Panel Unit Root Tests in the Presence of Cross-Sectional Dependencies: Comparison and Implications for Modelling

Christian Gengenbach, Franz C. Palm, Jean-Pierre Urbain*

Department of Quantitative Economics,
Universiteit Maastricht, The Netherlands

First version: September 16, 2004
This version: March 12, 2008

Abstract

Several panel unit root tests that account for cross section dependence using a common factor structure have been proposed in the literature recently. Pesaran's (2007) cross-sectionally augmented unit root tests are designed for cases where cross-sectional dependence is due to a single factor. The Moon and Perron (2004) tests which use defactored data is similar in spirit but can account for multiple common factors. The Bai and Ng (2004a) tests allow to determine the source of non-stationarity by testing for unit roots in the common factors and the idiosyncratic factors separately. Breitung and Das (2008) and Sul (2007) propose panel unit root tests when cross-section dependence is present possibly due to common factors, but the common factor structure is not fully exploited.

This paper makes four contributions: (1) it compares the testing procedures in terms of similarities and difference in the data generation process, tests, null and alternative hypotheses considered, (2) using Monte Carlo results it compares the small sample properties of the tests in models with up to two common factors, (3) it provides an application which illustrates the use of the tests, and (4) finally it discusses the use of the tests in modelling in general.

*Universiteit Maastricht, Faculty of Economics and Business Administration, P.O. Box 616, 6200 MD Maastricht, The Netherlands. Corresponding author: Christian Gengenbach, email: C.Gengenbach@ke.unimaas.nl. The authors would like to thank Jörg Breitung and four anonymous referees for helpful comments and suggestions. The usual disclaimer applies.

1 Introduction

For many economic applications it is important to know whether an observed time series is stationary or non-stationary. For example, to test the validity of Purchasing Power Parity (PPP) one should examine the properties of the real exchange rates. One needs to look at the behavior of differences in real per capita output growth to test for growth convergence. Therefore, unit root tests are an important tool for econometric analysis. However, univariate unit root tests are known to lack power for samples of small or medium size. Unfortunately, for many macroeconomic variables data is available only for a small sample span. But, since studies investigating for example PPP or growth convergence are concerned with the behavior of similar data series from several countries, a natural attempt is to pool the information contained in a data panel. Indeed, that is the general idea of panel unit root tests, and they only differ in the way the information is pooled. Unfortunately, simple pooling is only valid if the units of the panel are independent of each other and sufficiently homogenous. Independence however is unlikely to hold in most applications of panel unit root tests. In cross-country analysis there might be common influences to all panel members, e.g. in PPP-studies one usually uses a common numeraire country to calculate real exchange rates.

In early approaches to panel unit root testing, the often unrealistic assumption of cross-sectional independence is made. For instance, the tests proposed in Levin, Lin and Chu (2002) and Im, Pesaran and Shin (2003), denoted respectively as LLC and IPS, assume cross-sectional independence, but allow for heterogeneity of the form of individual deterministic effects (constant and/or linear time trend) and heterogenous serial correlation structure of the error terms. Both methods test the same null hypothesis of non-stationarity, but differ in terms of the considered alternative and hence, in the way information is pooled. Levin, Lin and Chu (2002) study balanced panels with N cross-sectional units and T time series observations. They assume a homogenous first order autoregressive parameter and their test is based on the pooled t-statistic of the estimator. Im, Pesaran and Shin (2003) allow unbalanced panels with N cross sectional units and T_i time series observations for each $i = 1, \ldots, N$. They propose a standardized average of individual ADF statistics to test the pooled unit root null hypothesis against a heterogenous alternative. Both methods assume cross-sectional independence among panel units except for a common time effect. In that case, the derived results remain valid if cross-sectional averages are subtracted from the data.

Attention has been drawn recently to the assumption of cross-sectional independence on which the asymptotic results of both procedures rely. Among the first to analyze the effect of cross-sectional correlation on panel unit root tests was O'Connell (1998). Using Monte Carlo simulations he shows that the LLC test severely suffers from cross-correlation in terms of increased size and reduced power. He suggests using FGLS estimation to overcome this problem. However, estimation of the error covariance matrix becomes infeasible as N and T grow large. Flôres, Jorion, Preumont and Szafarz (1999) use SUR estimation of the (possibly heterogenous) AR parameter, and determine critical values for their test via

Monte Carlo simulations. Their methodology has the disadvantage that it requires extensive simulations to determine critical values and does only account for contemporaneous cross-sectional correlation. In simulation studies, Banerjee, Marcellino and Osbat (2004, 2005) assess the finite sample performance of panel unit root and cointegration tests when panel members are cross-correlated or even cross-sectionally cointegrated[1]. Their finding is, that all methods experience size distortions when panel members are cointegrated. This means that procedures such as the LLC or IPS test would over-reject the non-stationarity null when there are common sources of non-stationarity. This is analytically confirmed by Lyhagen (2000).

Recently, panel unit root tests have been proposed model cross-sectional correlation using a common factor representation of the data, or robust methods allowing for a general form of cross-sectional dependence, e.g. Chang (2002). The purpose of this paper is to study some of the new methods which assume a factor structure and compare them in terms of modelling, assumptions and statistical properties of the test statistics. A Monte Carlo study assesses the finite sample properties of the test statistics in terms of size and power in order to compare them.

Three different newly proposed unit root tests will be considered. Pesaran (2007) suggests a cross-sectionally augmented Dickey-Fuller (CADF) test where the standard DF regressions are augmented with cross-sectional averages of lagged levels and first differences of the individual series. He also considers a cross-sectional augmented IPS (CIPS) test, which is a simple average of the individual CADF-tests. The data generating process (DGP) is a simple dynamic linear heterogenous panel data model. The error term is assumed to have an unobserved one-common-factor structure accounting for cross-sectional correlation and an idiosyncratic component.

A second type of panel unit root tests has been proposed by Moon and Perron (2004). We consider two feasible t-statistics proposed by them to test for unit roots in a dynamic panel model allowing for fixed effects. The stationary error term follows a K-unobserved-common-factor model to which an idiosyncratic shock is added. The t-statistics are based on appropriately standardized pooled estimators of the first order serial correlation coefficients of the data series.

The third type of panel unit root tests has been proposed by Bai and Ng (2004a). In their "Panel Analysis of Non-stationarity in Idiosyncratic and Common Components" (PANIC) approach the space spanned by the unobserved common factors and idiosyncratic disturbances is consistently estimated without knowing whether they are stationary or integrated. Next, the number of independent stochastic trends driving the common factors is determined. Both individual and pooled individual statistics are proposed to test separately for unit roots in the unobserved common and idiosyncratic components of the data instead of the observed series. Both common and idiosyncratic components may be stationary or integrated.

These three panel unit root tests have been selected for the following reasons. First of all,

[1]The notation of panel cointegration tests refers to tests for cointegration between several variables of one panel member, in contrast to cointegration between panel members.

the model specifications are sufficiently close to each other and some are partly nested to allow for comparison. At the same time, the test procedures differ in important ways to make it interesting to compare their properties and provide some guidelines for the empirical analysis of non-stationary panel data. Second, in all the approaches an unobserved common factor structure is assumed to explain cross-sectional correlation. Common factor structures have several advantages. Statistical estimation and testing methods, and selection procedures for the number of factors are at the disposal of the empirical researcher. The statistical properties of these procedures are in general well-understood. These method recently experienced a revival in the common features literature. Using common factors to explain cross-sectional correlation allows to deal with the curse of dimensionality problem in a natural way, which has been found to work well in empirical econometrics. Finally, common factor structures often result from theoretical considerations in economics. For instance the CAPM and the APT models used in finance are common factor models, and many intertemporal microeconomic models imply factor structures for the data.

The paper is organized as follows: In Section 2 we present the DGPs used in the three approaches mentioned above. Wherever one DGP is nested in another this will be pointed out. Also, the testing procedures used will be described in some detail. We briefly discuss which features of the three approaches will be compared. In Section 3, we present the results of an extensive simulation study which compares the three approaches to panel unit root testing for models with factor structures and two panel unit root tests proposed by Breitung and Das (2008) and by Sul (2007) which do not fully exploit factor structure. A PPP test using the described methods is presented in Section 4 as an illustrative example. Section 5 is devoted to conclusions. In particular, the implications of the findings for modeling in practice will be discussed.

2 Testing for unit roots in panel data when cross-sectional dependencies result from unobserved common factors

This section describes three approaches to panel unit root testing in the presence of cross-sectional correlation which employ factor models. In particular, the methods proposed by Pesaran (2007), Moon and Perron (2004) and Bai and Ng (2004a) will be presented. For reasons of comparison, it also briefly describes the panel unit root tests by Breitung and Das (2008) and by Sul (2007) which assume a factor structure but do not fully exploit it.

The factor structure used by all approaches is a convenient form to model cross-correlation, or even cointegration between panel members. Therefore, the (for pooled testing necessary) assumption of independence between the individual specific components of the data is far less restrictive than the assumption of independent cross-sections, underlying the IPS and LLC test.

A note on notation: Throughout this paper, M is used to denote a finite, generic constant. For a matrix A, $A > 0$ denotes that A is positive definite. Common factors which are denoted by f_t are always assumed to be stationary. Common factors denoted by F_t result from an autoregressive transformation of f_t. F_t has a unit root when there is a unit root in the autoregression. Whenever we refer to nonstationary common factors, this means nonstationarity of F_t.

2.1 Pesaran (2007): A dynamic panel model with one common factor

For a panel of observed data with N cross-sectional units and T time series observations, Pesaran (2007) uses a simple dynamic linear heterogenous model

$$Y_{i,t} = (1-\delta_i)\mu_i + \delta_i Y_{i,t-1} + u_{i,t}, \qquad i = 1,\ldots,N, \qquad t = 1,\ldots,T, \tag{1}$$

with given initial values $Y_{i,0}$ and a one-factor structure for the disturbance

$$u_{i,t} = \lambda_i f_t + e_{i,t}. \tag{2}$$

Considering serially uncorrelated disturbances, the idiosyncratic components, $e_{i,t}$, $i = 1,\ldots,N$, $t = 1,\ldots,T$ are assumed to be independently distributed both across i and t, have zero mean, variance σ_i^2, and finite forth-order moment. The common factor f_t is serially uncorrelated with mean zero and constant variance σ_f^2, and finite forth-order moment. Without loss of generality, σ_f^2 is set equal to one. $e_{i,t}$, λ_i and f_t are assumed to be mutually independent for all i and t.

It is convenient to write (1) and (2) as

$$\Delta Y_{i,t} = \alpha_i - (1-\delta_i)Y_{i,t-1} + \lambda_i f_t + e_{i,t}, \tag{3}$$

where $\alpha_i = (1-\delta_i)\mu_i$ and $\Delta Y_{i,t} = Y_{i,t} - Y_{i,t-1}$. The unit root hypothesis considered by Pesaran (2007), $\delta_i = 1$ for all i is tested against the possibly heterogenous alternative $\delta_i < 1$ for $i = 1,\ldots,N_1$, $\delta_i = 1$ for $i = N_1+1,\ldots,N$. Pesaran (2007) assumes that $\frac{N_1}{N}$, the fraction of the individual processes that is stationary, is non-zero and tends to some fixed value κ such that $0 < \kappa \leq 1$ as $N \to \infty$.

It is important to notice that any non-stationarity of the observations $Y_{i,t}$ in the setting considered by Pesaran (2007) is due to the presence of a unit root in the autoregressive part of (1), i.e. $\delta_i = 1$. For the unit root null hypothesis considered by Pesaran (2007), he proposes a test based on the t-ratio of the OLS estimate $\hat{b}_i$ in the following cross-sectionally augmented DF (CADF) regression

$$\Delta Y_{i,t} = a_i + b_i Y_{i,t-1} + c_i \bar{Y}_{t-1} + d_i \Delta\bar{Y}_t + \epsilon_{i,t}, \tag{4}$$

where $\bar{Y}_t = \frac{1}{N}\sum_{i=1}^{N} Y_{i,t}$, $\Delta\bar{Y}_t = \frac{1}{N}\sum_{i=1}^{N} \Delta Y_{i,t}$, and $\epsilon_{i,t}$ is the regression error.

The cross-sectional averages, $\bar{Y}_{t-1}$ and $\Delta\bar{Y}_t$, are included into (4) as a proxy for the unobserved common factor f_t. For analytical convenience when deriving the asymptotic properties, Pesaran (2007) replaces the usual estimator for σ_i^2 in the t-value for b_i by a slightly modified and also consistent one. He derives the asymptotic distribution of the modified t-statistic and shows that it is free of nuisance parameters as $N \to \infty$ for any fixed $T > 3$, as well as for the case where $N \to \infty$ followed by $T \to \infty$.

