ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Breitung, Jörg; Pesaran, Mohammad Hashem

Working Paper

Unit roots and cointegration in panels

CESifo Working Paper, No. 1565

Provided in Cooperation with:
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Breitung, Jörg; Pesaran, Mohammad Hashem (2005) : Unit roots and cointegration in panels, CESifo Working Paper, No. 1565, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:
https://hdl.handle.net/10419/19029

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

UNIT ROOTS AND COINTEGRATION IN PANELS

JOERG BREITUNG
M. HASHEM PESARAN

CESIFO WORKING PAPER NO. 1565
CATEGORY 10: EMPIRICAL AND THEORETICAL METHODS
OCTOBER 2005

An electronic version of the paper may be downloaded
- *from the SSRN website: www.SSRN.com*
- *from the CESifo website: www.CESifo-group.de*

CESifo *Working Paper No. 1565*

UNIT ROOTS AND COINTEGRATION IN PANELS

Abstract

This paper provides a review of the literature on unit roots and cointegration in panels where the time dimension (T) and the cross section dimension (N) are relatively large. It distinguishes between the first generation tests developed on the assumption of the cross section independence, and the second generation tests that allow, in a variety of forms and degrees, the dependence that might prevail across the different units in the panel. In the analysis of cointegration the hypothesis testing and estimation problems are further complicated by the possibility of cross section cointegration which could arise if the unit roots in the different cross section units are due to common random walk components.

JEL Code: C12, C15, C22, C23.

Keywords: panel unit roots, panel cointegration, cross section dependence, common effects.

Joerg Breitung
University of Bonn
Department of Economics
Institute of Econometrics
Adenauerallee 24 – 42
53113 Bonn
Germany
breitung@uni-bonn.de

M. Hashem Pesaran
Cambridge University
Sidgwick Avenue
Cambridge, CB3 9DD
United Kingdom
mhp1@econ.cam.ac.uk

We are grateful to Jushan Bai, Badi Baltagi, George Kapetanios, Uwe Hassler, Serena Ng, Elisa Tosetti, Ron Smith, and Joakim Westerlund for comments on a preliminary version of this paper.

1 Introduction

Recent advances in time series econometrics and panel data analysis have focussed attention on unit root and cointegration properties of variables observed over a relatively long span of time across a large number of cross section units, such as countries, regions, companies or even households. Such panel data sets have been used predominately in testing the purchasing power parity and output convergence, although the panel techniques have also been adapted more recently to the analysis of business cycle synchronization, house price convergence, regional migration and household income dynamics. This paper provides a review of the theoretical literature on testing for unit roots and cointegration in panels where the time dimension (T), and the cross section dimension (N) are relatively large. In cases where N is large (say over 100) and T small (less than 10) the analysis can proceed only under restrictive assumptions such as dynamic homogeneity and/or local cross section dependence as in spatial autoregressive or moving average models. In cases where N is small (less than 10) and T is relatively large standard time series techniques applied to systems of equations, such as the Seemingly Unrelated Regression Equations (SURE), can be used and the panel aspect of the data should not pose new technical difficulties.

One of the primary reasons behind the application of unit root and cointegration tests to a panel of cross section units was to gain statistical power and to improve on the poor power of their univariate counterparts. This was supported by the application of what might be called the first generation panel unit root tests to real exchange rates, output and inflation. For example, the augmented Dickey-Fuller test is typically not able to reject the hypothesis that the real exchange rate is nonstationary. In contrast, panel unit root tests applied to a collection of industrialized countries generally find that real exchange rates are stationary, thereby lending empirical support to the purchasing power parity hypothesis (e.g. Coakley and Fuertes (1997) and Choi (2001)).

Unfortunately, testing the unit root and cointegration hypotheses by using panel data instead of individual time series involves several additional complications. First, panel data generally introduce a substantial amount of unobserved heterogeneity, rendering the parameters of the model cross section specific. Second, in many empirical applications, particularly the application to the real exchange rates mentioned above, it is inappropriate to assume that the cross section units are independent. To overcome these difficulties, variants of panel unit root tests are developed that allow for dif-

ferent forms of cross sectional dependence.[1] Third, the panel test outcomes are often difficult to interpret if the null of the unit root or cointegration is rejected. The best that can be concluded is that "a significant fraction of the cross section units is stationary or cointegrated". The panel tests do not provide explicit guidance as to the size of this fraction or the identity of the cross section units that are stationary or cointegrated. Fourth, with unobserved $I(1)$ (i.e. integrated of order unity) common factors affecting some or all the variables in the panel, it is also necessary to consider the possibility of cointegration between the variables across the groups (cross section cointegration) as well as within group cointegration. Finally, the asymptotic theory is considerably more complicated due to the fact that the sampling design involves a time as well as a cross section dimension. For example, applying the usual Dickey-Fuller test to a panel data set introduces a bias that is not present in the case of a univariate test. Furthermore, a proper limit theory has to take into account the relationship between the increasing number of time periods and cross section units (cf. Phillips and Moon 1999).

By comparison to panel unit root tests, the analysis of cointegration in panels is still at an early stages of its developments. So far the focus of the panel cointegration literature has been on residual based approaches, although there has been a number of attempts at the development of system approaches as well. As in the case of panel unit root tests, such tests are developed based on homogenous and heterogeneous alternatives. The residual based tests were developed to ward against the "spurious regression" problem that can also arise in panels when dealing with $I(1)$ variables. Such tests are appropriate when it is known *a priori* that at most there can be only one within group cointegration in the panel. System approaches are required in more general settings where more than one within group cointegrating relation might be present, and/or there exist unobserved common $I(1)$ factors.

Having established a cointegration relationship, the long-run parameters can be estimated efficiently using techniques similar to the ones proposed in the case of single time series models. Specifically, fully-modified OLS procedures, the dynamic OLS estimator and estimators based on a vector error correction representation were adopted to panel data structures. Most approaches employ a homogenous framework, that is, the cointegration vectors are assumed to be identical for all panel units, whereas the short-run parameters are panel specific. Although such an assumption seems plausible

[1] In fact the application of the second generation panel unit root tests to real exchange rates tend to over-turn the earlier test results that assume the cross section units are independently distributed. See Moon and Perron (2004) and Pesaran (2005).

for some economic relationships (like the PPP hypothesis mentioned above) there are other behavioral relationships (like the consumption function or money demand), where a homogeneous framework seems overly restrictive. On the other hand, allowing all parameters to be individual specific would substantially reduce the appeal of a panel data study. It is therefore important to identify parameters that are likely to be similar across panel units whilst at the same time allowing for sufficient heterogeneity of other parameters. This requires the development of appropriate techniques for testing the homogeneity of a sub-set of parameters across the cross section units. When N is small relative to T, standard likelihood ratio based statistics can be used. Groen and Kleibergen (2003) provide an application. Testing for parameter homogeneity in the case of large panels poses new challenges that require further research. Some initial attempts are made in Pesaran, Smith and Im (1996), Phillips and Sul (2003a) and Pesaran and Yamagata (2005).

This paper reviews some recent work in this rapidly developing research area and thereby updating the earlier excellent surveys of Banerjee (1999), Baltagi and Kao (2000) and Choi (2004). The remainder of the paper is organized as follows: Section 2 sets out the basic model for the panel unit root tests and describes the first generation panel unit root tests. Second generation panel unit root tests are described in Section 3, and a brief account of the small sample properties of the panel unit root tests is provided in Section 4. General issues surrounding panel cointegration, including the problem of cross-section cointegration, are discussed in Section 5. Residual-based and system approaches to testing for cointegration in panels are reviewed in Sections 6 and 7; and estimation of the cointegration relations in panels is discussed in Section 8. Panels with unobserved common factors, allowing for cross-section cointegration, are reviewed in Section 9. Some concluding remarks are provided in Section 10.

