and Simulated Data", manuscript, UC Berkeley.

[12] Im, K., Pesaran, M.H. (2003), "On the Panel Unit Root Tests Using Nonlinear Instrumental Variables", manuscript, Cambridge University.

[13] Im, K., Pesaran, M.H., Shin, Y. (2003), "Testing for Unit Roots in Heterogeneous Panels", *Journal of Econometrics*, 115, 53-74.

[14] Ioannidis, C., Peel, D.A., Venetis, I.A. (2003), "On the Persistence of Profits: Some New Results", manuscript, Cardiff Business School.

[15] Levin, A. and Lin, C.F. (1992), "Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties" UCSD, Discussion Paper No. 92-93.

[16] Levin, A. and Lin, C.F. (1993), "Unit Root Tests in Panel Data: New Results" UCSD, Discussion Paper No. 93-56.

[17] MacDonald, R. (1996), "Panel Unit Root Tests and Real Exchange Rates", *Economics Letters*, 50, 7-11.

[18] Maddala. G.S. and 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, 631-652.

[19] Moon, H.R and Perron, P. (2003), Testing for a Unit Root in Panels with Dynamic Factors", forthcoming in *Journal of Econometrics.*

[20] Moon, H.R and Perron, P. (2003b), "Panel Evidence on Unit Roots in Exchange Rates and Interest Rates with Cross-Sectional Dependence", manuscript, USC.

[21] Moon, H.R, Perron, P., Phillips, P.C.B. (2003), "Incidental Trends and the Power of Panel Unit Root Tests", manuscript, Cowles Foundation.

[22] Moon, H.R and Phillips, P.C.B. (2003), "GMM Estimation of Autoregressive Roots Near Unity with Panel Data", forthcoming in *Econometrica.*

[23] O'Connell, P.G.J. (1998), "The Overvaluation of Purchasing Power Parity", *Journal of International Economics*, 44, 1-19.

[24] Park, J.Y. and Phillips, P.C.B. (1999), "Asymptotics for Nonlinear Transformations of Integrated Time Series", *Econometric Theory*, 15, 269-298.

[25] Park, J.Y. and Phillips, P.C.B. (2001), "Nonlinear Regression with Integrated Time Series", *Econometrica*, 69, 117-162.

[26] Pesaran, M.H. (2002), "Estimation and Inference in Large Heterogeneous Panels with Cross Section Dependence", *University of Cambridge DAE Working Paper No. 0305* and *CESifo Working Paper Series No. 869, October 2002.*

[27] Pesaran, M.H. (2003), "A Simple Panel Unit Root Test in the Presence of Cross Section Dependence", manuscript, Cambridge University.

[28] Phillips, P.C. B. and Moon, H.R. (1999), "Linear Regression Limit Theory for Nonstationary Panel Data", *Econometrica*, 67, 1057-1111.

[29] Phillips, P.C. B. and Schmidt, P. (1992), "Testing for a Unit Root in the Presence of Deterministic Trends", *Oxford Bulletin of Economics and Statistics*, 54, 257-287.

[30] Phillips, P.C. B and Sul, D. (2003), "Dynamic Panel Estimation and Homogeneity Testing under Cross-Section Dependence", *Econometrics Journal*, 6, 217-259.

[31] Ploberger, W. and Phillips, P.C.B. (2002), "Optimal Testing for Unit Roots in Panel Data", manuscript, Yale University.

[32] Quah, (1994), "Exploring Cross-Section Variations for Unit Root Inference in Dynamic Data", *Economics Letters*, 44, 9-19.

APPENDIX: TABLES

Local to unity power when incidental trends are present

	$(10,20)$	$(10,50)$	$(10,100)$	$(20,50)$	$(20,100)$	$(50,100)$
MWu1	1.000	1.000	1.000	1.000	1.000	1.000
MWu2	1.000	1.000	1.000	1.000	1.000	1.000
IPS	0.000	0.000	0.000	0.000	0.000	0.000
P	0.000	0.000	0.000	0.000	0.000	0.000
BD1	0.000	0.000	0.000	0.000	0.000	0.000
BD2	0.001	0.000	0.000	0.000	0.000	0.000
Ch	0.686	0.022	0.063	0.025	0.055	0.020
PSul1	0.000	0.000	0.000	0.000	0.000	0.000
PSul2	1.000	1.000	1.000	1.000	1.000	1.000
PSul3	0.997	1.000	1.000	1.000	1.000	1.000
MP1	0.014	0.012	0.012	0.000	0.000	0.000
MP2	0.009	0.008	0.007	0.000	0.000	0.000
PP	0.994	0.998	0.995	1.000	1.000	1.000
MP	0.996	1.000	0.992	1.000	1.000	1.000
MPP	0.109	0.007	0.005	0.000	0.000	0.000
BN	0.913	0.991	0.999	1.000	1.000	1.000

Table 1.1.a. Case of low cross sectional dependence; ε_{it} white noise.