In line with Im, Pesaran and Shin (2003), Pesaran (2007) proposes a cross-sectional augmented version of the IPS-test

$$CIPS = \frac{1}{N}\sum_{i=1}^{N} CADF_i, \tag{5}$$

where $CADF_i$ is the cross-sectionally augmented Dickey-Fuller statistic for the i-th cross-sectional unit given by the t-ratio of b_i in the CADF regression (4). Due to the presence of the common factor, the $CADF_i$ statistics will not be cross-sectionally independent[2]. Thus, a central limit theorem cannot be applied to derive the limiting distribution of the $CIPS$ statistic, and it is shown to be non-standard even for large N. Furthermore, to ensure the existence of moments for the distribution of $CADF_i$ in finite samples, Pesaran (2007) advocates the use of a truncated version of the $CIPS$ test, where for positive constants K_1 and K_2 such that $Pr[-K_1 < CADF_i < K_2]$ is sufficiently large, values of $CADF_i$ smaller than $-K_1$ or larger than K_2 are replaced by the respective bounds. Pesaran (2007) provides values for K_1 and K_2 obtained by simulations.

The presentation above outlines the procedure for serially uncorrelated disturbances. If there is serial correlation present in the common factors or idiosyncratic errors, additional lags of $\Delta Y_{i,t}$ and its cross-sectional average $\Delta \bar{Y}_t$ have to be included in the ADF regression (4).

2.2 Moon and Perron (2004): A dynamic panel model with K common factors

For a panel of observed data with N cross-sectional units and T time series observations, Moon and Perron (2004) model the DGP for $Y_{i,t}$ as an AR(1) process and assume, similar to Pesaran (2007), that common factors are present in the error term. They assume a K-factor model for the error term $u_{i,t}$

$$Y_{i,t} = (1-\delta_i)\mu_i + \delta_i Y_{i,t-1} + u_{i,t}, \tag{6}$$

$$u_{i,t} = \lambda_i' f_t + e_{i,t}, \tag{7}$$

for $i = 1, \ldots, N$ and $t = 1, \ldots, T$, where f_t is a $(K \times 1)$ vector of common factors, λ_i is the corresponding vector of factor loadings for cross-section i, and $e_{i,t}$ is an idiosyncratic disturbance term.

[2]Under the null hypothesis of a unit root, $CADF_i$ converges to a functional of Brownian motions, say $G(W_f, W_i)$, where W_f and W_i are Brownian motions driven by the common factor and idiosyncratic error, respectively.

The DGPs considered by Pesaran (2007) and Moon and Perron (2004) are identical if a single common factor is present in the composite error term. For the components of the composite error term in (7) similar assumptions are made as by Pesaran (2007). The idiosyncratic part $e_{i,t}$ follows a stationary and invertible infinite MA process, and is cross-sectionally uncorrelated, so that $e_{i,t} = \Gamma_i(L)\varepsilon_{i,t}$, where $\Gamma_i(L) = \sum_{j=0}^{\infty} \gamma_{i,j} L^j$ and $\varepsilon_{i,t} \sim i.i.d.(0,1)$ across i and t. Also the common factors f_t are assumed to have a stationary, invertible $MA(\infty)$ representation, i.e. $f_t = \Phi(L)\eta_t$. Here, $\Phi(L) = \sum_{j=0}^{\infty} \phi_j L^j$ is a K-dimensional lag polynomial and $\eta_t \sim i.i.d.(0, I_K)$. Furthermore, the covariance matrix of f_t is (asymptotically) positive definite. Although more than one common factor are permitted to be present in the data, some maximum number $\bar{K}(\geq K)$ is supposed to be known. Also, redundant factors, i.e. factors that asymptotically influence only a finite number of observed series, are excluded by imposing $\frac{1}{N}\sum_{i=1}^{N} \lambda_i \lambda_i' \rightarrow^p \Sigma_\lambda > 0$. Furthermore, short-run and long-run variances, $\sigma_{e_i}^2 \big(= \sum_{j=0}^{\infty} \gamma_{i,j}^2 \big)$ and $\omega_{e_i}^2 \big(= (\sum_{j=0}^{\infty} \gamma_{i,j})^2 \big)$, as well as the one sided long-run covariance $\varphi_{e_i} \big(= \sum_{l=1}^{\infty} \sum_{j=0}^{\infty} \gamma_{i,j}\gamma_{i,j+l} \big)$ are supposed to exist for all idiosyncratic disturbances $e_{i,t}$. Additionally, these parameters are assumed to have non-zero cross-sectional averages, $\sigma_e^2 = \frac{1}{N}\sum_{i=1}^{N} \sigma_{e_i}^2$, $\omega_e^2 = \frac{1}{N}\sum_{i=1}^{N} \omega_{e_i}^2$ and $\varphi_e^2 = \frac{1}{N}\sum_{i=1}^{N} \varphi_{e_i}^2$.

The unit root null hypothesis considered by Moon and Perron (2004) is $H_0 : \delta_i = 1$ for all $i = 1, \ldots, N$, which is tested against the heterogenous alternative $H_1 : \delta_i < 1$ for some i[3]. To test this hypothesis, two modified t-statistics are suggested, based on pooled estimation of the first-order serial correlation coefficient of the data. The estimation and testing procedure relies on de-factoring the data by a projection onto the space orthogonal to that spanned by the common factors. For that purpose, the matrix of factor loading $\Lambda = (\lambda_1, \ldots, \lambda_N)'$ has to be estimated to construct a projection matrix $Q_\Lambda = I_N - \Lambda(\Lambda'\Lambda)^{-1}\Lambda'$.

Imposing $\delta_i = \delta$ for all i, the pooled OLS estimator, denoted as $\hat{\delta}_{pooled}$, is T-consistent for 1 under the unit root null, as well as under the local alternative considered by Moon and Perron (2004). The usual t-ratio to test this hypothesis has a non-standard limiting distribution, due to the persistent cross-sectional correlation introduced by the common factors. From the residuals of the pooled regression (under the null where the intercept is equal to zero)

$$\hat{u}_{i,t} = Y_{i,t} - \hat{\delta}_{pooled} Y_{i,t-1}, \tag{8}$$

the matrix of factor loadings is estimated by the method of principal components[4]. With the estimator $\hat{\Lambda}$ one can then construct an estimator of the projection matrix denoted as $Q_{\hat{\Lambda}_K}$. Additionally, consistent estimates of the above defined nuisance parameters can be obtained

[3]To analyze local power properties of their test, Moon and Perron (2004) consider the following local alternative hypothesis:

$$\delta_i = 1 - \frac{\theta_i}{\sqrt{N}T},$$

where θ_i is a random variable with mean μ_θ on finite support $[0, \bar{M}]$. The considered null hypothesis is $H_0' : \mu_\theta = 0$, which is tested against the local alternative $H_1' : \mu_\theta > 0$.

[4]The principal component estimator is in general not unique. Moon and Perron (2004) use the normalization $\frac{1}{T}\sum_{t=1}^{T} f_t f_t' = I_K$ and re-scale the obtained estimate.

non-parametrically from the de-factored residuals $\hat{e} = \hat{u}Q_{\hat{\Lambda}_K}$, where $\hat{u} = (\hat{u}_1, \ldots, \hat{u}_N)$ with $\hat{u}_i = (\hat{u}_{i,1}, \ldots, \hat{u}_{i,T})'$. Denote the estimates as $\hat{\varphi}_{e_i}$ and $\hat{\omega}^2_{e_i}$, and their cross-sectional averages as $\hat{\varphi}_e$ and $\hat{\omega}^2_e$. Then the modified pooled estimator of δ suggested by Moon and Perron (2004) is

$$\delta^*_{pooled} = \frac{\sum_{t=2}^{T} Y'_{t-1} Q_{\hat{\Lambda}_K} Y_t - NT\hat{\varphi}_e}{\sum_{t=2}^{T} Y'_{t-1} Q_{\hat{\Lambda}_K} Y_{t-1}}, \tag{9}$$

where $Y_t = (Y_{1,t}, \ldots, Y_{N,t})'$. Based on this estimator, the following two t-statistics can be used to test the pooled unit root null hypothesis,

$$t^*_a = \frac{\sqrt{N}T(\hat{\delta}^*_{pooled} - 1)}{\sqrt{\frac{2\hat{\phi}^4_e}{\hat{\omega}^4_e}}} \tag{10}$$

and

$$t^*_b = \sqrt{N}T(\hat{\delta}^*_{pooled} - 1)\sqrt{\frac{1}{NT^2}\sum_{t=2}^{T} Y'_{t-1} Q_{\hat{\Lambda}_K} Y_{t-1}}\left(\frac{\hat{\omega}_e}{\hat{\phi}^2_e}\right), \tag{11}$$

where $\hat{\phi}^4_e = \frac{1}{N}\sum_{i=1}^{N} \hat{\phi}^4_{e_i}$, $\hat{\phi}^4_{e_i} = \hat{\omega}^4_{e_i}$. Moon and Perron (2004) analyze the asymptotic behavior of the two statistics as $N \to \infty$ and $T \to \infty$ with[5] $\liminf_{(N,T\to\infty)} \frac{\log T}{\log N} > 1$. Both test statistics have a limiting standard normal distribution under the null, and diverge under the stationary alternative.

2.3 Bai and Ng (2004a): A common factor model with unobserved common and idiosyncratic components of unknown order of integration.

In contrast to Pesaran (2007) or Moon and Perron (2004), the PANIC model of Bai and Ng (2004a) permits the non-stationarity in a panel of observed data to come either from a common source, or from the idiosyncratic errors, or from both[6]. Therefore, they focus on consistent estimation of the common factors and error terms, to test the properties of these series separately.

The model Bai and Ng (2004a) consider describes the observed data $Y_{i,t}$ as the sum of a deterministic part, a common (stochastic) component, and the idiosyncratic error. In particular,

$$Y_{i,t} = D_{i,t} + \lambda'_i F_t + E_{i,t} \qquad i = 1, \ldots, N,\ t = 1, \ldots, T, \tag{12}$$

where as before λ_i is a $(K \times 1)$ vector of factor loadings, F_t is a $(K \times 1)$ vector of common factors[7], and $E_{i,t}$ is an error term. The deterministic component $D_{i,t}$ contains either a constant α_i or a linear trend $\alpha_i + \beta_i t$. As the two aforementioned approaches, Bai and Ng (2004a) consider a balanced panel with N cross-sectional units and, T time series observations.

[5] The restriction on the relative divergence rate of N and T is necessary, as f_t and $e_{i,t}$ are unobserved.

[6] Under the unit root null the data in the Pesaran's (2007) or Moon and Perron's (2004) model contains a common, as well as an idiosyncratic stochastic trend.

[7] K is assumed to be known here.

The common factors are assumed to follow an $AR(1)$ process, such that

$$F_t = F_{t-1} + f_t, \tag{13}$$

where $f_t = \Phi(L)\eta_t$, $\Phi(L) = \sum_{j=1}^{\infty} \phi_j L^j$ is a $K-$dimensional lag polynomial and $\text{rank}\big(\Phi(1)\big) = k_1$. So, F_t contains $k_1 \leq K$ independent stochastic trends and consequently $K - k_1$ stationary components. The shock η_t is assumed to be $i.i.d.(0, \Sigma_\eta)$ with finite fourth-order moment. The idiosyncratic terms are allowed to be either I(0) and I(1), and are also modelled as $AR(1)$ processes

$$E_{i,t} = \delta_i E_{i,t-1} + e_{i,t}, \tag{14}$$

where $e_{i,t}$ follows a mean zero, stationary, invertible MA process, such that $e_{i,t} = \Gamma_i(L)\varepsilon_{i,t}$ with $\varepsilon_{i,t} \sim i.i.d.(0, \sigma^2_{\varepsilon_i})$. Bai and Ng (2004a) do not assume cross-sectional independence of the idiosyncratic term[8] from the outset, but impose it later to validate pooled testing. The assumption that Σ_η is not (necessarily) a diagonal matrix is more general than the corresponding assumption in Moon and Perron (2004), where the innovations of the common factors are assumed to be uncorrelated. The short-run covariance matrix of ΔF_t has full rank while the long-run covariance matrix has reduced rank and hence permits cointegration among the common factors. As in Moon and Perron (2004), (asymptotically) redundant factors are ruled out.

In this setup, the goal of PANIC is to determine the number of non-stationary factors k_1, and to test for each $i = 1, \ldots, N$, whether $\delta_i = 1$. Bai and Ng (2004) suggest using principal components to consistently estimate the unobserved components F_t and $E_{i,t}$. However, to derive consistent estimates even if some elements of F_t and $E_{i,t}$ are I(1), a suitable transformation of $Y_{i,t}$ is used. In particular, if the DGP does not contain a deterministic linear trend, the first differences of the data are employed, while in the presence of a deterministic linear trend, demeaned first-differences are used. So, in the former case $y_{i,t} = \Delta Y_{i,t} = Y_{i,t} - Y_{i,t-1}$, while in the latter $y_{i,t} = \Delta Y_{i,t} - \Delta \bar{Y}_{i,t}$, where $\Delta \bar{Y}_{i,t} = \frac{1}{T-1} \sum_{t=2}^{T} \Delta Y_{i,t}$. As the estimated common factors and idiosyncratic errors, denoted as $\hat{f}_t$ and $\hat{e}_{i,t}$ respectively, are derived applying the method of principal components to first-differenced or de-trended data, Bai and Ng (2004a) propose to re-accumulate them to remove the effect of possible overdifferencing. This yields

$$\hat{F}_t = \sum_{s=2}^{t} \hat{f}_s, \tag{15}$$

$$\hat{E}_{i,t} = \sum_{s=2}^{t} \hat{e}_{i,s}. \tag{16}$$

These estimates are now individually tested for unit roots.