2 First Generation Panel Unit Root Tests

2.1 The Basic Model

Assume that time series $\{y_{i0}, \ldots, y_{iT}\}$ on the cross section units $i = 1, 2, ..., N$ are generated for each i by a simple first-order autoregressive, AR(1), process

$$y_{it} = (1 - \alpha_i)\mu_i + \alpha_i y_{i,t-1} + \varepsilon_{it}, \tag{1}$$

where the initial values, y_{i0}, are given, and the errors ε_{it} are identically, independently distributed (i.i.d.) across i and t with $E(\varepsilon_{it}) = 0$, $E(\varepsilon_{it}^2) = \sigma_i^2 < \infty$ and $E(\varepsilon_{it}^4) < \infty$. These processes can also be written equivalently

as simple Dickey-Fuller (DF) regressions

$$\Delta y_{it} = -\phi_i \mu_i + \phi_i y_{i,t-1} + \varepsilon_{it} \ , \tag{2}$$

where $\Delta y_{it} = y_{it} - y_{i,t-1}$, $\phi_i = \alpha_i - 1$. In further developments of the model it is also helpful to write (1) or (2) in mean-deviations forms $\tilde{y}_{it} = \alpha_i \tilde{y}_{i,t-1} + \varepsilon_{it}$, where $\tilde{y}_{it} = y_{it} - \mu_i$. The corresponding DF regression in $\tilde{y}_{it}$ is given by

$$\Delta \tilde{y}_{it} = \phi_i \tilde{y}_{i,t-1} + \varepsilon_{it}. \tag{3}$$

The null hypothesis of interest is

$$H_0 : \ \phi_1 = \cdots = \phi_N = 0, \tag{4}$$

that is, all time series are independent random walks. We will consider two alternatives:

$$\begin{aligned} H_{1a} \ &: \ \phi_1 = \cdots = \phi_N \equiv \phi \text{ and } \ \phi < 0 \\ H_{1b} \ &: \ \phi_1 < 0 \ , \ \cdots \ , \ \phi_{N_0} < 0, \ \ N_0 \leq N. \end{aligned}$$

Under H_{1a} it is assumed that the autoregressive parameter is identical for all cross section units (see, for example, Levin and Lin (1993, LL), and Levin, Lin and Chu 2002). This is called the *homogeneous alternative.* H_{1b} assumes that N_0 of the N $(0 < N_0 \leq N)$ panel units are stationary with individual specific autoregressive coefficients. This is referred to as the heterogeneous alternatives (see, for example, Im, Pesaran and Shin (2003, IPS). For the consistency of the test it is assumed that $N_0/N \to \kappa > 0$ as $N \to \infty$. Different panel testing procedures can be developed depending on which of the two alternatives is being considered. The panel unit root statistics motivated by the first alternative, H_{1a}, pools the observations across the different cross section units before forming the "pooled" statistic, whilst the tests developed against the heterogeneous alternatives, H_{1b}, operates directly on the test statistics for the individual cross section units using (standardized) simple averages of the underlying individual statistics or their suitable transformations such as rejection probabilities. Despite the differences in the way the two tests view the alternative hypothesis both tests can be consistent against both types of the alternatives. Also interpretation of the outcomes of both tests is subject to similar considerations discussed in the introduction. When the null hypothesis is rejected one can only conclude that a significant fraction of the AR(1) processes in the panel does not contain unit roots.

2.2 Derivation of the Tests

The various first generation panel unit roots proposed in the literature can be obtained using the pooled log-likelihood function of the individual Dickey-Fuller regressions given by (2).

$$\ell_{NT}(\boldsymbol{\phi},\boldsymbol{\theta}) = \sum_{i=1}^{N}\left\{-\frac{T}{2}\log 2\pi\sigma_i^2 - \frac{1}{2\sigma_i^2}\sum_{t=1}^{T}\left(\Delta y_{it} + \phi_i\mu_i - \phi_i y_{i,t-1}\right)^2\right\}, \quad (5)$$

where $\boldsymbol{\phi} = (\phi_1, ..., \phi_N)'$, $\boldsymbol{\theta}_i = (\mu_i, \sigma_i^2)'$ and $\boldsymbol{\theta} = (\boldsymbol{\theta}_1', ..., \boldsymbol{\theta}_N')'$. In the case of the homogeneous alternatives, H_{1a}, where $\phi_i = \phi$, the maximum likelihood estimator of ϕ is given by

$$\hat{\phi}(\boldsymbol{\theta}) = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}\sigma_i^{-2}\Delta y_{it}\left(y_{i,t-1} - \mu_i\right)}{\sum_{i=1}^{N}\sum_{t=1}^{T}\sigma_i^{-2}\left(y_{i,t-1} - \mu_i\right)^2}. \quad (6)$$

The nuisance cross-section specific parameters $\boldsymbol{\theta}_i$ can be estimated either under the null or the alternative hypothesis. Under the null hypothesis μ_i is unidentified, but as we shall see it is often replaced by y_{i0}, on the implicit (identifying) assumption that $\tilde{y}_{i0} = 0$ for all i. For this choice of μ_i the effective number of time periods used for estimation of ϕ_i is reduced by one. Under the alternative hypothesis the particular estimates of μ_i and σ_i^2 chosen naturally depend on the nature of the alternatives envisaged. Under homogeneous alternatives, $\phi_i = \phi < 0$, the ML estimates of μ_i and σ_i^2 are given as non-linear functions of $\hat{\phi}$. Under heterogeneous alternatives ϕ_i and σ_i^2 can be treated as free parameters and estimated separately for each i.

Levin, Lin and Chu (2002) avoid the problems associated with the choice of the estimators for μ_i and base their tests on the t-ratio of ϕ in the pooled fixed-effects regression

$$\Delta y_{it} = a_i + \phi y_{i,t-1} + \varepsilon_{it},\ \varepsilon_{it} \backsim i.i.d.(0, \sigma_i^2).$$

The t-ratio of the FE estimator of ϕ is given by

$$\tau_\phi = \frac{\sum_{i=1}^{N}\widehat{\sigma}_i^{-2}\Delta\mathbf{y}_i'\mathbf{M}_\tau\mathbf{y}_{i,-1}}{\sqrt{\sum_{i=1}^{N}\widehat{\sigma}_i^{-2}\left(\mathbf{y}_{i,-1}'\mathbf{M}_\tau\mathbf{y}_{i,-1}\right)}} \quad (7)$$

where $\Delta\mathbf{y}_i = (\Delta y_{i1}, \Delta y_{i2}, \ldots, \Delta y_{iT})'$, $\mathbf{y}_{i,-1} = (y_{i0}, y_{i1}, \ldots, y_{i,T-1})'$, $\mathbf{M}_\tau = \mathbf{I}_T - \boldsymbol{\tau}_T(\boldsymbol{\tau}_T'\boldsymbol{\tau}_T)^{-1}\boldsymbol{\tau}_T'$, $\boldsymbol{\tau}_T$ is a $T \times 1$ vector of ones,

$$\hat{\sigma}_i^2 = \frac{\Delta\mathbf{y}_i'\mathbf{M}_i\Delta\mathbf{y}_i}{T-2}, \quad (8)$$

$\mathbf{M}_i = \mathbf{I}_T - \mathbf{X}_i(\mathbf{X}_i'\mathbf{X}_i)^{-1}\mathbf{X}_i'$, and $\mathbf{X}_i = (\boldsymbol{\tau}_T, \mathbf{y}_{i,-1})$.