	$(10,20)$	$(10,50)$	$(10,100)$	$(20,50)$	$(20,100)$	$(50,100)$
MWu1	0.957	1.000	1.000	1.000	1.000	1.000
MWu2	0.988	1.000	1.000	1.000	1.000	1.000
IPS	0.000	0.000	0.000	0.000	0.000	0.000
P	0.000	0.000	0.000	0.000	0.000	0.000
BD1	0.009	0.000	0.000	0.000	0.000	0.000
BD2	0.081	0.001	0.000	0.000	0.000	0.000
Ch	0.254	0.116	0.169	0.161	0.216	0.193
PSul1	0.000	0.000	0.000	0.000	0.000	0.000
PSul2	0.999	1.000	1.000	1.000	1.000	1.000
PSul3	1.000	1.000	1.000	1.000	1.000	1.000
MP1	0.012	0.005	0.018	0.000	0.000	0.000
MP2	0.004	0.002	0.009	0.000	0.000	0.000
PP	0.996	1.000	1.000	1.000	1.000	1.000
MP	0.996	1.000	1.000	1.000	1.000	1.000
MPP	0.023	0.000	0.000	0.000	0.000	0.000
BN	0.817	0.911	0.928	0.970	0.967	0.987

Table 1.1.b. Case of large cross sectional dependence; ε_{it} white noise.

	$(10,20)$	$(10,50)$	$(10,100)$	$(20,50)$	$(20,100)$	$(50,100)$
MWu1	1.000	1.000	1.000	1.000	1.000	1.000
MWu2	0.999	1.000	1.000	1.000	1.000	1.000
IPS	0.000	0.001	0.001	0.000	0.000	0.000
P	0.000	0.000	0.000	0.000	0.000	0.000
BD1	0.021	0.000	0.000	0.000	0.000	0.000
BD2	0.048	0.000	0.000	0.000	0.000	0.000
Ch	0.058	0.033	0.002	0.000	0.000	0.000
PSul1	0.000	0.000	0.000	0.000	0.000	0.000
PSul2	0.999	1.000	1.000	1.000	1.000	1.000
PSul3	0.992	1.000	1.000	1.000	1.000	1.000
MP1	0.017	0.012	0.017	0.000	0.000	0.000
MP2	0.013	0.010	0.011	0.000	0.000	0.000
PP	0.582	0.155	0.019	0.644	0.196	0.002
MP	0.435	0.080	0.005	0.469	0.066	0.000
MPP	0.000	0.000	0.000	0.000	0.000	0.000
BN	0.282	0.487	0.451	0.964	0.964	0.987

Table 1.2.a. Case of low cross sectional dependence; ε_{it} modelled as an AR(1) process.

	$(10,20)$	$(10,50)$	$(10,100)$	$(20,50)$	$(20,100)$	$(50,100)$
MWu1	0.971	0.999	1.000	1.000	1.000	1.000
MWu2	0.987	1.000	1.000	1.000	1.000	1.000
IPS	0.000	0.000	0.000	0.000	0.000	0.000
P	0.000	0.000	0.000	0.000	0.000	0.000
BD1	0.037	0.000	0.000	0.000	0.000	0.000
BD2	0.171	0.005	0.000	0.000	0.000	0.000
Ch	0.137	0.035	0.029	0.027	0.014	0.015
PSul1	0.006	0.002	0.001	0.000	0.000	0.000
PSul2	0.996	1.000	1.000	1.000	1.000	1.000
PSul3	0.969	1.000	1.000	1.000	1.000	1.000
MP1	0.012	0.008	0.018	0.000	0.000	0.000
MP2	0.003	0.004	0.008	0.000	0.000	0.000
PP	0.715	0.466	0.353	0.447	0.302	0.003
MP	0.747	0.433	0.263	0.491	0.220	0.000
MPP	0.000	0.000	0.000	0.000	0.000	0.000
BN	0.655	0.820	0.864	0.935	0.953	0.987

Table 1.2.b. Case of large cross sectional dependence; ε_{it} modelled as an AR(1) process.