For the idiosyncratic components, Bai and Ng (2004a) suggest to compute an ADF statistic

[8] Bai and Ng(2004a) allow for some weak cross-sectional dependence of the shock terms driving the $e_{i,t}$. The full set of assumptions can be found in their paper.

based on up to p lags. Denote the t-statistic to test the unit root hypothesis for each $\hat{E}_{i,t}$ as $ADF^c_{\hat{E}_i}$ or $ADF^\tau_{\hat{E}_i}$, depending on whether a constant, or a constant and linear trend is included in the DGP. Bai and Ng (2004a) derive the limiting distributions, which are non-standard. For the case where a constant is present in the DGP given by (12), the distribution coincides with the usual Dickey-Fuller (DF) distribution where no constant is included in the estimation. The 5% critical value is -1.95. If the DGP in (12) contains a constant and a linear trend, the limiting distribution is proportional to the reciprocal of a Brownian bridge. Critical values for this distribution are not tabulated yet, and have to be simulated.

Both ADF statistics given above do not have the advantage of a standard normal limiting distribution, as do the other panel unit root tests described so far. That is due to the fact that the panel information has only been used to consistently estimate $E_{i,t}$, but not to analyze its dynamic properties. Only if independence among the error terms is assumed, pooled testing is valid. In that case, Bai and Ng (2004a) propose a Fisher-type test[9] as suggested in Maddala and Wu (1999), using the correction proposed by Choi (2001). The test statistic, denoted as $P^c_{\hat{E}}$ or $P^\tau_{\hat{E}}$ depending on the deterministic specification, is given by

$$P^c_{\hat{E}}, P^\tau_{\hat{E}} = \frac{-2\sum_{i=1}^{N} \log \pi_i - 2N}{\sqrt{4N}}, \tag{17}$$

where π_i is the p-value of the ADF test for the i-th cross-section. These two panel unit root test statistics have standard normal limiting distributions.

Depending on whether there is just one, or several common factors, Bai and Ng (2004a) suggest to use either an ADF test based on up to p lags, or a rank test for $\hat{F}_t$. Denote the t-statistic for the unit root hypothesis as $ADF^c_{\hat{F}}$ when only a constant is accounted for, and as $ADF^\tau_{\hat{F}}$ in the linear trend case. Then, Bai and Ng (2004a) derive their limiting distributions, which coincide with the DF distributions for the cases where only a constant, or a constant and a linear trend are included in the ADF estimation. The asymptotic 5% critical values are -2.86 and -3.41, respectively.

If there are $K > 1$ common factors, Bai and Ng (2004a) suggest an iterative procedure, comparable to the Johansen trace test for cointegration to select k_1. They use demeaned or de-trended factor estimates, depending on whether (12) contains just a constant, or a constant and linear trend. Define $\tilde{F}_t = \hat{F}_t - \bar{\hat{F}}_t$ with $\bar{\hat{F}}_t = \frac{1}{T-2}\sum_{t=2}^{T} \hat{F}_t$ in the former case. In the latter, let $\tilde{F}_t$ denote the residuals from a regression of $\hat{F}_t$ on a constant and linear trend. Using $\tilde{F}_t$, the following steps describe the proposed test.
Starting with $m = K$,

1. Let $\hat{\beta}_\perp$ be the m eigenvectors associated with the m largest eigenvalues of $\frac{1}{T^2}\sum_{t=2}^{T} \tilde{F}_t \tilde{F}_t'$. Let $\hat{X}_t = \hat{\beta}_\perp' \tilde{F}_t$. Two statistics can be considered:

2. (a) Let $K(j) = 1 - \frac{j}{J+1}$, $j = 1, \ldots, J$;

[9] In principal, also an IPS-type test using a standardized average of the above described t-statistics should be possible. See also Bai and Ng (2007).

i. Let $\hat{\xi}_t$ be the residuals from estimating a VAR(1) in $\hat{X}_t$, and let

$$\hat{\Sigma}_1 = \sum_{j=1}^{J} K(j)\big(\frac{1}{T}\sum_{t=2}^{T} \hat{\xi}_{t-j}\hat{\xi}_t'\big).$$

ii. Let $\hat{\nu}_c(m)$ be the smallest eigenvalue of

$$\widehat{\Phi}_c(m) = \frac{1}{2}\Big[\sum_{t=2}^{T}(\hat{X}_t\hat{X}_{t-1}' + \hat{X}_{t-1}\hat{X}_t') - T(\hat{\Sigma}_1 + \hat{\Sigma}_1')\Big]\Big(\sum_{t=2}^{T}\hat{X}_{t-1}\hat{X}_{t-1}'\Big)^{-1}.$$

iii. Denote $T[\hat{\nu}_c(m) - 1]$ as $MQ_c^c(m)$ in the constant only case, or as $MQ_c^\tau(m)$ in the linear trend case.

(b) For p fixed that does not depend on N or T,

i. Estimate a Var(p) in $\Delta\hat{X}_t$ in order to obtain $\widehat{\Pi}(L) = I_m - \widehat{\Pi}_1 L - \cdots - \widehat{\Pi}_p L^P$. Filter $\hat{X}_t$ by $\widehat{\Pi}(L)$ to get $\hat{x}_t = \widehat{\Pi}(L)\hat{X}_t$.

ii. Let $\hat{\nu}_f(m)$ be the smallest eigenvalue of

$$\widehat{\Phi}_f(m) = \frac{1}{2}\Big[\sum_{t=2}^{T}(\hat{x}_t\hat{x}_{t-1}' + \hat{x}_{t-1}\hat{x}_t')\Big]\Big(\sum_{t=2}^{T}\hat{x}_{t-1}\hat{x}_{t-1}'\Big)^{-1}.$$

iii. Denote $T[\hat{\nu}_f(m) - 1]$ as $MQ_f^c(m)$ in the constant only case, or as $MQ_f^\tau(m)$ in the linear trend case.

3. If $H_0 : k_1 = m$ is rejected, set $m = m - 1$ and return to Step 1. Otherwise, set $\hat{k}_1 = m$ and stop.

For the $MQ_c^{c,\tau}$ and $MQ_f^{c,\tau}$ statistics described above, Bai and Ng (2004a) derive limiting distributions, which are again non-standard, and they provide 1%, 5%, and 10% critical values for all four statistics and for various values of m.

The PANIC procedure has the advantage that the estimated common factors and idiosyncratic components are consistent whether they are stationarity or non-stationarity. This is due to the practice of estimating the unobserved components from the first-differenced (or de-trended) data, and re-accumulating the estimates to remove the effect of possible overdifferencing if the factors or errors are stationary. Hence, the obtained estimates could also be used for stationarity tests, which is discussed in Bai and Ng (2004b).

2.4 Alternative panel unit root tests in the presence of cross-sectional dependencies

The three approaches to panel unit root testing presented in the previous sections explicitly account for the common factors employed to model the cross-sectional dependence in the data by using methods that require large N to be valid. In this section we introduce alternative panel unit root tests which do not necessarily exploit the common factor structure, and could provide alternatives to the aforementioned tests in small N panels. In particular, we will consider two test statistics proposed by Breitung and Das (2008) and the tests proposed by Sul (2007).

Breitung and Das (2008)

Breitung and Das (2008) study the behaviour of several panel unit root tests when cross-sectional dependence in the data is present in the form of a common factor. The DGP they employ is similar to that of Bai and Ng (2004a) presented in Section 2.3, Equations (12) to (14). However, Breitung and Das (2008) focus on the special case where (13) can is replaced by

$$F_t = \rho F_{t-1} + f_t,$$

with the scalar first order autoregressive parameter $|\rho| \leq 1$. They consider test statistics on the "reduced form" regression equation below, which is obtained when $\delta_i = \delta$ for all i and $\rho = \delta$:

$$\Delta Y_t = \phi Y_{t-1} + u_t, \tag{18}$$

where $\Delta Y_t = (\Delta Y_{1,t}, \ldots, \Delta Y_{N,t})'$, $Y_{t-1} = (Y_{1,t-1}, \ldots, Y_{N,t-1})'$, $u_t = (u_{1,t}, \ldots, u_{N,t})'$ with $u_{i,t} = \lambda_i f_t + e_{i,t}$ and $\phi = (\delta - 1)$. Breitung and Das (2008) present their analysis for a DGP and model without individual specific constant or time trend. The deterministic component in (12) has been assumed to be zero in this case. If a model with individual specific constant is employed, Breitung and Das (2008) suggest to remove it by considering data in deviation from the first observation, $Y^*_{i,t} = Y_{i,t} - Y_{i,0}$.

Breitung and Das (2008) particularly consider a robust OLS t-statistic t_{rob} and a GLS t-statistic t_{gls} to test for the unit root null hypothesis $\phi = 0$ against the homogenous alternative $\phi < 0$. The robust OLS statistic is given by

$$t_{rob} = \frac{\sum_{t=1}^{T} Y'_{t-1} \Delta Y_t}{\left(\sum_{t=1}^{T} Y'_{t-1} \hat{\Omega} Y_{t-1}\right)^{\frac{1}{2}}},$$

with $\hat{\Omega} = \sum_{t=1}^{T} \hat{u}_t \hat{u}'_t$ where $\hat{u}_t = \Delta Y_t - \hat{\phi} Y_{t-1}$ are the OLS residuals. The GLS statistic, t_{gls}, is given by

$$t_{gls} = \frac{\sum_{t=1}^{T} Y'_{t-1} \hat{\Omega}^{-1} \Delta Y_t}{\left(\sum_{t=1}^{T} Y'_{t-1} \hat{\Omega}^{-1} Y_{t-1}\right)^{\frac{1}{2}}}.$$

Note that this statistic can only be computed for $T > N$, as otherwise $\hat{\Omega}$ is singular. Also, if a common factor structure is assumed for the data, one could exploit this in for the GLS statistic by taking the factor structure into account when estimating the covariance matrix Ω. For the static factor model with orthonormal factors, $\Omega = \Lambda\Lambda' + \Sigma$, where Λ is the $N \times k$ matrix of factor loadings and Σ is the covariance matrix of the idiosyncratic innovations. Estimates of Λ and Σ can be obtained using a principal component approach as in Bai and Ng (2004a) or Moon and Perron (2004). If there is higher order serial correlation present in the residuals, a Newey-West type estimator for Ω can be employed, or an ADF regression estimated in the first step.

Breitung and Das (2008) consider 3 cases in their analysis, where the reduced form (18) is misspecified in cases 2 and 3, namely an I(1) common factor combined with I(1) idiosyncratic components, an I(1) common factor and I(0) idiosyncratic components (cross-member cointegration) and the case where a unit root is present in the idiosyncratic component but the common factor is I(0). If $\frac{N^3}{T} \to 0$, t_{gls} is asymptotically normally distributed in the first and third case, while it diverges in the second case. t_{rob} converges to a Dickey-Fuller distribution in the first case if there is a single common factor. It is equivalent to an ADF test on the first principal component of Y_t in that case. In the other cases, the test is not valid.

Sul (2007)

Sul (2007) proposes to use recursive mean adjustment for panel unit root tests to increase their power. Similar to Moon and Perron (2004), Sul (2007) models cross-sectional dependence by employing a common factor structure for the error term. The DGP is similar to that given in Equations (6) and (7). To account for the cross-sectional dependence, Sul (2007) suggests a (feasible) GLS statistic to test for the unit root null hypothesis $\delta_i = 1$ for all i against the heterogenous alternative $\delta_i < 1$ for some i in

$$Y_{i,t} = (1 - \delta_i)\mu_i + \delta_i Y_{i,t-1} + u_{i,t}, \tag{19}$$

The test procedure follows multiple steps, where the regression can be augmented by lagged first differences of $Y_{i,t}$ to account for higher order serial correlation in the residuals:

1. Run the following regression for each unit individually

$$Y_{i,t} - c_{i,t-1} = \delta_i(Y_{i,t-1} - c_{i,t-1}) + \sum_{j=1}^{p_i} \varphi_{ij} \Delta Y_{i,t-j} + \epsilon_{i,t}, \tag{20}$$

   where $c_{i,t-1} = (t-1)^{-1} \sum_{s=1}^{t-1} Y_{i,s}$ is the recursive mean, to obtain the LS estimator $\hat{\delta}_i$.

2. If $\hat{\delta}_i > 1$ set $\hat{\delta}_i = 1$ and run the regression

$$Y_{i,t} - \hat{\delta}_i Y_{i,t-1} = a_i + \sum_{j=1}^{p} \varphi_{ij} \Delta Y_{i,t-j} + \varepsilon_{i,t}. \tag{21}$$

   Construct the sample covariance matrix $\hat{\Omega} = (T - p - 1)^{-1} \sum_{t=p+1}^{T} \hat{\varepsilon}_t \hat{\varepsilon}_t'$, where $\hat{\varepsilon}_t = (\hat{\varepsilon}_{1,t}, \ldots, \hat{\varepsilon}_{N,t})'$ are the vectors of residuals from the previous regression.