The construction of a test against H_{1b} is less clear because the alternative consists of a set of inequality conditions. Im, Pesaran and Shin (1995, 2003) suggest the mean of the individual specific t-statistics[2]

$$\bar{\tau} = \frac{1}{N}\sum_{i=1}^{N} \tau_i \ , \tag{9}$$

where

$$\tau_i = \frac{\Delta\mathbf{y}_i'\mathbf{M}_\tau\mathbf{y}_{i,-1}}{\hat{\sigma}_i\left(\mathbf{y}_{i,-1}'\mathbf{M}_\tau\mathbf{y}_{i,-1}\right)^{1/2}}, \tag{10}$$

is the Dickey-Fuller t-statistic of cross section unit i.[3] LM versions of the t-ratios of ϕ and ϕ_i, that are analytically more tractable, can also be used which are given by

$$\tilde{\tau}_\phi = \frac{\sum_{i=1}^{N} \tilde{\sigma}_i^{-2}\Delta\mathbf{y}_i'\mathbf{M}_\tau\mathbf{y}_{i,-1}}{\sqrt{\sum_{i=1}^{N} \tilde{\sigma}_i^{-2}\left(\mathbf{y}_{i,-1}'\mathbf{M}_\tau\mathbf{y}_{i,-1}\right)}}, \tag{11}$$

and

$$\tilde{\tau}_i = \frac{\Delta\mathbf{y}_i'\mathbf{M}_\tau\mathbf{y}_{i,-1}}{\tilde{\sigma}_i\left(\mathbf{y}_{i,-1}'\mathbf{M}_\tau\mathbf{y}_{i,-1}\right)^{1/2}}, \tag{12}$$

where $\tilde{\sigma}_i^2 = (T-1)^{-1}\Delta\mathbf{y}_i'\mathbf{M}_\tau\Delta\mathbf{y}_i$. It is easily established that the panel unit root tests based on τ_ϕ and $\tilde{\tau}_\phi$ in the case of the pooled versions, and those based on $\bar{\tau}$ and $\bar{\tilde{\tau}} = N^{-1}\sum_{i=1}^{N}\tilde{\tau}_i$, in the case of their mean group versions are asymptotically equivalent.

2.3 Null Distribution of the Tests

To establish the distribution of $\tilde{\tau}_\phi$ and $\bar{\tilde{\tau}}$, we first note that under $\phi_i = 0$, $\Delta\mathbf{y}_i = \sigma_i\mathbf{v}_i = \sigma_i(v_{i1}, v_{i2}, ..., v_{iT})'$, where $\mathbf{v}_i \backsim (\mathbf{0}, \mathbf{I}_T)$ and $\mathbf{y}_{i,-1}$ can be written

[2] Andrews (1998) has considered optimal tests in such situations. His directed Wald statistic that gives a high weights to alternatives close to the null (i.e. the parameter c in Andrews (1998) tends to zero) is equivalent to the mean of the individual specific test statistics.

[3] The mean of other unit-root test statistics may be used as well. For example, Smith et al. (2004) suggest to use the mean of the weighted symmetric test statistic proposed for single time series by Park and Fuller (1995) and Fuller (1996, Section 10.1.3), or the Max-ADF test proposed by Leybourne (1995) based on the maximum of the original and the time reversed Dickey-Fuller test statistics.

as

$$\mathbf{y}_{i,-1} = y_{i0}\boldsymbol{\tau}_T + \sigma_i \mathbf{s}_{i,-1}, \tag{13}$$

where y_{i0} is a given initial value (fixed or random), $\mathbf{s}_{i,-1} = (s_{i0}, s_{i1}, \ldots, s_{i,T-1})'$, with $s_{it} = \sum_{j=1}^{t} v_{ij}$, $t = 1, 2, ..., T$, and $s_{i0} = 0$. Using these results in (11) and (12) we have

$$\tilde{\tau}_\phi = \frac{\sum_{i=1}^{N}\left(\frac{\sqrt{T-1}\mathbf{v}_i'\mathbf{M}_\tau \mathbf{s}_{i,-1}}{\mathbf{v}_i'\mathbf{M}_\tau \mathbf{v}_i}\right)}{\sqrt{\sum_{i=1}^{N}\left(\frac{\mathbf{s}_{i,-1}'\mathbf{M}_\tau \mathbf{s}_{i,-1}}{\mathbf{v}_i'\mathbf{M}_\tau \mathbf{v}_i}\right)}}, \tag{14}$$

$$\bar{\tilde{\tau}} = N^{-1}\sum_{i=1}^{N}\frac{\sqrt{T-1}\mathbf{v}_i'\mathbf{M}_\tau \mathbf{s}_{i,-1}}{\left(\mathbf{v}_i'\mathbf{M}_\tau \mathbf{v}_i\right)^{1/2}\left(\mathbf{s}_{i,-1}'\mathbf{M}_\tau \mathbf{s}_{i,-1}\right)^{1/2}}.$$

It is clear that under the null hypothesis both test statistics are free of nuisance parameters and their critical values can be tabulated for all combinations of N and T assuming, for example, that ε_{it} (or v_{it}) are normally distributed. Therefore, in the case where the errors, ε_{it}, are serially uncorrelated an exact sample panel unit root test can be developed using either of the test statistics and no adjustments to the test statistics are needed. The main difference between the two tests lies in the way information on individual units are combined and their relative small sample performance would naturally depends on the nature of the alternative hypothesis being considered.

Asymptotic null distributions of the tests can also be derived depending on whether $(T, N) \to \infty$, sequentially, or when both N and $T \to \infty$, jointly. To derive the asymptotic distributions we need to work with the standardized versions of the test statistics

$$Z_{LL} = \frac{\tau_\phi - E\left(\tau_\phi\right)}{\sqrt{Var\left(\tau_\phi\right)}}, \tag{15}$$

and

$$Z_{IPS} = \frac{\sqrt{N}\left[\bar{\tau} - E(\tau_i)\right]}{\sqrt{Var(\tau_i)}}, \tag{16}$$

assuming that T is sufficiently large such that the second order moments of τ_i and τ_ϕ exist. The conditions under which τ_i has a second order moment are discussed in IPS and it is shown that when the underlying errors are normally distributed the second order moments exist for $T > 5$. For non-normal distributions the existence of the moments can be ensured by basing the IPS test on suitably truncated versions of the individual t-ratios. (see

Pesaran (2005) for further details). The exact first and second order moments of τ_i and $\tilde{\tau}_i$ for different values of T are given in IPS (2003, Table 1). Using these results it is also possible to generalize the IPS test for unbalanced panels. Suppose the number of time periods available on the ith cross section unit is T_i, the standardized IPS statistics will now be given by

$$Z_{IPS} = \frac{\sqrt{N}\left[\bar{\tau} - N^{-1}\sum_{i=1}^{N} E(\tau_{iT_i})\right]}{\sqrt{N^{-1}\sum_{i=1}^{N} Var(\tau_{iT_i})}}, \tag{17}$$

where $E(\tau_{iT_i})$ and $Var(\tau_{iT_i})$ are, respectively, the exact mean and variance of the DF statistics based on T_i observations. IPS show that for all finite $T_i > 6$, $Z_{IPS} \xrightarrow{d} \mathcal{N}(0,1)$ as $N \to \infty$. Similar results follow for the LL test.