	$(10,20)$	$(10,50)$	$(10,100)$	$(20,50)$	$(20,100)$	$(50,100)$
MWu1	0.056	0.621	0.949	1.000	1.000	1.000
MWu2	0.075	0.999	1.000	1.000	1.000	1.000
IPS	0.021	0.000	0.000	0.000	0.000	0.000
P	0.090	0.999	1.000	1.000	1.000	1.000
BD1	0.000	0.000	0.000	0.000	0.000	0.000
BD2	0.011	0.000	0.000	0.000	0.000	0.000
Ch	1.000	1.000	1.000	1.000	1.000	1.000
PSul1	0.096	0.125	0.125	0.026	0.011	0.019
PSul2	0.299	0.623	0.774	0.984	0.993	1.000
PSul3	0.151	0.866	0.995	0.999	1.000	1.000
MP1	0.000	0.000	0.000	0.000	0.000	0.000
MP2	0.000	0.001	0.000	0.000	0.000	0.000
PP	1.000	1.000	1.000	1.000	1.000	1.000
MP	1.000	1.000	1.000	1.000	1.000	1.000
MPP	1.000	1.000	1.000	0.002	0.004	0.000
BN	0.992	1.000	1.000	1.000	1.000	1.000

Table 1.3.a. Case of low cross sectional dependence; ε_{it} modelled as an MA(1) process.

	$(10,20)$	$(10,50)$	$(10,100)$	$(20,50)$	$(20,100)$	$(50,100)$
MWu1	0.146	0.562	0.892	1.000	1.000	1.000
MWu2	0.147	0.895	1.000	1.000	1.000	1.000
IPS	0.056	0.007	0.005	0.001	0.000	0.000
P	0.036	0.225	0.351	0.073	0.105	0.059
BD1	0.003	0.000	0.000	0.000	0.000	0.000
BD2	0.047	0.000	0.000	0.000	0.000	0.000
Ch	0.862	0.965	0.980	0.994	0.996	0.999
PSul1	0.078	0.099	0.087	0.018	0.012	0.013
PSul2	0.288	0.632	0.796	0.987	0.997	1.000
PSul3	0.135	0.855	1.000	1.000	1.000	1.000
MP1	0.000	0.000	0.000	0.000	0.000	0.000
MP2	0.000	0.000	0.000	0.000	0.000	0.000
PP	1.000	1.000	1.000	1.000	1.000	1.000
MP	1.000	1.000	1.000	1.000	1.000	1.000
MPP	0.996	0.999	1.000	0.004	0.000	0.000
BN	0.981	0.989	0.982	0.999	0.998	0.999

Table 1.3.b. Case of large cross sectional dependence; ε_{it} modelled as an MA(1) process.

	$(10,20)$	$(10,50)$	$(10,100)$	$(20,50)$	$(20,100)$	$(50,100)$
MWu1	0.985	1.000	1.000	1.000	1.000	1.000
MWu2	0.978	1.000	1.000	1.000	1.000	1.000
IPS	0.000	0.000	0.003	0.000	0.000	0.000
P	0.000	0.000	0.000	0.000	0.000	0.000
BD1	0.066	0.001	0.000	0.000	0.000	0.000
BD2	0.122	0.002	0.000	0.000	0.000	0.000
Ch	0.164	0.040	0.000	0.002	0.000	0.000
PSul1	0.001	0.000	0.000	0.000	0.000	0.000
PSul2	0.992	1.000	1.000	1.000	1.000	1.000
PSul3	0.964	1.000	1.000	1.000	1.000	1.000
MP1	0.021	0.010	0.022	0.000	0.000	0.000
MP2	0.011	0.009	0.019	0.000	0.000	0.000
PP	0.501	0.083	0.010	0.200	0.014	0.009
MP	0.336	0.039	0.003	0.060	0.000	0.000
MPP	0.000	0.000	0.000	0.000	0.000	0.000
BN	0.534	0.282	0.198	0.877	0.849	0.899

Table 1.4.a. Case of low cross sectional dependence; ε_{it} modelled as an ARMA(1,1) process.

	$(10,20)$	$(10,50)$	$(10,100)$	$(20,50)$	$(20,100)$	$(50,100)$
MWu1	0.964	1.000	1.000	1.000	1.000	1.000
MWu2	0.981	1.000	1.000	1.000	1.000	1.000
IPS	0.000	0.000	0.001	0.000	0.000	0.000
P	0.000	0.000	0.000	0.000	0.000	0.000
BD1	0.055	0.000	0.000	0.000	0.000	0.000
BD2	0.219	0.011	0.000	0.000	0.000	0.000
Ch	0.082	0.023	0.013	0.008	0.001	0.003
PSul1	0.004	0.008	0.007	0.000	0.000	0.000
PSul2	0.982	1.000	1.000	1.000	1.000	1.000
PSul3	0.945	1.000	1.000	1.000	1.000	1.000
MP1	0.023	0.012	0.018	0.000	0.000	0.000
MP2	0.010	0.006	0.010	0.000	0.000	0.000
PP	0.461	0.151	0.065	0.061	0.008	0.009
MP	0.482	0.119	0.019	0.040	0.002	0.000
MPP	0.000	0.000	0.000	0.000	0.000	0.000
BN	0.617	0.787	0.832	0.939	0.952	0.976

Table 1.4.b. Case of large cross sectional dependence; ε_{it} modelled as an ARMA(1,1) process.