3. Project $(Y_{i,t} - c_{i,t-1})$ and $(Y_{i,t-1} - c_{i,t-1})$ on the lagged first differences

$$\begin{aligned}(Y_{i,t} - c_{i,t-1}) &= \sum_{j=1}^{p} \phi_{ij} \Delta Y_{i,t-j} + \xi_{i,t}, \\ (Y_{i,t-1} - c_{i,t-1}) &= \sum_{j=1}^{p} \zeta_{ij} \Delta Y_{i,t-j} + \xi_{i,t-1}.\end{aligned}$$

4. Define $\hat{\omega}'_{ij}$ as the ij^{th} element of $\hat{\Omega}^{-1}$, one can now obtain the pooled FGLS estimator of δ and the associated t-statistic as

$$\hat{\delta}_{fglsrma} = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} \hat{\omega}'_{ij} \sum_{t=p+1}^{T} \hat{\xi}_{i,t-1} \hat{\xi}_{j,t}}{\sum_{i=1}^{N} \sum_{j=1}^{N} \hat{\omega}'_{ij} \sum_{t=p+1}^{T} \hat{\xi}^2_{i,t-1}}, \quad (22)$$

$$t_{fglsrma} = \frac{\hat{\delta}_{fglsrma} - 1}{\sqrt{\sum_{i=1}^{N} \sum_{j=1}^{N} \hat{\omega}'_{ij} \sum_{t=p+1}^{T} \hat{\xi}^2_{i,t-1}}}. \quad (23)$$

Sul (2007) shows that the $t_{fglsrma}$ converges to a Dickey-Fuller distribution, and he provides finite sample critical values to account for finite sample bias.

Similar to Breitung and Das (2008), Sul's (2007) $t_{fglsrma}$ effectively tests for a unit root in the idiosyncratic component of the data if the error term $u_{i,t}$ in (19) permits a common factor structure. To test for a unit root in the common component, Sul (2007) proposes to apply a recursive mean adjusted covariate augmented DF test to the cross-sectional averages of the data, $\bar{Y}_t = N^{-1} \sum_{i=1}^{N} Y_{i,t}$. The steps of the procedure are similar to the ones outlined above, and the resulting t-statistic is denoted as t_{crma}. Sul (2007) provides some evidence that his test is precise and powerful, especially when T is larger than N, a case for which it has been designed.

2.5 Differences and similarities

This section discusses differences and similarities of the panel unit root tests relying on a factor structure, presented in the previous subsections. For all considered DGPs, we can write the data as the sum of the deterministic component ($D_{i,t}$), a "common component" ($CC_{i,t}$) and an "idiosyncratic component" ($IC_{i,t}$) such that

$$Y_{i,t} = D_{i,t} + CC_{i,t} + IC_{i,t}.$$

For the DGP of Bai and Ng (2004a), we have $CC_{i,t} = \lambda'_i F_t$ and $IC_{i,t} = E_{i,t}$. For a DGP as assumed by Pesaran (2007) or Moon and Perron (2004) given in e.g. (6)-(7) where the common factor structure is assumed for the error term, we obtain for a simple AR(1),

$$Y_{i,t} = (1 - \delta_i^t)\mu_i + \lambda'_i \sum_{s=0}^{t-1} \delta_i^{t-s} f_{t-s} + \sum_{s=0}^{t-1} \delta_i^{t-s} e_{i,t-s}.$$

Hence, $CC_{i,t} = \lambda_i' \sum_{s=0}^{t-1} \delta_i^{t-s} f_{t-s}$ and $IC_{i,t} = \sum_{s=0}^{t-1} \delta_i^{t-s} e_{i,t-s}$ for those DGPs. The approaches to panel unit root testing presented above may differ in terms of assumptions made which place restriction on the DGP, in particular whether the order of integration is allowed to differ between $CC_{i,t}$ and $IC_{i,t}$ and thus whether the possibility of cross-member cointegration is excluded or not, and the number of common factors. Furthermore, the presented test statistics are applied to different data components. For example, the Moon and Perron (2004) tests apply to the idiosyncratic component only, as has been shown by Breitung and Das (2008) and forcefully argued by Bai and Ng (2007).

DGP

The DGP assumed by Pesaran (2007) for a single common factor and Moon and Perron (2004) for $K \geq 1$ restrict the common and idiosyncratic component to have the same order of integration. Bai and Ng (2004a) explicitly allow the order of integration to differ between $CC_{it,}$ and $IC_{i,t}$ and they allow for the presence of $K \geq 1$ factors. Sul (2007) considers a DGP similar to Bai and Ng (2004a) as well and proposes to proxy a single common factor with the cross-sectional average of the data. Breitung and Das (2008) analyze the behaviour of their tests in DGPs as assumed by Bai and Ng (2004a).

Null Hypothesis and Tested Data Component

All considered tests have non-stationarity as null hypothesis. The statistics proposed by Pesaran (2007) and Moon and Perron test defactored data ($IC_{i,t}$) for a unit root. The common component is not tested, although it is non-stationary if $IC_{i,t}$ is non-stationary given that the assumptions on the DGP are true. Bai and Ng (2004a) suggest test statistics for the idiosyncratic and common component separately, where the null hypothesis is non-stationarity of the given component. Breitung and Das (2008) formulate the null hypothesis in terms of the reduced form regression (18) as a unit root in the observed data. However, they show that their FGLS statistic effectively tests for a unit root in the idiosyncratic component, while their robust OLS statistic is equivalent to an ADF test for the first principal component only if both $CC_{i,t}$ and $IC_{i,t}$ are non-stationary. Sul's (2007) FGLS statistic also tests for a unit root in the idiosyncratic component, while cross-sectional averages are used as a proxy for a single common factor and tested for a unit root with the t_{crma} test.

The $CIPS$ test of Pesaran (2007), the tests of Moon and Perron (2004), the $P_{\hat{E}}^{c,\tau}$ statistics of Bai and Ng (2004a) and the FGLS statistics proposed by Breitung and Das (2008) and Sul (2007) are pooled tests for the null hypothesis that the defactored data are unit root processes for all i. All approaches except Breitung and Das (2008) use a heterogenous alternative, namely that some series have a unit root and some do not. Moon and Perron (2004) use a pooled estimator of the first order autoregressive coefficient δ_i in the construction of their statistics. Similarly, the FGLS tests of Breitung and Das (2008) and Sul (2007) are based on pooled estimators $\hat{\delta}$. The individual specific $CADF$ statistic of Pesaran (2007) and the

$ADF_{\hat{E}}^{c,\tau}$ statistic of Bai and Ng (2004a) test for a unit root in the idiosyncratic component for a given i, and the alternative hypothesis is stationarity of that component.

Bai and Ng's (2004a) $ADF_{\hat{F}}^{c,\tau}$ statistic and Sul's (2007) t_{crma} test for a unit root in a single common factor. Also, Bai and Ng (2004a) allow for more than one common factor and the $MQ_c^{c,\tau}$ and $MQ_f^{c,\tau}$ statistics are designed to determine the number of independent stochastic trends in F_t.

Panel dimensions N and T

The three type of tests proposed by Pesaran (2007), Moon and Perron (2004) and Bai and Ng (2004a) are designed for large N and T due to the estimation of the common factor(s) either by using principal components or by including the cross-sectional mean as proposed by Pesaran (2007). The FGLS tests of Breitung and Das (2008) and Sul (2007) on the other hand can only be constructed if $T > N$.

Cointegration

While Pesaran (2007) and Moon and Perron (2004) exclude the possibility of cointegration among the $Y_{i,t}$, as well as between the observed data and the common factors, Bai and Ng (2004a) include both possibilities in their model. In particular, if $k_1 \geq 1$ and $E_{i,t}(= IC_{i,t})$ is stationary for some i, then the observed data and the common factors are cointegrated for those i with cointegrating vector $(1, -\lambda_i')'$. Furthermore, if all idiosyncratic errors are I(0), then the orthogonalization matrix used by Moon and Perron (2004) to eliminate the common factors, Q_Λ, serves as cointegration matrix for the $Y_{i,t}$. So, Bai and Ng's (2004a) procedure can be used as a cointegration test[10], by investigating the hypotheses $k_1 \geq 1$ and all idiosyncratic errors are stationary[11]. Breitung and Das (2008) consider the case of cross-member cointegration in their analysis, however their tests are not able to detect it. Sul's (2007) tests could be used to detect cross-member cointegration, namely if the $t_{fglsrma}$ statistic rejects a unit root for the idiosyncratic component while the t_{crma} test fails to reject the unit root for the cross-sectional averages.

Common Factors and Estimation of K

For the tests proposed by Moon and Perron (2004) and Bai and Ng (2004a), an important aspect in application is the selection of the number of common factors K. Consistent estimation of K is discussed in Bai and Ng (2002) for a factor model as given by (12) with stationary errors, and also briefly treated in Moon and Perron (2004). It should be noted that while the information criteria designed to estimate the number of common factors work

[10]What is meant here is a cointegration test between panel members, in contrast to panel cointegration tests. The latter ones are used to test for cointegration between several variables for the same i.

[11]Note that the null hypothesis for the ADF tests using the estimated error terms remains that of non-stationarity. Rejecting the unit root hypothesis for all i is thus one part of not rejecting cointegration between panel members.

well in simulations, their application in practice is difficult as they are usually observed to select the maximum number of common factors allowed.

In terms of computational burden, all procedures are rather easy to implement. Pesaran (2007) provides tables with critical values for his tests. The PANIC procedure of Bai and Ng (2004a) also requires some tabulated critical values for the rank test statistics $MQ^c_{(\cdot)}$ and $MQ^f_{(\cdot)}$, as well as for the $ADF^\tau_{\hat{E}_{i,t}}$ statistic. Also, a procedure to calculate the p-values of $ADF^c_{\hat{E}_{i,t}}$ and $ADF^\tau_{\hat{E}_{i,t}}$ is needed to implement the suggested pooled tests. Sul (2007) also provides simulated finite sample critical values for his test statistics.

3 Small sample performance: Monte Carlo results

3.1 Monte Carlo simulation setup

In this section we study the small sample performance of the tests proposed by Pesaran (2007), Moon and Perron (2004) and Bai and Ng (2004a) for various types of DGPs. Furthermore, we consider the robust OLS t-test t_{rob} and the FGLS t-test t_{GLS} described in Breitung and Das (2008) and the recursive mean adjusted FGLS test $t_{\rho fglsrma}$ and the recursive mean adjusted test for the average data proposed by Sul (2007). All considered DGPs with one exception have the following structure which corresponds to Bai and Ng's (2004a) framework:

$$\begin{aligned} Y_{i,t} &= \lambda_i' F_t + E_{i,t}, \\ F_{m,t} &= \varphi F_{m,t-1} + f_{m,t}, \\ E_{i,t} &= \delta_i E_{i,t-1} + e_{i,t}, \end{aligned} \tag{24}$$

with $i = 1, \ldots, N$, $t = 1, \ldots, T$ and $m = 1, \ldots, K$. We consider three different values for N and T each, namely 20, 50 and 100[12]. The method of principle components estimates the space spanned by the common factors when N is large. We have chosen N and T at least equal to 20 to assure that common factors are estimated with sufficient precision or approximated reasonably well by cross-sectional averages. Notice that the regularity condition $N \neq T$ needed for some tests is not satisfied in some cases. First a single common factor is considered, which is generated by a first order autoregression, or a random walk when $\varphi = 1$. We also consider the case of two common factors which are generated using the same parameter values for φ and σ^2_f, but different drawings for the error terms. The idiosyncratic terms $E_{i,t}$ are also generated by a first order autoregression or random walk with first order moving average, depending on whether or not $\delta_i = 1$.

[12] Pesaran (2007) reports Monte Carlo results for his tests with N, T = 10, 20, 30, 50, 100, Moon and Perron (2004) choose N = 10, 20 and T = 100, 300, Bai and Ng (2004a) report results for N = 40, T = 100 while Bai and Ng (2007) choose N, T = 20, 50, 100, Sul (2007) performs simulations with N = 5, 10, 15, 20 and T = 50, 100, 200, and Breitung and Das (2008) select N = 10, 20, 50 and T = 20, 50, 100.

In addition, a DGP as assumed by Pesaran (2007) and Moon and Perron (2004) is used:

$$\begin{aligned} Y_{i,t} &= \delta_i Y_{i,t-1} + u_{i,t}, \\ u_{i,t} &= \lambda_i f_t + e_{i,t}. \end{aligned} \quad (25)$$

In (24) and (25) the error terms are generated as MA(1) processes such that

$$\begin{aligned} f_{m,t} &= \eta_{m,t} + \gamma_m \eta_{m,t-1}, \\ e_{i,t} &= \varepsilon_{i,t} + \rho_i \varepsilon_{i,t-1}. \end{aligned}$$

The shocks are drawn from independent normal distributions, such that $\eta_t \sim i.i.d.N(0, \Sigma_f^2)$, with $\Sigma_f^2 = \sigma_f^2 I_K$, and $\varepsilon_{i,t} \sim i.i.d.N(0,1)$. We consider three different values for the signal-to-noise ratio, such that $\sigma_f^2 = 0.5$, 1 and 2[13]. The MA parameters γ_m and ρ_i are independently, uniformly distributed on $[0.2, 0.5]$. The factor loading λ_i are uniformly distributed on $[-1, 3]$[14].