To establish the asymptotic distribution of the panel unit root tests in the case of $T \to \infty$, we first note that for each i

$$\tau_i \xrightarrow{d} \eta_i = \frac{\int_0^1 \widetilde{W}_i(a) d\widetilde{W}_i(a)}{\int_0^1 \widetilde{W}_i(a)^2 da},$$

where $\widetilde{W}_i(a)$ is a demeaned Brownian motion defined as $\widetilde{W}_i(a) = W_i(a) - \int_0^1 W_i(a)da$ and $W_1(a), \ldots, W_N(a)$ are independent standard Brownian motions. The existence of the moments of η_i are established in Nabeya (1999) who also provides numerical values for the first six moments of the DF distribution for the three standard specifications; namely models with and without intercepts and linear trends. Therefore, since the individual Dickey-Fuller statistics $\tau_1, \ldots, \tau_N$ are independent, it follows that $\eta_1, \eta_2, ...\eta_N$ are also independent with finite moments. Hence, by standard central limit theorems we have

$$Z_{IPS} \xrightarrow[T\to\infty]{d} \frac{\sqrt{N}\left[\bar{\eta} - E(\eta_i)\right]}{\sqrt{Var(\eta_i)}} \xrightarrow[N\to\infty]{d} \mathcal{N}(0,1),$$

where $\bar{\eta} = N^{-1}\sum_{i=1}^{N} \eta_i$. Similarly,

$$Z_{LL} = \frac{\tau_\phi - E(\tau_\phi)}{\sqrt{Var(\tau_\phi)}} \xrightarrow[(T,N)\to\infty]{d} \mathcal{N}(0,1).$$

To simplify the exposition the above asymptotic results are derived using a sequential limit theory, where $T \to \infty$ is followed by $N \to \infty$. However, Phillips and Moon (1999) show that sequential convergence does not imply joint convergence so that in some situations the sequential limit theory may break down. In the case of models with serially uncorrelated errors, IPS (2003) show that the t-bar test is in fact valid for N and $T \to \infty$ jointly.

Also as we shall see it is conjectured that the IPS test is valid for the case of serially correlated errors as N and $T \to \infty$ so long as $N/T \to k$ where k is a finite non-zero constant.

Maddala and Wu (1999) and Choi (2001) independently suggested a test against the heterogenous alternative H_{1b} that is based on the p-values of the individual statistic as originally suggested by Fisher (1932). Let π_i denote the p-value of the individual specific unit-root test applied to cross-section unit i. The combined test statistic is

$$\overline{\pi} = -2 \sum_{i=1}^{N} \log(\pi_i). \tag{18}$$

Another possibility would be to use the inverse normal test defined by

$$Z_{INV} = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \Phi^{-1}(\pi_i), \tag{19}$$

where $\Phi(\cdot)$ denotes the cdf of the standard normal distribution. An important advantage of this approach is that it is possible to allow for different specifications (such as different deterministic terms and lag orders) for each panel unit.

Under the null hypothesis $\overline{\pi}$ is χ^2 distributed with $2N$ degrees of freedom. For large N the transformed statistic

$$\overline{\pi}^* = -\frac{1}{\sqrt{N}} \sum_{i=1}^{N} [\log(\pi_i) + 1], \tag{20}$$

is shown to have a standard normal limiting null distribution as $T, N \to \infty$, sequentially.

2.4 Asymptotic Power of the Tests

It is interesting to compare the asymptotic power of the test statistics against the sequence of local alternatives

$$H_\ell \ : \quad \alpha_{i,NT} = 1 - \frac{c_i}{T\sqrt{N}} \ . \tag{21}$$

Following Breitung (2000) and Moon, Perron and Phillips (2003) the asymptotic distribution under H_ℓ is obtained as $Z_j \xrightarrow{d} \mathcal{N}(-\bar{c}\,\theta_j, 1)$, $j = LL, IPS$, where $\bar{c} = \lim_{N\to\infty} N^{-1} \sum_{i=1}^{N} c_i$ and

$$\theta_1 = \sqrt{E\left(\int_0^1 \widetilde{W}_i(a)^2 da\right)}, \ \theta_2 = \frac{E\left(\sqrt{\int_0^1 \widetilde{W}_i(a)^2 da}\right)}{\sqrt{Var(\tau_i)}} \ .$$

It is interesting to note that the local power of both test statistics depend on the mean $\bar{c}$. Accordingly, the test statistics do not exploit the deviations from the mean value of the autoregressive parameter.

Moon, Perron and Phillips (2003) derive the most powerful test statistic against the local alternative (21). Assume that we (randomly) choose the sequence $c_1^*, \ldots, c_N^*$ instead of the unknown values $c_1, \ldots, c_N$. The point optimal test statistic is constructed using the (local-to-unity) pseudo differences

$$\Delta_{c_i^*} y_{it} = y_{it} - (1 - c_i^*/T\sqrt{N}) y_{i,t-1} \text{ for } t = 1, \ldots, T.$$

For the model without individual constants and homogeneous variances the point optimal test results in the statistic

$$V_{NT} = \frac{1}{\widehat{\sigma}^2} \left(\sum_{i=1}^{N} \sum_{t=1}^{T} (\Delta_{c_i^*} y_{it})^2 - (\Delta y_{it})^2 \right) - \frac{1}{2}\kappa^2 ,$$

where $E(c_i^*)^2 = \kappa^2$. Under the sequence of local alternatives (21) Moon et al. (2003, Theorem 7) derive the limiting distribution as

$$V_{NT} \xrightarrow{d} \mathcal{N}\left(-E(c_i c_i^*), 2\kappa^2\right).$$

The upper bound of the local power is achieved with $c_i = c_i^*$, that is, if the local alternatives used to construct the test coincide with the actual alternative. Unfortunately, in practice it seems extremely unlikely that one could select values of c_i^* that are perfectly correlated with the true values, c_i. If, on the other hand, the variates c_i^* are independent of c_i, then the power is smaller than the power of a test using identical values $c_i^* = c^*$ for all i. This suggests that if there is no information about variation of c_i, then a test cannot be improved by taking into account a possible heterogeneity of the alternative.

2.5 Heterogeneous Trends

To allow for more general mean functions we consider the model:

$$y_{it} = \boldsymbol{\delta}_i' \mathbf{d}_{it} + \tilde{y}_{it} , \tag{22}$$

where $\mathbf{d}_{it}$ represents the deterministics and $\Delta\tilde{y}_{it} = \phi_i \tilde{y}_{i,t-1} + \varepsilon_{it}$. For the model with a constant mean we let $\mathbf{d}_{it} = 1$ and the model with individual specific time trends $\mathbf{d}_{it}$ is given by $\mathbf{d}_{it} = (1, t)'$. Furthermore, structural breaks in the mean function can be accommodated by including (possibly individual specific) dummy variables in the vector $\mathbf{d}_{it}$. The parameter vector

$\boldsymbol{\delta}_i$ is assumed to be unknown and has to be estimated. For the Dickey-Fuller test statistic the mean function is estimated under the alternative, that is, for the model with a time trend $\hat{\boldsymbol{\delta}}_i'\mathbf{d}_{it}$ results from a regression of y_{it} on a constant and t ($t = 1, 2, \ldots, T$). Alternatively, the mean function can also be estimated under the null hypothesis (cf. Schmidt and Phillips, 1992) or under a local alternative (Elliott et al., 1996).