Three different types of non-stationarity are considered as null hypothesis, as well as different settings for the stationary alternative hypothesis. In particular, we consider the following 5 cases, where 1 to 4 use the DGP given by (24) and 5 uses DGP (25)[15]:

1. Common and idiosyncratic unit roots

$$H_0^A : \varphi = 1, \text{ and } \delta_i = 1 \text{ for all } i.$$

2. Common unit root, nearly stationary idiosyncratic components

$$H_0^B : \varphi = 1, \text{ and } \delta_i \sim U[0.8, 1] \text{ for all } i,$$

3. Stationary common component, integrated idiosyncratic components

$$H_0^C : \varphi = 0.95, \text{ and } \delta_i = 1 \text{ for all } i,$$

4. Stationary common and idiosyncratic components

$$H_A^A : \varphi = 0.95 \text{ and } \delta_i \sim U[0.8, 1].$$

5. Stationary data using a DGP as given by (25) with heterogenous roots

$$H_A^E : \delta_i \sim U[0.8, 1] \text{ for all } i.$$

[13] In the tables we only report the values for $\sigma_f^2 = 1$. The other results are available at http://www.personeel.unimaas.nl/J.Urbain/.

[14] Consistency of the test procedure of Pesaran (2007) requires a non-zero mean for the factor loadings. This assumption is not necessary for the other approaches.

[15] Please note that under setup 1 (24) and (25) are equivalent. In cases 4 and 5 we have stationarity provided $\delta_i \neq 1$.

The results are obtained with GAUSS 8.0 using 1000 replications. The reported rejection frequencies are based on 5% nominal size. All power results are size unadjusted. For Pesaran's (2007) $CADF$ and $CIPS$ we use the critical values reported in Tables 1b and 3b of his paper. Results for Moon and Perron's (2004) statistics, Bai and Ng's (2004a) $P^c_{\hat{E}}$ statistic and Breitung and Das (2008) t_{rob} and t_{gls} are based on a critical value from the standard normal distribution. Rejection frequencies of the $ADF^c_{\hat{E}}$ and $ADF^c_{\hat{F}}$ statistics are obtained using the critical values from DF distributions for the no intercept and intercept only cases, respectively. Critical values for the MQ^c_c and MQ^c_f are provided in Table 1 of Bai and Ng (2004a). For Sul's (2007) $t_{\rho fglsrma}$ test we use finite sample critical values reported in Table 5 of Sul (2007) and for the t_{crma} we use the asymptotic critical value of -1.88. When obtaining the t_{crma} statistic we use $Y_{1,t}$ as covariate and calculate the cross-sectional averages over the remaining $N-1$ panel members such that $\bar{Y}_t = (N-1)^{-1}\sum_{i=2}^{N} Y_{i,t}$.

Similar to Moon and Perron (2004), we use the Andrews-Monahan (1992) estimator employing the quadratic spectral kernel in the estimation of the nuisance parameters for the t^*_a and t^*_b statistics. For Bai and Ng's (2004a) $ADF^c_{\hat{E}}$ and $ADF^c_{\hat{F}}$ and Pesaran's (2007) $CADF$ and $CIPS$ we use the Akaike information criterion (AIC) to determine the lag length, starting with a maximum lag length of $p_{max} = 6$. For the test of Sul (2007) and Breitung and Das (2008) we use the Bayesian information criterion (BIC). For the MQ^c_c statistic we use the Bartlett kernel with a bandwidth as suggested in Andrews (1991). The lag length for the MQ^c_f statistic is determined using the criteria proposed by Aznar and Salvador (2002).

Although the considered DGPs do not include deterministic components, we do account for individual fixed effects in the simulation by including constants in the regressions. Following the advise of Breitung and Das (2008) for the t_{rob} and t_{gls} test we consider data in deviation from the initial observation to remove the effect of an individual specific constant[16].

The finite sample performance of the considered test statistics depend on these choices. For reasons of comparison, we follow the original authors with the choices they report or we select a procedure that performs better in terms of size in our simulations.

3.2 Monte Carlo results

A general finding is that the presence of serial correlation[17] leads to size distortions for almost all statistics when T is small, which can be quite strong in some cases and even persist for $T = 100$. For a single common factor, the signal-to-noise ratio seems to have little to no effect on the tests proposed by Pesaran (2007) and Bai and Ng (2004a). For two common factors in the DGP, Bai and Ng's (2004a) MQ^c_c and MQ^c_f statistics usually select maximum possible number of common stochastic trends, leading to low size and low power for these tests when the auto-regressive root is close to unity. The FGLS statistics of Breitung and Das (2008)

[16] As already noted by Breitung and Das (2008), applying the tests to demeaned data leads to dependence on nuisance parameter unless applied to the GLS transformed data, and severe finite sample size distortions.

[17] Results for the case of i.i.d. $N(0,1)$ error terms $e_{i,t}$ and f_t in (24) are not included in this version of the paper. They are are available at http://www.personeel.unimaas.nl/J.Urbain/.

and Sul (2007) behave quite similarly in terms of size and power. Sul's (2007) t_{crma} statistic applied to the cross-sectional averages of the data has similar size properties as Bai and Ng's (2004) $ADF^c_{\hat{F}}$. Power properties of the two tests are similar too for most cases.

The results in Table 1 are obtained for the case where a unit root is present in the common factors and in all idiosyncratic errors. Both statistics proposed by Pesaran (2007), the $CADF$[18] and the $CIPS$ test show size distortions when T is small (20), which are stronger for the $CIPS$ test. For a single common factor those size distortions are reduced as T increases and for $T = 100$ the tests are only slightly over-sized. For $K = 2$, size distortions increasing in the signal-to-noise ratio remain even for large T, in particular for the $CIPS$ test. Both statistics proposed by Moon and Perron (2004) show slight size distortions which seem to increase with the signal-to-noise ratio when K is correctly specified. The size distortions are decreasing in T and higher for t^*_a than for t^*_b. The later is actually undersized for small signal-to-noise ratios. If K is misspecified, both statistics show strong size distortions increasing in the signal-to-noise ratio, but size distortions are lower when K is over-estimated. Bai and Ng's $ADF^c_{\hat{E}}$ and $ADF^c_{\hat{F}}$ statistics for the extracted individual idiosyncratic error series and the single common factor respectively, are oversized for small $T(= 20)$ but size distortions decrease as T gets large. The pooled statistic $P^c_{\hat{E}}$ has strong size distortions when T is small and size increases in N. For $T = 100$, size rages from 0.12 to 0.18 for the different values of N. Similar to Moon and Perron's (2004) tests, size distortions are less severe when the number of common factors is over-specified if $\hat{K} \neq K$. Both rank statistics MQ^c_c and MQ^c_f usually pick the maximum number of possible common stochastic trends, leading to good properties when K is specified correctly but failure to estimate the correct number of common factors if $\hat{K} = 3$ is used. Breitung and Das's (2008) t_{rob} is under-sized for small T with rejection frequencies increasing in T but decreasing with N, leading to rejection frequencies between 0.00 and 0.11. The t_{gls} test has a size of about 0.05 for $N = 20$ and is under-sized for $N = 50$, similarly to Sul's $t_{\rho fglsrma}$ tests. The t_{crma} test for the cross-sectional averages is slightly oversized with size distortions decreasing in T. All four statistics behave similarly whether a single or two common factors are present in the data.

Insert Tables 1 about here

Table 2 considers the case of a unit root in the common factors and near-unit roots in the idiosyncratic factors, i.e. the case of cross-member cointegration. For $K = 1$, Pesaran's (2007) $CADF$ statistic has an average rejection frequency of about 0.32 for $T = 20$ and between 0.17 and 0.21 for larger T. The rejection frequencies of the $CIPS$ test are high and go to 1 for large N and T. For $K = 2$, rejection frequencies are reduced, in particular for $CIPS$ where they also decrease as the signal-to-noise ratio increases. Both statistics proposed by Moon and Perron (2004) have rejection increasing to 1 in N and T, with rejection frequencies for t^*_a slightly higher than those for t^*_b when the correct number of common factors is employed.

[18]Entries for the CADF-statistics are average rejection frequencies of the individual unit root tests.

When K is under-estimated, rejection frequencies are strongly reduced. Bai and Ng's $ADF^c_{\hat{E}}$ statistic has an average power increasing from about 0.23 to 0.48 as both N and T increase. The pooled $P^c_{\hat{E}}$ test has a power of 1 for almost all combinations of N and T considered, when K is correctly specified or over-specified. When a single common factor is extracted but two common factors are present in the data, rejection frequencies are reduced. The $ADF^c_{\hat{F}}$ tests has some size distortions, but rejection frequencies decrease from about 0.40 for $T = 20$ to 0.07 to 0.10 for $T = 100$. The MQ^c_c and MQ^c_f statistics again pick the maximum number of possible trends, leading to good properties only when K is correctly specified. Rejection frequencies for Breitung and Das's (2008) t_{rob} statistics decrease for higher signal-to-noise ratios, whereas they increase with T. The t_{gls} statistic has rejection frequencies between 0.44 and 0.64, increasing with T. Sul's (2007) $t_{\rho fglrma}$ statistic has similar rejection frequencies ranging between 0.38 and 0.66, which also increase in T. The t_{crma} test is slightly oversized with size distortions lower for $T = 100$.

Insert Table 2 about here

Table 3 covers the case of integrated idiosyncratic errors combined with a stationary common factor. The statistics proposed by Pesaran (2007) behave similar to the case of I(1) idiosyncratic and common component (Table 1), but size is slightly reduced for the $CIPS$ test, which is now under-sized for $T = 100$ and $K = 1$. Moon and Perron's (2004) t^*_a and t^*_b also behave similar to Table 1 but have slightly higher rejection frequencies, increasing in the signal-to-noise ratio in particular for $K = 2$. When K is misspecified, rejection frequencies for both statistics increase in N, T and the signal-to-noise ratio. Bai and Ng's (2004a) $ADF^c_{\hat{E}}$ and $P^c_{\hat{E}}$ tests have sizes close to the one shown in Table 1. The power of the $ADF^c_{\hat{F}}$ is smaller than 0.20 for $T \geq 50$. The MQ^c_c and MQ^c_f statistics fail to detect the correct number of common stochastic trends. Breitung and Das's (2008) t_{rob} test has size increasing in T but decreasing in N. The t_{gls} and Sul's (2007) $t_{fglsrma}$ tests are slightly over-sized for $N = 20$ and under-sized for $N = 50$, with size increasing in the signal-to-noise ratio. The t_{crma} test has rejection frequencies ranging from 0.15 to 0.24, increasing in N and T but decreasing as the signal-to-noise ratio increases. Also, rejection frequencies for the tests of Breitung and Das (2008) and Sul (2006) are slightly larger when $K = 2$.

Insert Tables 3 about here

Tables 4 and 5 consider stationary data. For Table 4 the DGP is given by (24) with I(0) idiosyncratic and common components. Pesaran's (2007) $CADF$ has low power while the power of the $CIPS$ test is relatively high and increasing in N, reaching 1 for $N, T = 100$ when $K = 1$. For $K = 2$, power is reduced and furthermore decreasing in the signal-to-noise ratio. Moon and Perron's (2004) tests both have power increasing to 1 as N and T increase. The average power of Bai and Ng's (2004a) $ADF^c_{\hat{E}}$ is relatively low (0.52 for $N, T = 100$) while the pooled test $P^c_{\hat{E}}$ has a power of 1 for $N > 20$ or $T > 20$. The power of the $ADF^c_{\hat{F}}$ is low and both rank tests MQ^c_c and MQ^c_f fail to select the correct number of common stochastic trends. Breitung and Das's (2008) t_{rob} test has power increasing in T but decreasing the signal-to-noise ratio. The t_{gls} test has a power between 0.55 and 0.84, increasing in T but decreasing in N, similar to Sul's (2007) $t_{fglsrma}$ test which has power between 0.51 and 0.87. Power for these 3 tests is increased for $K = 2$. Rejection frequencies for the t_{crma} are 0.10 and 0.27 when $K = 1$ and 0.07 and 0.30 for $K = 2$, increasing in T but decreasing in N and the signal-to-noise ratio.

Table 5 considers stationary data generated using (25). Rejection frequencies for most tests are reduced and now decrease as the signal-to-noise ration increases, in particular for Moon and Perron's (2004) t_a^* and t_b^* and Breitung and Das's (2008) t_{gls} and Sul's (2007) $t_{fglsrma}$ tests. Bai and Ng's (2004a) $ADF_{\hat{F}}^c$ has a higher power now, but it is still relatively low. Sul's (2007)t_{crma} test also has an increased power now, increasing in N, T and the signal-to-noise ratio.