Including deterministic terms may have an important effect on the asymptotic properties of the test. Let $\Delta\widehat{\tilde{y}}_t$ and $\widehat{\tilde{y}}_{i,t-1}$ denote estimates for $\Delta\tilde{y}_{it} = \Delta y_{it} - E(\Delta y_{it})$ and $\tilde{y}_{i,t-1} = y_{i,t-1} - E(y_{i,t-1})$. In general, running the regression

$$\Delta\widehat{\tilde{y}}_{it} = \phi\widehat{\tilde{y}}_{i,t-1} + e_{it}$$

does not render a t-statistic with a standard normal limiting distribution due to the fact that $\widehat{\tilde{y}}_{i,t-1}$ is correlated with e_{it}. For example, if $\mathbf{d}_{it}$ is an individual specific constant such that $\widehat{\tilde{y}}_{i,t-1} = y_{it} - T^{-1}(y_{i0} + \cdots + y_{i,T-1})$ we obtain under the null hypothesis

$$\lim_{T\to\infty} \frac{1}{T} E\left\{\sum_{t=1}^{T} e_{it}\widehat{\tilde{y}}_{i,t-1}\right\} = -\sigma_i^2/2 \ .$$

It follows that the t-statistic of $\phi = 0$ tends to $-\infty$ as N or T tends to infinity.

To correct for the bias, Levin et al. (2002) suggested using the correction terms

$$a_T(\hat{\boldsymbol{\delta}}) = E\left(\frac{1}{\sigma_i^2 T}\sum_{t=1}^{T}\Delta\widehat{\tilde{y}}_{it}\widehat{\tilde{y}}_{i,t-1}\right) \tag{23}$$

$$b_T^2(\hat{\boldsymbol{\delta}}) = \frac{Var\left(T^{-1}\sum_{t=1}^{T}\Delta\widehat{\tilde{y}}_{it}\widehat{\tilde{y}}_{i,t-1}\right)}{\sigma_i^2 E\left(T^{-1}\sum_{t=1}^{T}\widehat{\tilde{y}}_{i,t-1}^2\right)} \tag{24}$$

where $\boldsymbol{\delta} = (\hat{\boldsymbol{\delta}}_1', \hat{\boldsymbol{\delta}}_2', ..., \hat{\boldsymbol{\delta}}_N')'$, and $\hat{\boldsymbol{\delta}}_i$ is the estimator of the coefficients of the deterministics, $\mathbf{d}_{it}$, in the OLS regression of y_{it} on $\mathbf{d}_{it}$. The corrected, standardized statitic is given by

$$Z_{LL}(\hat{\boldsymbol{\delta}}) = \frac{\left[\sum_{i=1}^{N}\sum_{t=1}^{T}\Delta\widehat{\tilde{y}}_{it}\widehat{\tilde{y}}_{i,t-1}/\widehat{\sigma}_i^2\right] - NTa_T(\hat{\boldsymbol{\delta}})}{b_T(\hat{\boldsymbol{\delta}})\sqrt{\sum_{i=1}^{N}\sum_{t=1}^{T}\widehat{\tilde{y}}_{i,t-1}^2/\widehat{\sigma}_i^2}} \ .$$

Levin et al. (2002) present simulated values of $a_T(\hat{\boldsymbol{\delta}})$ and $b_T(\hat{\boldsymbol{\delta}})$ for models with constants, time trends and various values of T. A problem is, however, that for unbalanced data sets no correction terms are tabulated.

Alternatively, the test statistic may be corrected such that the adjusted t-statistic

$$Z^*_{LL}(\hat{\boldsymbol{\delta}}) = [Z_{LL}(\hat{\boldsymbol{\delta}}) - a^*_T(\hat{\boldsymbol{\delta}})]/b^*_T(\hat{\boldsymbol{\delta}})$$

is asymptotically standard normal. Harris and Tzavalis (1999) derive the small sample values of $a^*_T(\hat{\boldsymbol{\delta}})$ and $b^*_T(\hat{\boldsymbol{\delta}})$ for T fixed and $N \to \infty$. Therefore, their test statistic can be applied for small values of T and large values of N.

An alternative approach is to avoid the bias – and hence the correction terms – by using alternative estimates of the deterministic terms. Breitung and Meyer (1994) suggest using the initial value y_{i0} as an estimator of the constant term. As argued by Schmidt and Phillips (1992), the initial value is the best estimate of the constant given the null hypothesis is true. Using this approach the regression equation for a model with a constant term becomes

$$\Delta y_{it} = \phi^*(y_{i,t-1} - y_{i0}) + v_{it} \ .$$

Under the null hypothesis, the pooled t-statistic of $H_0 : \phi^* = 0$ has a standard normal limit distribution.

For a model with a linear time trend a minimal invariant statistic is obtained by the transformation (cf. Ploberger and Phillips 2002)

$$x^*_{it} = y_{it} - y_{i0} - \frac{t}{T}(y_{iT} - y_{i0}) \ .$$

In this transformation subtracting y_{i0} eliminates the constant and $(y_{iT} - y_{i0})/T = (\Delta y_{i1} + \cdots + \Delta y_{iT})/T$ is an estimate of the slope of the individual trend function.

To correct for the mean of Δy_{it} a Helmert transformation can be used

$$\Delta y^*_{it} = s_t \left[\Delta y_{it} - \frac{1}{T-t}(\Delta y_{i,t+1} + \cdots + \Delta y_{iT}) \right] \ , t = 1, \ldots, T-1$$

where $s^2_t = (T-t)/(T-t+1)$ (cf. Arellano, 2003, p. 17). Using these transformations the regression equation becomes

$$\Delta y^*_{it} = \phi^* x^*_{i,t-1} + v_{it} \ . \tag{25}$$

It is not difficult to verify that under the null hypothesis $E(\Delta y^*_{it} x^*_{i,t-1}) = 0$ and, thus, the t-statistic for $\phi^* = 0$ is asymptotically standard normally distributed (cf. Breitung, 2000).

The test against heterogeneous alternatives H_{1b} can be easily adjusted for individual specific deterministic terms such as linear trends or seasonal dummies. This can be done by computing IPS statistics, defined by (16) and (17) for the balanced and unbalanced panels, using Dickey-Fuller t-statistics based on DF regressions including the deterministics $\boldsymbol{\delta}_i'\mathbf{d}_{it}$, where $\mathbf{d}_{it} = 1$ in the case of a constant term, $\mathbf{d}_{it} = (1, t)'$ in the case of models with a linear time trend and so on. The mean and variance corrections should, however, be computed to match the nature of the deterministics. Under a general setting IPS (2003) have shown that the Z_{IPS} statistic converges in distribution to a standard normal variate as $N, T \to \infty$, jointly.

In a straightforward manner it is possible to include dummy variables in the vector $\mathbf{d}_{it}$ that accommodate structural breaks in the mean function (see, e.g., Murray and Papell, 2002; Tzavalis, 2002; Carrion-I-Sevestre, Del Barrio and Lopez-Bazo, 2004; Breitung and Candelon, 2005; Im, Lee and Tieslau, 2005).