Insert Tables 4-5 about here

From the Monte Carlo simulations, several general conclusions can be drawn. The presence of serial correlation in the error term leads to size distortions which can be quite large in small samples. The Moon and Perron (2004) tests, the tests of Pesaran (2007), the $P_{\hat{E}}^c$ and $ADF_{\hat{E}}^c$ statistics of Bai and Ng (2004a) and the FGLS statistics proposed by Breitung and Das (2008) and Sul (2007) indeed test for a unit root in the idiosyncratic component, and reject a unit root if it is present in the common factor alone. The pooled CIPS test of Pesaran (2007) and $P_{\hat{E}}^c$ test of Bai and Ng (2004a) are more powerful than the individual test statistics they are based on, CADF and $ADF_{\hat{E}}^c$ respectively. However, the pooled tests show higher size distortions for small T. The CIPS test has good size and power for large N and T if a single common factor is present. However, an additional common factor leads to size distortions and reduced power. The $P_{\hat{E}}^c$ statistic has high power, but some size distortions remain even for $N, T = 100$. The t_a^* statistic has slightly larger size distortions that the t_b^* test, with power being high for both statistics. The later three statistics are not distorted by the presence of a second common factor if K is correctly specified in the estimation. If K is misspecified, the statistics exhibit size distortions, but over-estimating K seems to be less harmful in terms of power. The two FGLS statistics are slightly undersized for $N = 50$ but have a high power. Also, their performance remains good in terms of size and power if two factors are included in the data.

Bai and Ng's (2004a) $ADF_{\hat{F}}^c$ statistic and Sul's (2007) t_{crma} statistic have been proposed test whether there is a unit root in a single common factor. The $ADF_{\hat{F}}^c$ has low power and some size distortions even for large N, T. Sul's (2007) t_{crma} test shows similar size and power in most cases, but has a higher power when the DGP given in (25) is used. Bai and Ng's (2004a) MQ_c^c and MQ_f^c are designed to test for the number of common stochastic trends if more than one common factor is present, but have very low power against alternatives close to a unit root.

We have not studied the issue of which test to choose if the common factor model representation is not appropriate to describe cross-sectional dependence. Bootstrap unit root tests might be used in such an instance, but this question is left for future research. Furthermore, we do not consider DGPs with idiosyncratic linear deterministic trends. Moon and Perron (2004) show that their tests have no local power in that case, but all other authors consider propose their tests for such DGPs as well. Bai and Ng (2007) provide simulation results for some tests for DGPs including idiosyncratic linear deterministic trends.

4 An illustrative application: Testing for PPP using the new approaches

This section presents an application of the new panel unit root tests described in Section 2 to illustrate their use in an empirical study of the validity of purchasing power parity (PPP). For this purpose we consider the potential existence of a unit root in real exchange rate series that are constructed as

$$Y_{i,t} = s_{i,t} - p_t^* + p_{i,t}, \tag{26}$$

where $s_{i,t}$ is the ln of country i's nominal exchange rate versus some numeraire currency, p_t^* is the ln of the aggregate price level in the numeraire country, and $p_{i,t}$ is the ln of country i's domestic aggregate price level.

The numerous analyzes of PPP in the literature do not come to a common conclusion with respect to PPP. Some studies report stronger rejection of the unit root null, if the German Mark instead of the US Dollar is used as a numeraire currency. Also, studies using univariate unit root or cointegration tests reject PPP, while tests using panel methods as the LLC or IPS test tend to find evidence in favor of it, see for example Oh (1996). However, as was already discussed in the introduction, several studies have analyzed the properties of early panel unit root tests in the presence of cross-sectional dependence since then, and argued against their use for PPP tests. Lyhagen (2000) analytically derives the cross-correlation structure in a panel of real exchange rates, constructed with a common numeraire country. He also derives the effect of the common stochastic trend in the data introduced by the numeraire on the limiting distributions of various panel statistics. In Monte Carlo simulations, he finds size distortions similar to those reported by Banerjee et al. ((2004) and (2005)).

In the analysis presented in this section, monthly data from 14 European countries is considered. The data set includes information on the nominal exchange rates of local currency versus US Dollar ($US) for Austria, Denmark, Finland, France, Germany, Greece, Italy, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and the UK. Furthermore, the Consumer Price Index (CPI) as a proxy for aggregate price levels is included for those 14 countries and the US. The sample includes monthly observations on all variables for the period from February 1986 to September 2002, so 200 observations. For such a sample, one can expect to find high correlation between panel units, due to a high degree of economic integration and political co-operation. As far as monetary policy is concerned, the most important mechanism of co-operation is the European Monetary System (EMS), to which some panel members belong, and which finally led to the introduction of the Euro as a common currency in some countries.

As a starting point of the analysis, the real exchange rate series are individually tested for a unit root using an ADF test. The lag length p is set to 12 for all countries. The individual ADF statistics are shown in Table 6. Only for the UK, the unit root null can be rejected for both real exchange rate series. Using the real exchange rate versus DM, also for Switzerland the ADF test rejects at a 5% significance level. These findings are representative for those of studies using univariate tests. The problem here is that it remains unclear whether the non-rejection of the unit root is due to a failure of PPP, or the low power of the ADF test against near unit root alternative.

Next, the panel unit root tests described in Section 2 are performed. For each test, it is assumed that a single common factor is present in the data. Given that the real exchange rate series are constructed using a common base currency, this assumption seems reasonable. For the tests of Pesaran (2007) and Moon and Perron (2004), the data representation in (1)-(2) is assumed to be valid. The results of the $CADF_i$ tests suggested by Pesaran (2007) are given in Table 6, and Table 7 presents Pesaran's (2007) $CIPS$ statistic and those proposed by Moon and Perron (2004). Except for the French real exchange rate when measured against the German Mark, the $CADF_i$ statistics fail to reject the unit root null. Also, the pooled $CIPS$ test does not reject the null in both panels. This provides some evidence against PPP. The t_a^* and t_b^* statistics of Moon and Perron (2004) do not provide such a clear picture. While the former one rejects PPP in both panels, the latter one does not reject it when the US Dollar is used as a numeraire currency.

The results for the panel unit root tests proposed by Breitung and Das (2008) and by Sul (2007) are given in Table 7. While the unit root null hypothesis is not rejected by any test for real exchange rates constructed with the US as base country, the t_{gls} test of Breitung and Das (2008) and the $t_{fglsrma}$ test of Sul (2007) reject the unit root when real exchange rates are constructed with Germany as base country.

Insert Tables 6-7 about here

For the application of the Bai and Ng (2004a) procedure, it is assumed that the data can be represented as in (12). With this representation, there is an interpretive problem. Clearly, if both F_t and $E_{i,t}$ are stationary, the real exchange rate is stationary and PPP holds in the long run at least. Also, if both common and idiosyncratic components are I(1), PPP can be rejected. But, if just the common factors are non-stationary the real exchange rate series are pairwise cointegrated along the cross-section but individually non-stationary, so that PPP in the usual sense does not hold between panel members and the base country. However, in the special case $\lambda_i = \lambda_j$, the cointegrating vector for $Y_{i,t}^B$ and $Y_{j,t}^B$ is $[1, -1]$, where the superscript

B denotes the base country. Then PPP holds between countries i and j, since

$$Y_{i,t}^{B} - Y_{j,t}^{B} = s_{i,t}^{B} - s_{j,t}^{B} + p_{i,t} - p_{j,t} = s_{i,t}^{j} + p_{i,t} - p_{j,t} = Y_{i,t}^{j} \sim I(0). \tag{27}$$

The results for the test statistics suggested by Bai and Ng (2004a) are presented in Table 7. Most of the individual tests for the idiosyncratic errors, as well as the test for the common factor reject the unit root. Also, the pooled error test rejects the unit root for both panels of real exchange rates. This provides some evidence in favor of PPP.

5 Conclusion

In this paper several panel unit root tests that account for cross section dependence assuming or using a common factor structure have been proposed in the literature, notably Pesaran (2007), Moon and Perron (2004), Bai and Ng (2004a), Breitung and Das (2008) and Sul (2007). There are often valid theoretical and empirical reasons why a common factor structure can be expected to yield sensible results. Therefore, panels with dynamic factors are of interest in economic modelling.

We have studied these approaches to unit root testing in panels with dynamic factors, compared them in terms of DGP, tests, null and alternative hypotheses. We have studied the small sample behavior of the tests proposed in a common framework and discussed their use in econometric modelling. In addition, we have applied them in an empirical study of purchasing power parity.

The main conclusions are:

- In the case where the observed non-stationarity is only due to a non-stationary common factor, the individual series are pairwise cointegrated along the cross sectional dimension. Only the Bai and Ng (2004a) and Sul (2007) tests allow for this type of structure to be detected, if the unit root is rejected for the idiosyncratic component but not for the common factor.

- The $ADF_{\hat{F}}^{c}$ for testing for the presence of unit roots in a single common factor is found to have low power. Similarly, in a multi-factor setting, the MQ_{c}^{c} and MQ_{f}^{c} tests fail to distinguish high but stationary serial correlation from non-stationarity in the common factors. For the one factor model, Bai and Ng's (2004) $ADF_{\hat{F}}^{c}$ test has similar size and power than Sul's (2007) t_{crma} test in most cases, except when a DGP as given in (25) is employed in which case the later test is more powerful.

- Testing the idiosyncratic component for a unit root: Pesaran's (2007) CADF and CIPS tests are indeed designed for testing for unit roots when cross-sectional dependence is due to a single common factor, and size and power are adversely affected by a second common factor. The pooled CIPS test has better power properties than the individual specific CADF tests. Similarly, Bai and Ng's (2004a) pooled $P_{\hat{E}}^{c}$ tests is more powerful than the individual specific $ADF_{\hat{E}}^{c}$ in detecting unit roots in the idiosyncratic components, although the former can have strong size distortion when the time dimension of the panel is small. However, the $P_{\hat{E}}^{c}$ and $ADF_{\hat{E}}^{c}$ statistic can accommodate to more than one common factor. The Moon and Perron (2004) tests can also account for multiple common factors. The two tests proposed by Moon and Perron (2004) are found to have similar small sample power, but the t_a^* statistic is found to have slightly larger size distortions than the t_b^*. When the FGLS tests considered by Breitung and Das (2008) and Sul (2007) can be computed, i.e. when $N < T$, they provide good alternatives to test for unit roots in the idiosyncratic component.

- When the number of common factors is unknown and has to be selected, it is less harmful in terms of power to include too many factors than too few in the test procedures of Bai and Ng (2004a) and Moon and Perron (2004). The statistics exhibits size distortions if the number of common factors is misspecified.

References

[1] Andrews, D.W.K. (1991), Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation, *Econometrica* 59, 817-858.

[2] Andrews, D.W.K. and Monahan, C. (1992), An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator, *Econometrica* 60, 953-966.

[3] Aznar, A. and Salvador, M. (2002), Selecting the Rank of the Cointegration Space and the Form of the Intercept Using an Information Criterion, *Econometric Theory* 18, 926-947.

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

[5] Bai, J. and Ng, S. (2004a), A PANIC Attack on Unit Roots and Cointegration, *Econometrica* 72, 1127-1177.

[6] Bai, J. and Ng, S. (2004b), A New Look at Panel Testing of Stationarity and the PPP Hypothesis, in *Indentification and Inference in Econometric Models: Essays in Honor of Thomas J. Rothenberg*, Don Andrews and James Stock (eds.), Cambridge University Press.

[7] Bai, J. and Ng, S. (2007), Panel Unit Root Tests with Cross-Section Dependence: A Further Investigation, Working Paper, Department of Economics, New York University.

[8] Banerjee, A., Marcellino, M. and Osbat, C. (2004), Some Cautions on the Use of Panel Methods for Integrated Series of Macro-Economic Data, *Econometrics Journal* 7, 322-340.

[9] Banerjee, A., Marcellino, M. and Osbat, C. (2005), Testing for PPP: Should We Use Panel Methods?, *Empirical Economics* 30, 77-91.

[10] Breitung, J. and Das, S. (2008), Testing for Unit Roots in Panels with a Factor Structure, Econometric Theory 24, 88-108.

[11] Chang, Y., 2002, Nonlinear IV Unit Root Tests in Panels with Cross-Sectional Dependency, *Journal of Econometrics* 110, 261-292.

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

[13] Flôres, R., Jorion, P. and Preumont, P.-Y., Szafarz, A. (1999) Multivariate unit root tests of the PPP hypothesis, *Journal of Empirical Finance* 6, 335-353.

[14] Im, K.S., Pesaran, M.H. and Shin, Y. (2003), Testing for unit roots in heterogenous panels, *Journal of Econometrics* 115, 53-74.

[15] Levin, A., Lin, C.F. and Chu C.S.J. (2002), Unit root tests in panel data: asymptotic and finite-sample properties, *Journal of Econometrics* 108, 1-24.

[16] Lyhagen, J. (2000), Why not use standard panel unit root test for testing PPP, mimeo, Stockholm School of Economics.

[17] Maddala, G.S. and Wu, S. (1999), A Comparative Study of Unit Root Tests with Panel Data and Cointegration Tests, *Oxford Bulletin of Economics and Statistics* 61, 631-652.