2.6 Short-run Dynamics

If it is assumed that the error in the autoregression (1) is a serially correlated stationary process, the short-run dynamics of the errors can be accounted for by including lagged differences

$$\Delta y_{it} = \boldsymbol{\delta}_i'\mathbf{d}_{it} + \phi_i y_{i,t-1} + \gamma_{i1}\Delta y_{i,t-1} + \cdots + \gamma_{i,p_i}\Delta y_{i,t-p_i} + \varepsilon_{it} \ . \tag{26}$$

For example, the IPS statistics (16) and (17) developed for balanced and unbalanced panels can now be constructed using the ADF(p_i) statistics based on the above regressions. As noted in IPS (2003), small sample properties of the test can be much improved if the standardization of the IPS statistic is carried out using the simulated means and variances of $\tau_i(p_i)$, the t-ratio of ϕ_i computed based on ADF(p_i) regressions. This is likely to yield better approximations, since $E\left[\tau_i(p_i)\right]$, for example, makes use of the information contained in p_i while $E\left[\tau_i(0)\right] = E(\tau_i)$ does not. Therefore, in the serially correlated case IPS propose the following standardized t-bar statistic

$$Z_{IPS} = \frac{\sqrt{N}\left\{\bar{\tau} - \frac{1}{N}\sum_{i=1}^{N} E\left[\tau_i(p_i)\right]\right\}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N} Var\left[\tau_i(p_i)\right]}} \xrightarrow[(T,N)\to\infty]{d} \mathcal{N}(0,1). \tag{27}$$

The value of $E\left[\tau_i(p)\right]$ and $Var\left[\tau_i(p)\right]$ simulated for different values of T and p, are provided in Table 3 of IPS. These simulated moments also allow the IPS panel unit root test to be applied to unbalanced panels with serially correlated errors.

For tests against the homogenous alternatives, $\phi_1 = \cdots = \phi_N = \phi < 0$, Levin et al. (2002) suggest removing all individual specific parameters within a first step regression such that e_{it} ($v_{i,t-1}$) are the residuals from a regression of Δy_{it} ($y_{i,t-1}$) on $\Delta y_{i,t-1}, \ldots, \Delta y_{i,t-p_i}$ and d_{it}. In the second step the common parameter ϕ is estimated from a pooled regression

$$(e_{it}/\widehat{\sigma}_i) = \phi(v_{i,t-1}/\widehat{\sigma}_i) + \nu_{it} \ ,$$

where $\widehat{\sigma}_i^2$ is the estimated variance of e_{it}. Unfortunately, the first step regressions are not sufficient to remove the effect of the short-run dynamics on the null distribution of the test. Specifically,

$$\lim_{T\to\infty} E\left[\frac{1}{T-p}\sum_{t=p+1}^{T} e_{it}v_{i,t-1}/\sigma_i^2\right] \ = \ \frac{\bar{\sigma}_i}{\sigma_i}\ a_\infty(\hat{\boldsymbol{\delta}}) \ ,$$

where $\overline{\sigma}_i^2$ is the long-run variance and $a_\infty(\hat{\boldsymbol{\delta}})$ denotes the limit of the correction term given in (23). Levin et al. (2002) propose a nonparametric (kernel based) estimator for $\bar{\sigma}_i^2$

$$\bar{s}_i^2 = \frac{1}{T}\ \left[\sum_{t=1}^{T}\Delta\widehat{\widetilde{y}}_{it}^2 + 2\sum_{l=1}^{K}\left(\frac{K+1-l}{K+1}\right)\left(\sum_{t=l+1}^{T}\Delta\widehat{\widetilde{y}}_{it}\Delta\widehat{\widetilde{y}}_{i,t-l}\right)\right], \qquad (28)$$

where $\Delta\widehat{\widetilde{y}}_{it}$ denotes the demeaned difference and K denotes the truncation lag. As noted by Phillips and Ouliaris (1990), in a time series context the estimator of the long-run variance based on differences is inappropriate since under the stationary alternative $\bar{s}_i^2 \xrightarrow{p} 0$ and, thus, using this estimator yields an inconsistent test. In contrast, in the case of panels the use of $\bar{s}_i^2$ improves the power of the test, since with $\bar{s}_i^2 \xrightarrow{p} 0$ the correction term drops out and the test statistic tends to $-\infty$.

It is possible to avoid the use of a kernel based estimator of the long-run variance by using an alternative approach suggested by Breitung and Das (2005a). Under the null hypothesis we have

$$\gamma_i(L)\Delta y_{it} = \boldsymbol{\delta}_i'\mathbf{d}_{it} + \varepsilon_{it} \ ,$$

where $\gamma_i(L) = 1 - \gamma_{i1}L - \cdots - \gamma_{i,p_i}L^p$ and L is the lag operator. It follows that $\widetilde{g}_t = \gamma_i(L)[y_{it} - E(y_{it})]$ is a random walk with uncorrelated increments. Therefore, the serial correlation can be removed by replacing y_{it} by the pre-whitened variable $\hat{y}_{it} = \widehat{\gamma}_i(L)y_{it}$, where $\widehat{\gamma}_i(L)$ is an estimator of the lag polynomial obtained from the least-square regression

$$\Delta y_{it} = \boldsymbol{\delta}_i'\mathbf{d}_{it} + \gamma_{i1}\Delta y_{i,t-1} + \cdots + \gamma_{i,p_i}\Delta y_{i,t-p_i} + \varepsilon_{it} \ . \qquad (29)$$

This approach may also be used for modifying the "unbiased statistic" based on the t-statistic of $\phi^* = 0$ in (25). The resulting t-statistic has a standard normal limiting distribution as $T \to \infty$ is followed by $N \to \infty$.

Pedroni and Vogelsang (2005) have proposed a test statistic that avoids the specification of the short-run dynamics by using an autoregressive approximation. Their test statistic is based on the pooled variance ratio statistic

$$Z_{NT}^{w} = \frac{Tc_i(0)}{N\widehat{s}_i^2},$$

where $c_i(\ell) = T^{-1}\sum_{t=\ell+1}^{T} \widehat{\tilde{y}}_{it}\widehat{\tilde{y}}_{i,t-\ell}$, $\widehat{\tilde{y}}_{it} = y_{it} - \hat{\boldsymbol{\delta}}_i'\mathbf{d}_{it}$ and $\widehat{s}_i^2$ is the untruncated Bartlett kernel estimator defined as $\widehat{s}_i^2 = \sum_{\ell=-T+1}^{T+1}(1 - |\ell|/T)c_i(\ell)$. As has been shown by Kiefer and Vogelsang (2002) and Breitung (2002), the limiting distribution of such "nonparametric" statistics does not depend on nuisance parameters involved by the short run dynamics of the processes. Accordingly, no adjustment for short-run dynamics is necessary.

2.7 Other Approaches to Panel Unit Root Testing

An important problem of combining Dickey-Fuller type statistics in a panel unit root test is that they involve a nonstandard limiting distribution. If the panel unit root statistic is based on a standard normally distributed test statistic z_i, then $N^{-1/2}\sum_{i=1}^{N} z_i$ has a standard normal limiting distribution even for a finite N. In this case no correction terms need to be tabulated to account for the mean and the variance of the test statistic.

Chang (2002) proposes a nonlinear instrumental variable (IV) approach, where the transformed variable

$$w_{i,t-1} = y_{i,t-1}e^{-c_i|y_{i,t-1}|}$$

is used as an instrument for estimating ϕ_i in the regression $\Delta y_{it} = \phi_i y_{i,t-1} + \varepsilon_{it}$ (which may also include deterministic terms and lagged differences). Since $w_{i,t-1}$ tends to zero as $y_{i,t-1}$ tends to $\pm\infty$ the trending behavior of the nonstationary variable $y_{i,t-1}$ is eliminated. Using the results of Chang, Park and Phillips (2001), Chang (2002) showed that the Wald test of $\phi = 0$ based on the nonlinear IV estimator possesses a standard normal limiting distribution. Another important property of the test is that the nonlinear transformation also takes account of possible contemporaneous dependence among the cross section units. Accordingly, Chang's panel unit root test is also robust against cross-section dependence.