[18] Moon, H.R. and Perron, B. (2004), Testing for a Unit Root in Panels with Dynamic Factors, *Journal of Econometrics* 122, 81-126.

[19] O'Connell, P.G.J. (1998), The overvaluation of purchasing power parity, *Journal of International Economics* 44, 1-19.

[20] Oh, K.-Y. (1996), Purchasing Power Parity and unit root tests using panel data, *Journal of International Money and Finance* 15, 405-418.

[21] Pesaran, M.H. (2007), A Simple Panel Unit Root Test in the Presence of Cross Section Dependence, *Journal of Applied Econometrics*, 22, 265-312.

[22] Phillips, P. and Sul, D. (2003), Dynamic Panel Estimation and Homogeneity Testing Under Cross Section Dependence, *Econometrics Journal* 6, 217-259.

[23] Sul, D. (2007), Panel Unit Root Tests Under Cross Section Dependence with Recursive Mean Adjustment, mimeo, University of Auckland.

Table 1: Finite sample (average) rejection rates for DGP (24) with I(1) common factor(s) and I(1) idiosyncratic components.

K	**N**	**T**	t_a^*	t_b^*	$P_{\hat{E}}^c$	$ADF_{\hat{E}}^c$	$ADF_{\hat{F}}^c$	CADF	CIPS	t_{rob}	t_{gls}	$t_{fglsrma}$	t_{crma}
1	20	20	0.09	0.04	0.64	0.16	0.37	0.28	0.52	0.03	-	-	-
1	20	50	0.08	0.05	0.15	0.07	0.15	0.12	0.16	0.09	0.09	0.07	0.12
1	20	100	0.09	0.05	0.12	0.06	0.08	0.07	0.07	0.10	0.06	0.05	0.07
1	50	20	0.11	0.06	0.86	0.16	0.37	0.28	0.59	0.00	-	-	-
1	50	50	0.06	0.04	0.20	0.07	0.14	0.12	0.16	0.04	-	-	-
1	50	100	0.06	0.04	0.14	0.06	0.07	0.07	0.07	0.08	0.01	0.01	0.08
1	100	20	0.09	0.05	0.96	0.16	0.35	0.28	0.64	0.00	-	-	-
1	100	50	0.05	0.04	0.28	0.07	0.13	0.12	0.16	0.00	-	-	-
1	100	100	0.05	0.04	0.18	0.06	0.09	0.07	0.05	0.02	-	-	-

K	**N**	**T**	t_a^*	t_b^*	$P_{\hat{E}}^c$	MQ_c^c	MQ_f^c	CADF	CIPS	t_{rob}	t_{gls}	$t_{fglsrma}$	t_{crma}
2	20	20	0.11	0.05	0.63	1.00	0.97	0.32	0.53	0.04	-	-	-
2	20	50	0.08	0.05	0.16	1.00	1.00	0.15	0.30	0.09	0.09	0.06	0.12
2	20	100	0.10	0.05	0.12	1.00	1.00	0.09	0.24	0.09	0.06	0.07	0.10
2	50	20	0.14	0.10	0.82	1.00	0.98	0.31	0.59	0.01	-	-	-
2	50	50	0.05	0.03	0.20	1.00	1.00	0.15	0.34	0.03	-	-	-
2	50	100	0.06	0.04	0.14	1.00	1.00	0.08	0.25	0.09	0.01	0.01	0.08
2	100	20	0.12	0.09	0.96	1.00	0.99	0.31	0.60	0.00	-	-	-
2	100	50	0.06	0.04	0.31	1.00	1.00	0.14	0.33	0.01	-	-	-
2	100	100	0.04	0.03	0.15	1.00	1.00	0.08	0.28	0.01	-	-	-

			$\hat{\mathbf{K}}=\mathbf{1}$						$\hat{\mathbf{K}}=\mathbf{3}$				
K	**N**	**T**	t_a^*	t_b^*	$P_{\hat{E}}^c$	$ADF_{\hat{E}}^c$	$ADF_{\hat{F}}^c$		t_a^*	t_b^*	$P_{\hat{E}}^c$	MQ_c^c	MQ_f^c
2	20	20	0.31	0.22	0.58	0.16	0.40		0.25	0.22	0.61	0.00	0.01
2	20	50	0.32	0.23	0.25	0.07	0.12		0.18	0.14	0.16	0.00	0.00
2	20	100	0.32	0.23	0.21	0.06	0.08		0.15	0.11	0.13	0.00	0.00
2	50	20	0.38	0.32	0.75	0.16	0.39		0.37	0.36	0.82	0.00	0.01
2	50	50	0.38	0.33	0.32	0.07	0.13		0.25	0.26	0.19	0.00	0.00
2	50	100	0.41	0.37	0.26	0.06	0.09		0.22	0.21	0.12	0.00	0.00
2	100	20	0.45	0.42	0.78	0.16	0.38		0.37	0.36	0.92	0.00	0.01
2	100	50	0.48	0.46	0.42	0.07	0.14		0.30	0.29	0.30	0.00	0.00
2	100	100	0.50	0.49	0.38	0.06	0.10		0.22	0.21	0.18	0.00	0.00

Finite sample (average) rejection rates for Pesaran's (2007) CADF and CIPS statistics, Moon and Perron's (2004) t_a^* and t_b^* statistics, Bai and Ng's (2004a) $ADF_{\hat{E}}^c$, $P_{\hat{E}}^c$, and $ADF_{\hat{F}}^c$ statistics, Breitung and Das's (2008) t_{rob} and t_{gls} statistics, and Sul's (2007) $t_{fglsrma}$ and t_{crma} statistics. Proportions of repetitions when Bai and Ng's (2004a) MQ_c^c and MQ_f^c statistics chose the correct number of common stochastic trends. Finite sample (average) rejection rates for Moon and Perron's (2004) t_a^* and t_b^* statistics, and Bai and Ng's (2004a) $ADF_{\hat{E}}^c$, $P_{\hat{E}}^c$, and $ADF_{\hat{F}}^c$ statistics, the proportions of repetitions when Bai and Ng's (2004a) MQ_c^c and MQ_f^c statistics chose the correct number of common stochastic trends, when the number of common factors is misspecified. K denotes the number of common factors in the DGP. $\hat{K}$ specifies the number of common factors used when testing if $\hat{K}$ is different from K. Rejection frequencies are based on 5% cutoff values from Pesaran (2007), Tables 1b and 3b, Sul (2007) Table 5, Bai and Ng (2004a) Table 1, 5% cutoff values of the standard normal distribution, or 5% Dickey-Fuller critical values for the test statistics as specified in the text. Results are obtained with GAUSS 8.0 using 1000 replications.

Table 2: Finite sample (average) rejection rates for DGP (24) with I(1) common factor(s) and I(0) idiosyncratic components.

K	**N**	**T**	t_a^*	t_b^*	$P_{\hat{E}}^c$	$ADF_{\hat{E}}^c$	$ADF_{\hat{F}}^c$	CADF	CIPS	t_{rob}	t_{gls}	$t_{fglsrma}$	t_{crma}
1	20	20	0.57	0.40	0.94	0.23	0.40	0.32	0.68	0.06	-	-	-
1	20	50	0.85	0.76	1.00	0.25	0.14	0.17	0.55	0.19	0.45	0.39	0.10
1	20	100	0.92	0.85	1.00	0.44	0.07	0.18	0.85	0.16	0.62	0.65	0.09
1	50	20	0.75	0.67	1.00	0.23	0.38	0.31	0.79	0.02	-	-	-
1	50	50	0.98	0.96	1.00	0.26	0.13	0.18	0.71	0.11	-	-	-
1	50	100	1.00	0.99	1.00	0.46	0.09	0.19	0.98	0.17	0.56	0.59	0.07
1	100	20	0.99	0.85	1.00	0.24	0.38	0.32	0.86	0.02	-	-	-
1	100	50	1.00	1.00	1.00	0.27	0.14	0.19	0.81	0.07	-	-	-
1	100	100	1.00	1.00	1.00	0.48	0.09	0.21	1.00	0.11	-	-	-
K	**N**	**T**	t_a^*	t_b^*	$P_{\hat{E}}^c$	MQ_c^c	MQ_f^c	CADF	CIPS	t_{rob}	t_{gls}	$t_{fglsrma}$	t_{crma}
2	20	20	0.48	0.37	0.91	1.00	0.98	0.33	0.59	0.05	-	-	-
2	20	50	0.84	0.75	0.99	1.00	1.00	0.16	0.44	0.13	0.42	0.35	0.11
2	20	100	0.93	0.88	1.00	1.00	1.00	0.13	0.48	0.12	0.64	0.62	0.09
2	50	20	0.72	0.62	1.00	1.00	0.98	0.33	0.63	0.04	-	-	-
2	50	50	0.98	0.96	1.00	1.00	1.00	0.17	0.51	0.06	-	-	-
2	50	100	1.00	0.99	1.00	1.00	1.00	0.12	0.55	0.11	0.47	0.56	0.06
2	100	20	0.85	0.81	1.00	1.00	0.98	0.34	0.68	0.03	-	-	-
2	100	50	1.00	1.00	1.00	1.00	1.00	0.16	0.51	0.04	-	-	-
2	100	100	1.00	1.00	1.00	1.00	1.00	0.12	0.54	0.07	-	-	-
			$\hat{\mathbf{K}} = \mathbf{1}$						**$\hat{\mathbf{K}} = \mathbf{3}$**				
K	**N**	**T**	t_a^*	t_b^*	$P_{\hat{E}}^c$	$ADF_{\hat{E}}^c$	$ADF_{\hat{F}}^c$		t_a^*	t_b^*	$P_{\hat{E}}^c$	MQ_c^c	MQ_f^c
2	20	20	0.46	0.40	0.69	0.19	0.38		0.41	0.31	0.84	0.00	0.01
2	20	50	0.56	0.49	0.56	0.14	0.15		0.78	0.69	0.96	0.00	0.00
2	20	100	0.60	0.54	0.61	0.18	0.08		0.90	0.86	1.00	0.00	0.00
2	50	20	0.58	0.56	0.79	0.19	0.39		0.54	0.49	0.99	0.00	0.03
2	50	50	0.57	0.56	0.64	0.14	0.14		0.95	0.92	1.00	0.00	0.00
2	50	100	0.65	0.62	0.70	0.19	0.09		0.99	0.99	1.00	0.00	0.00
2	100	20	0.64	0.62	0.83	0.20	0.37		0.68	0.64	1.00	0.00	0.03
2	100	50	0.67	0.66	0.66	0.13	0.12		1.00	0.99	1.00	0.00	0.00
2	100	100	0.70	0.69	0.73	0.18	0.08		1.00	1.00	1.00	0.00	0.00

See notes Table 1.

Table 3: Finite sample (average) rejection rates for DGP (24) with I(0) common factor(s) and I(1) idiosyncratic components.

K	**N**	**T**	t_a^*	t_b^*	$P_{\hat{E}}^c$	$ADF_{\hat{E}}^c$	$ADF_{\hat{F}}^c$	CADF	CIPS	t_{rob}	t_{gls}	$t_{fglsrma}$	t_{crma}
1	20	20	0.13	0.06	0.61	0.16	0.41	0.27	0.47	0.04	-	-	-
1	20	50	0.11	0.06	0.12	0.07	0.19	0.11	0.10	0.22	0.11	0.10	0.18
1	20	100	0.12	0.06	0.12	0.06	0.19	0.06	0.02	0.43	0.09	0.09	0.20
1	50	20	0.17	0.10	0.84	0.16	0.40	0.27	0.55	0.00	-	-	-
1	50	50	0.09	0.06	0.20	0.07	0.17	0.11	0.08	0.08	-	-	-
1	50	100	0.08	0.06	0.12	0.06	0.19	0.05	0.01	0.31	0.01	0.01	0.23
1	100	20	0.13	0.08	0.95	0.16	0.40	0.27	0.58	0.00	-	-	-
1	100	50	0.08	0.06	0.25	0.07	0.16	0.11	0.06	0.01	-	-	-
1	100	100	0.07	0.06	0.15	0.06	0.19	0.06	0.01	0.09	-	-	-
K	**N**	**T**	t_a^*	t_b^*	$P_{\hat{E}}^c$	MQ_c^c	MQ_f^c	CADF	CIPS	t_{rob}	t_{gls}	$t_{fglsrma}$	t_{crma}
2	20	20	0.16	0.08	0.59	0.00	0.00	0.32	0.54	0.06	-	-	-
2	20	50	0.13	0.08	0.14	0.00	0.00	0.14	0.24	0.30	0.10	0.11	0.21
2	20	100	0.15	0.09	0.10	0.00	0.00	0.08	0.15	0.57	0.11	0.14	0.29
2	50	20	0.23	0.16	0.78	0.00	0.00	0.31	0.59	0.02	-	-	-
2	50	50	0.12	0.08	0.14	0.00	0.00	0.14	0.26	0.13	-	-	-
2	50	100	0.14	0.10	0.11	0.00	0.00	0.07	0.12	0.53	0.01	0.04	0.24
2	100	20	0.25	0.19	0.94	0.00	0.00	0.31	0.59	0.00	-	-	-
2	100	50	0.15	0.11	0.21	0.00	0.00	0.14	0.31	0.06	-	-	-
2	100	100	0.12	0.09	0.12	0.00	0.00	0.08	0.18	0.28	-	-	-

			$\hat{\mathbf{K}} = \mathbf{1}$						$\hat{\mathbf{K}} = \mathbf{3}$				
K	**N**	**T**	t_a^*	t_b^*	$P_{\hat{E}}^c$	$ADF_{\hat{E}}^c$	$ADF_{\hat{F}}^c$		t_a^*	t_b^*	$P_{\hat{E}}^c$	MQ_c^c	MQ_f^c
2	20	20	0.45	0.33	0.63	0.17	0.42		0.42	0.38	0.57	0.00	0.00
2	20	50	0.62	0.48	0.39	0.10	0.17		0.52	0.45	0.14	0.00	0.00
2	20	100	0.77	0.64	0.48	0.10	0.15		0.59	0.52	0.10	0.00	0.00
2	50	20	0.56	0.50	0.82	0.17	0.41		0.62	0.60	0.79	0.00	0.00
2	50	50	0.74	0.66	0.50	0.09	0.16		0.76	0.75	0.15	0.00	0.00
2	50	100	0.89	0.83	0.63	0.09	0.18		0.89	0.89	0.11	0.00	0.00
2	100	20	0.64	0.60	0.87	0.17	0.41		0.67	0.67	0.90	0.00	0.00
2	100	50	0.85	0.84	0.68	0.10	0.19		0.81	0.81	0.24	0.00	0.00
2	100	100	0.97	0.95	0.82	0.10	0.20		0.94	0.94	0.14	0.00	0.00

See notes Table 1.