It should be noted that $w_{i,t-1} \in [-(c_i e)^{-1}, (c_i e)^{-1}]$ with a maximum (minimum) at $y_{i,t-1} = 1/c_i$ ($y_{i,t-1} = -1/c_i$). Therefore, the choice of the parameter c_i is crucial for the properties of the test. First, the parameter should

be proportional to inverse of the standard deviations of Δy_{it}. Furthermore, in her simulation experiments Chang (2002) seems to have calibrated the values of c_i according to the sample sizes. However, as emphasized by Im and Pesaran (2003) the asymptotic theory for this test is based on the fact that c_i is a constant.

An alternative approach to obtain an asymptotically standard normal test statistic is to adjust the given samples in all cross-sections so that they all have sums of squares $y_{i1}^2 + \cdots + y_{ik_i}^2 = \sigma_i^2 c T^2 + h_i$, where $h_i \xrightarrow{p} 0$ as $T \to \infty$. In other words, the panel data set becomes an unbalanced panel with k_i time periods in the i'th unit. Chang calls this setting the "equi-squared sum contour", whereas the traditional framework is called the "equi-sample-size contour". The nice feature of this approach is that it yields asymptotically standard normal test statistics. An important drawback is, however, that a large number of observations may be discarded by applying this contour which may result in a severe loss of power.

Hassler, Demetrescu and Tarcolea (2004) have suggested to use the LM statistic for a fractional unit root as an asymptotically normally distributed test statistic. This test statistic is uniformly most powerful against fractional alternatives of the form $(1-L)^d y_{it} = \varepsilon_{it}$ with $d < 1$. Although usually panel unit root tests are used to decide whether the series are $I(1)$ or $I(0)$, it can be argued that fractional unit root tests also have a good (albeit not optimal) power against the $I(0)$ alternative (e.g. Robinson, 1994).

As in the time series case it is possible to test the null hypothesis that the series are stationary against the alternative that (at least some of) the series are nonstationary. The test suggested by Tanaka (1990) and Kwiatkowski et al. (1992) is designed to test the hypothesis $H_0^* : \theta_i = 0$ in the model

$$y_{it} = \boldsymbol{\delta}_i' \mathbf{d}_{it} + \theta_i r_{it} + u_{it} \ , \ t = 1, \ldots, T, \tag{30}$$

where Δr_{it} is white noise with unit variance and u_{it} is stationary. The cross-section specific KPSS statistic is

$$\kappa_i = \frac{1}{T^2 \bar{\sigma}_{T,i}^2} \sum_{t=1}^{T} \widehat{S}_{it}^2 \ ,$$

where $\bar{\sigma}_{T,i}^2$ denotes a consistent estimator of the long-run variance of Δy_{it} and $\widehat{S}_{it} = \sum_{\ell=1}^{t} \left(y_{i\ell} - \widehat{\boldsymbol{\delta}}_i' \mathbf{d}_{i\ell} \right)$ is the partial sum of the residuals from a regression of y_{it} on the deterministic terms (a constant or a linear time trend). The individual test statistics can be combined as in the test suggested by IPS (2003) yielding

$$\bar{\kappa} = N^{-1/2} \frac{\sum_{i=1}^{N} \left[\kappa_i - E(\kappa_i) \right]}{\sqrt{Var(\kappa_i)}},$$

where asymptotic values of $E(\kappa_i)$ and $Var(\kappa_i)$ are derived in Hadri (2000) and values for finite T and $N \to \infty$ are presented Hadri and Larsson (2005).

The test of Harris, Leybourne and McCabe (2004) is based on the stationarity statistic

$$Z_i(k) = \sqrt{T}\ \widehat{c}_i(k)/\widehat{\omega}_{zi}(k),$$

where $\widehat{c}_i(k)$ denotes the usual estimator of the covariance at lag k of cross section unit i and $\widehat{\omega}^2_{zi}(k)$ is an estimator of the long-run variance of $z^k_{it} = (y_{it} - \widehat{\boldsymbol{\delta}}'_i \mathbf{d}_{it})(y_{i,t-k} - \widehat{\boldsymbol{\delta}}'_i \mathbf{d}_{i,t-k})$. The intuition behind this test statistic is that for a stationary and ergodic time series we have $E[\widehat{c}_i(k)] \to 0$ as $k \to \infty$. Since $\widehat{\omega}^2_z$ is a consistent estimator for the variance of $\widehat{c}_i(k)$ it follows that $Z_i(k)$ converges to a standard normally distributed random variable as $k \to \infty$ and $k/\sqrt{T} \to \delta < \infty$.

3 Second Generation Panel Unit Root Tests: Allowing for Cross Section Correlations[4]

So far we have assumed that the time series $\{y_{it}\}_{t=0}^T$ are independent across i. However, in many macroeconomic applications using country or regional data it is found that the time series are contemporaneously correlated. Prominent examples are the analysis of purchasing power parity and output convergence.[5] The literature on modelling of cross section dependence in large panels is still developing and in what follows we provide an overview of some of the recent contributions.

Cross section dependence can arise due to a variety of factors, such as omitted observed common factors, spatial spill over effects, unobserved common factors, or general residual interdependence that could remain even when all the observed and unobserved common effects are taken into account. Abstracting from common observed effects and residual serial correlation a general specification for cross sectional error dependence can be written as

$$\Delta y_{it} = -\mu_i \phi_i + \phi_i y_{i,t-1} + u_{it},$$

where

$$u_{it} = \boldsymbol{\gamma}'_i \mathbf{f}_t + \xi_{it}, \tag{31}$$

[4] A survey of the second generation panel unit root tests is also provided by Hurlin and Mignon (2004).

[5] See, for example, O'Connell (1998) and Phillips and Sul (2003b). Tests for cross section independence of errors with applications to output growth equations are considered in Pesaran (2004a).

or

$$\mathbf{u}_t = \mathbf{\Gamma}\mathbf{f}_t + \boldsymbol{\xi}_t, \tag{32}$$

$\mathbf{u}_t = (u_{1t}, u_{2t}, ..., u_{Nt})'$, $\mathbf{f}_t$ is an $m \times 1$ vector of serially uncorrelated unobserved common factors, and $\boldsymbol{\xi}_t = (\xi_{1t}, \xi_{2t}, ..., \xi_{Nt})'$ is an $N \times 1$ vector of serially uncorrelated errors with mean zero and the positive definite covariance matrix $\mathbf{\Omega}_\xi$, and $\mathbf{\Gamma}$ is an $N \times m$ matrix of factor loadings defined by $\mathbf{\Gamma} = (\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, ..., \boldsymbol{\gamma}_N)'$. [6] Without loss of generality the covariance matrix of $\mathbf{f}_t$ is set to $\mathbf{I}_m$, and it is assumed that $\mathbf{f}_t$ and $\boldsymbol{\xi}_t$ are independently distributed. If $\boldsymbol{\gamma}_1 = \cdots = \boldsymbol{\gamma}_N$, then $\theta_t = \boldsymbol{\gamma}'\mathbf{f}_t$ is a conventional "time effect" that can be removed by subtracting the cross section means from the data. In general it is assumed that $\boldsymbol{\gamma}_i$, the factor loading for the ith cross section unit, differs across i and represents draws from a given distribution.