Table 4: Finite sample (average) rejection rates for DGP (24) with I(0) common factor(s) and I(0) idiosyncratic components.

K	**N**	**T**	t_a^*	t_b^*	$P_{\hat{E}}^c$	$ADF_{\hat{E}}^c$	$ADF_{\hat{F}}^c$	CADF	CIPS	t_{rob}	t_{gls}	$t_{fglsrma}$	t_{crma}
1	20	20	0.66	0.49	0.94	0.23	0.43	0.31	0.64	0.08	-	-	-
1	20	50	0.91	0.83	1.00	0.26	0.17	0.16	0.48	0.36	0.56	0.51	0.18
1	20	100	0.95	0.90	1.00	0.48	0.17	0.16	0.74	0.63	0.81	0.83	0.24
1	50	20	0.86	0.78	1.00	0.24	0.41	0.30	0.73	0.03	-	-	-
1	50	50	0.99	0.98	1.00	0.27	0.17	0.17	0.59	0.25	-	-	-
1	50	100	1.00	1.00	1.00	0.49	0.20	0.17	0.95	0.62	0.69	0.75	0.13
1	100	20	0.95	0.94	1.00	0.24	0.41	0.31	0.82	0.04	-	-	-
1	100	50	1.00	1.00	1.00	0.29	0.18	0.17	0.69	0.20	-	-	-
1	100	100	1.00	1.00	1.00	0.52	0.15	0.19	1.00	0.51	-	-	-
K	**N**	**T**	t_a^*	t_b^*	$P_{\hat{E}}^c$	MQ_c^c	MQ_f^c	CADF	CIPS	t_{rob}	t_{gls}	$t_{fglsrma}$	t_{crma}
2	20	20	0.62	0.48	0.90	0.00	0.00	0.33	0.59	0.10	-	-	-
2	20	50	0.91	0.86	1.00	0.00	0.00	0.17	0.47	0.43	0.67	0.60	0.22
2	20	100	0.97	0.93	1.00	0.00	0.00	0.14	0.62	0.77	0.89	0.88	0.29
2	50	20	0.87	0.80	1.00	0.00	0.00	0.34	0.67	0.06	-	-	-
2	50	50	1.00	0.99	1.00	0.00	0.00	0.18	0.53	0.32	-	-	-
2	50	100	1.00	1.00	1.00	0.00	0.00	0.14	0.68	0.77	0.70	0.88	0.07
2	100	20	0.96	0.94	1.00	0.00	0.00	0.35	0.69	0.07	-	-	-
2	100	50	1.00	1.00	1.00	0.00	0.00	0.17	0.54	0.27	-	-	-
2	100	100	1.00	1.00	1.00	0.00	0.00	0.15	0.67	0.66	-	-	-
			$\hat{\mathbf{K}} = \mathbf{1}$						$\hat{\mathbf{K}} = \mathbf{3}$				
K	**N**	**T**	t_a^*	t_b^*	$P_{\hat{E}}^c$	$ADF_{\hat{E}}^c$	$ADF_{\hat{F}}^c$		t_a^*	t_b^*	$P_{\hat{E}}^c$	MQ_c^c	MQ_f^c
2	20	20	0.63	0.57	0.78	0.21	0.42		0.58	0.50	0.83	0.00	0.00
2	20	50	0.89	0.85	0.87	0.23	0.17		0.93	0.87	0.99	0.00	0.00
2	20	100	0.99	0.98	1.00	0.40	0.19		0.98	0.96	1.00	0.00	0.00
2	50	20	0.75	0.73	0.88	0.22	0.43		0.77	0.72	0.99	0.00	0.00
2	50	50	0.96	0.95	0.94	0.22	0.17		1.00	1.00	1.00	0.00	0.00
2	50	100	1.00	1.00	1.00	0.40	0.18		1.00	1.00	1.00	0.00	0.00
2	100	20	0.82	0.82	0.90	0.21	0.41		0.87	0.84	1.00	0.00	0.00
2	100	50	0.97	0.97	0.95	0.21	0.16		1.00	1.00	1.00	0.00	0.00
2	100	100	1.00	1.00	1.00	0.39	0.17		1.00	1.00	1.00	0.00	0.00

See notes Table 1.

Table 5: Finite sample (average) rejection rates for DGP (25) with I(0) common factor(s) I(0) idiosyncratic components.

K	**N**	**T**	t_a^*	t_b^*	$P_{\hat{E}}^c$	$ADF_{\hat{E}}^c$	$ADF_{\hat{F}}^c$	CADF	CIPS	t_{rob}	t_{gls}	$t_{fglsrma}$	t_{crma}
1	20	20	0.49	0.35	0.90	0.22	0.45	0.30	0.60	0.09	-	-	-
1	20	50	0.59	0.52	0.97	0.22	0.26	0.16	0.39	0.37	0.37	0.31	0.39
1	20	100	0.55	0.51	1.00	0.39	0.31	0.18	0.80	0.54	0.54	0.37	0.63
1	50	20	0.60	0.53	1.00	0.23	0.41	0.29	0.70	0.03	-	-	-
1	50	50	0.68	0.65	1.00	0.22	0.20	0.16	0.54	0.20	-	-	-
1	50	100	0.69	0.67	1.00	0.40	0.29	0.18	0.97	0.39	0.47	0.28	0.87
1	100	20	0.77	0.73	1.00	0.23	0.42	0.30	0.77	0.04	-	-	-
1	100	50	0.82	0.80	1.00	0.24	0.25	0.16	0.59	0.18	-	-	-
1	100	100	0.81	0.80	1.00	0.41	0.34	0.20	1.00	0.40	-	-	-
K	**N**	**T**	t_a^*	t_b^*	$P_{\hat{E}}^c$	MQ_c^c	MQ_f^c	CADF	CIPS	t_{rob}	t_{gls}	$t_{fglsrma}$	t_{crma}
2	20	20	0.41	0.29	0.83	0.00	0.00	0.33	0.57	0.11	-	-	-
2	20	50	0.45	0.38	0.87	0.00	0.00	0.17	0.44	0.37	0.28	0.21	0.27
2	20	100	0.48	0.41	1.00	0.00	0.00	0.18	0.73	0.55	0.37	0.25	0.40
2	50	20	0.61	0.53	0.98	0.00	0.00	0.33	0.68	0.07	-	-	-
2	50	50	0.62	0.56	1.00	0.00	0.00	0.18	0.56	0.29	-	-	-
2	50	100	0.61	0.57	1.00	0.00	0.00	0.19	0.84	0.57	0.25	0.18	0.95
2	100	20	0.69	0.65	1.00	0.00	0.00	0.33	0.69	0.07	-	-	-
2	100	50	0.70	0.67	1.00	0.00	0.00	0.17	0.57	0.24	-	-	-
2	100	100	0.67	0.65	1.00	0.00	0.00	0.18	0.81	0.45	-	-	-

			$\hat{\mathbf{K}} = \mathbf{1}$						$\hat{\mathbf{K}} = \mathbf{3}$				
K	**N**	**T**	t_a^*	t_b^*	$P_{\hat{E}}^c$	$ADF_{\hat{E}}^c$	$ADF_{\hat{F}}^c$		t_a^*	t_b^*	$P_{\hat{E}}^c$	MQ_c^c	MQ_f^c
2	20	20	0.51	0.43	0.83	0.22	0.44		0.53	0.46	0.75	0.00	0.00
2	20	50	0.57	0.50	0.94	0.25	0.27		0.66	0.60	0.84	0.00	0.00
2	20	100	0.62	0.55	1.00	0.45	0.40		0.67	0.62	0.99	0.00	0.00
2	50	20	0.69	0.65	0.93	0.47	0.46		0.70	0.67	0.96	0.00	0.00
2	50	50	0.82	0.78	0.99	0.27	0.27		0.88	0.87	1.00	0.00	0.00
2	50	100	0.85	0.82	1.00	0.37	0.36		0.87	0.86	1.00	0.00	0.00
2	100	20	0.75	0.74	0.96	0.22	0.43		0.80	0.78	0.99	0.00	0.00
2	100	50	0.82	0.81	0.99	0.24	0.22		0.94	0.93	1.00	0.00	0.00
2	100	100	0.86	0.84	1.00	0.47	0.31		0.94	0.93	1.00	0.00	0.00

See notes Table 1.

Table 6: Unit root test statistics for individual series of real exchange rates.

	ADF_i		$CADF_i$		$ADF^c_{\hat{E}}$	
Country	$Y^{\$}_{i,t}$	$Y^{DM}_{i,t}$	$Y^{\$}_{i,t}$	$Y^{DM}_{i,t}$	$Y^{\$}_{i,t}$	$Y^{DM}_{i,t}$
Austria	-1.4546	-1.9706	-1.4507	-1.7930	-2.0148^{**}	-1.0179
Denmark	-1.7667	-1.8571	-2.6133	-2.3326	-2.5865^{**}	-12.3785^{**}
Finland	-1.4707	-1.7878	-0.8203	-1.5242	-1.7758^{*}	-14.8259^{**}
France	-1.3803	-1.9697	-2.1514	-3.5277^{**}	-5.3264^{**}	-7.9788^{**}
Germany	-1.3537	-	-1.6298	-	-8.1950^{**}	-
Greece	-1.3602	-2.0247	-1.2353	-2.4531	-0.1947	-0.5564
Italy	-1.1971	-1.9952	-1.5815	-2.2758	-1.9926^{**}	-3.4528^{**}
NL	-1.8189	-1.1789	-1.8504	0.4805	-12.4864^{**}	-2.4592^{**}
Norway	-1.6895	-1.8398	-1.4439	-2.4683	-3.2075^{**}	-2.6420^{**}
Portugal	-1.9609	-1.9522	-0.6468	-1.5935	-1.3352	-2.2390^{**}
Spain	-1.0189	-1.9702	-0.6811	-1.2196	-1.6539^{*}	-1.5992
Sweden	-1.3288	-1.9717	-2.3255	-1.5175	-13.3224^{**}	-10.7084^{**}
Switzerland	-2.0188	-3.0893^{**}	-1.9543	-2.4054	-2.8777^{**}	-14.4867^{**}
UK	-3.2172^{**}	-2.8032^{*}	-1.7646	-2.7754	-3.7003^{**}	-12.4749^{**}
factor	-	-	-	-	-7.6314^{**}	-7.749^{**}

$*$ *indicates rejection at 10% significance level;*

$**$ *indicates rejection at 5% significance level.*

Table 7: Pooled unit root test statistics panels of real exchange rates.

	Pesaran (2007)	Moon and Perron (2004)		Bai and Ng (2004a)
	CIPS	t^*_a	t^*_b	$P^c_{\hat{E}}$
$q^{\$}_{i,t}$	-1.5821	-0.1214	-2.9890^{**}	16.6123^{**}
$q^{DM}_{i,t}$	-1.9543	-0.0358	-0.2076	18.4179^{**}

	Breitung and Das (2008)		Sul (2007)	
	t_{rob}	t_{gls}	$t_{fglsrma}$	t_{crma}
$q^{\$}_{i,t}$	-0.9687	0.1789	0.5708	-0.6639
$q^{DM}_{i,t}$	-1.0903	-3.3914^{**}	-4.0404^{**}	0.6835

$*$ *indicates rejection at 10% significance level;*

$**$ *indicates rejection at 5% significance level.*