Under the above assumptions and conditional on $\boldsymbol{\gamma}_i$, $i = 1, 2, ..., N$, the covariance matrix of the composite errors, $\mathbf{u}_t$, is given by $\mathbf{\Omega} = \mathbf{\Gamma}\mathbf{\Gamma}' + \mathbf{\Omega}_\xi$. It is clear that without further restrictions the matrices $\mathbf{\Gamma}$ and $\mathbf{\Omega}_\xi$ are not separately identified. The properties of $\mathbf{\Omega}$ also crucially depend on the relative eigenvalues of $\mathbf{\Gamma}\mathbf{\Gamma}'$ and $\mathbf{\Omega}_\xi$, and their limits as $N \to \infty$. Accordingly two cases of cross-section dependence can be distinguished: *(i) Weak dependence.* In this case it is assumed that the eigenvalues of $\mathbf{\Omega}$ are bounded as $N \to \infty$. This assumption rules out the presence of unobserved common factors, but allows the cross section units to be, for example, spatially correlated with a finite number of "neighbors". *(ii) Strong dependence.* In this case some eigenvalues of $\mathbf{\Omega}$ are $O(N)$, which arises when there are unobserved common factors. When N is fixed as $T \to \infty$ both sources of dependence could be present. But for $N \to \infty$ (and particularly when $N > T$) it seems only sensible to consider cases where $rank(\mathbf{\Gamma}) = m \geq 1$ and $\mathbf{\Omega}_\xi$ is a diagonal matrix.

A simple example of panel data models with weak cross section dependence is given by

$$\begin{bmatrix} \Delta y_{1t} \\ \vdots \\ \Delta y_{Nt} \end{bmatrix} = \begin{bmatrix} a_1 \\ \vdots \\ a_N \end{bmatrix} + \phi \begin{bmatrix} y_{1,t-1} \\ \vdots \\ y_{N,t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ \vdots \\ u_{Nt} \end{bmatrix} \tag{33}$$

or

$$\Delta\mathbf{y}_t = \mathbf{a} + \phi\mathbf{y}_{t-1} + \mathbf{u}_t, \tag{34}$$

where $a_i = -\phi\mu_i$ and $\Delta\mathbf{y}_t$, $\mathbf{y}_{t-1}$, $\mathbf{a}$ and $\mathbf{u}_t$ are $N \times 1$ vectors. The cross-section correlation is represented by a non-diagonal matrix

$$\mathbf{\Omega} = E(\mathbf{u}_t\mathbf{u}_t'), \text{ for all } t,$$

[6] The case where $\mathbf{f}_t$ and/or ξ_{it} might be serially correlated will be considered below.

with bounded eigenvalues. For the model without constants Breitung and Das (2005a) showed that the regression t-statistic of $\phi = 0$ in (34) is asymptotically distributed as $\mathcal{N}(0, v_\Omega)$ where

$$v_\Omega = \lim_{N\to\infty} \frac{tr(\mathbf{\Omega}^2/N)}{(tr\mathbf{\Omega}/N)^2} \ . \tag{35}$$

Note that $tr(\mathbf{\Omega})$ and $tr(\mathbf{\Omega}^2)$ are $O(N)$ and, thus, v_Ω converges to a constant that can be shown to be larger than one. This explains why the test ignoring the cross-correlation of the errors has a positive size bias.

Since (34) can be seen as a seemingly unrelated regression system, O'Connell (1998) suggests to estimate the system by using a GLS estimator (see also Flores, et al., 1999). Let $\widehat{\mathbf{\Omega}} = T^{-1}\sum_{t=1}^{T} \widehat{\mathbf{u}}_t\widehat{\mathbf{u}}_t'$ denote the sample covariance matrix of the residual vector. The GLS t-statistic is given by

$$t_{gls}(N) = \frac{\sum\limits_{t=1}^{T} \Delta\widetilde{\mathbf{y}}_t'\widehat{\mathbf{\Omega}}^{-1}\widetilde{\mathbf{y}}_{t-1}}{\sqrt{\sum\limits_{t=1}^{T} \widetilde{\mathbf{y}}_{t-1}'\widehat{\mathbf{\Omega}}^{-1}\widetilde{\mathbf{y}}_{t-1}}},$$

where $\widetilde{\mathbf{y}}_t$ is the vector of demeaned variables. Harvey and Bates (2003) derive the limiting distribution of $t_{gls}(N)$ for a fixed N and as $T \to \infty$, and tabulate its asymptotic distribution for various values of N. Breitung and Das (2005a) show that if $\widetilde{\mathbf{y}}_t = \mathbf{y}_t - \mathbf{y}_0$ is used to demean the variables and $T \to \infty$ is followed by $N \to \infty$, then the GLS t-statistic possesses a standard normal limiting distribution.

The GLS approach cannot be used if $T < N$ as in this case the estimated covariance matrix $\widehat{\mathbf{\Omega}}$ is singular. Furthermore, Monte Carlo simulations suggest that for reasonable size properties of the GLS test, T must be substantially larger than N (e.g. Breitung and Das, 2005a). Maddala and Wu (1999) and Chang (2004) have suggested a bootstrap procedure that improves the size properties of the GLS test.

An alternative approach based on "panel corrected standard errors" (PCSE) is considered by Jönsson (2005) and Breitung and Das (2005a). In the model with weak dependence, the variance of the OLS estimator $\widehat{\phi}$ is consistently estimated by

$$\widehat{var}(\widehat{\phi}) = \frac{\sum\limits_{t=1}^{T} \widetilde{\mathbf{y}}_{t-1}'\widehat{\mathbf{\Omega}}\widetilde{\mathbf{y}}_{t-1}}{\left(\sum\limits_{t=1}^{T} \widetilde{\mathbf{y}}_{t-1}'\widetilde{\mathbf{y}}_{t-1}\right)^2}.$$

If $T \to \infty$ is followed by $N \to \infty$ the robust t statistic $t_{rob} = \widehat{\phi}/\sqrt{\widehat{var}(\widehat{\phi})}$ is asymptotically standard normally distributed (Breitung and Das, 2005a).

If it is assumed that the cross correlation is due to common factors then the largest eigenvalue of the error covariance matrix, $\mathbf{\Omega}$, is $O_p(N)$ and the robust PCSE approach breaks down. Specifically, Breitung and Das (2005b) showed that in this case t_{rob} is distributed as the ordinary Dickey-Fuller test applied to the first principal component.

In the case of a single unobserved common factor, Pesaran (2005) has suggested a simple modification of the usual test procedure. Let $\bar{y}_t = N^{-1}\sum_{i=1}^{N} y_{it}$ and $\Delta\bar{y}_t = N^{-1}\sum_{i=1}^{N}\Delta y_{it} = \bar{y}_t - \bar{y}_{t-1}$. The cross section augmented Dickey-Fuller (CADF) test is based on the following regression

$$\Delta y_{it} = a_i + \phi_i y_{i,t-1} + b_i\bar{y}_{t-1} + c_i\Delta\bar{y}_t + e_{it} \ .$$

In this regression the additional variables $\Delta\bar{y}_t$ and $\bar{y}_{t-1}$ are $\sqrt{N}$-consistent estimators for the rescaled factors $\bar{\gamma}f_t$ and $\bar{\gamma}\sum_{j=0}^{t-1} f_j$, where $\bar{\gamma} = N^{-1}\sum_{i=1}^{N}\gamma_i$. Pesaran (2005) showed that the distribution of the regression t-statistic for $\phi_i = 0$ is free of nuisance parameters. To test the unit root hypothesis in a heterogenous panel the average of the N individual CADF t-statistics (or suitably truncated version of them) can be used. Coakley et al. (2005) apply the CADF test to real exchange rates of 15 OECD countries.

A similar approach was proposed by Moon and Perron (2004) and Phillips and Sul (2003a). The test of Moon and Perron (2004) is based on a principal components estimator of $m < N$ common factors $f_{1t}, \ldots, f_{mt}$ in (31). The number of common factors can be consistently determined by using the information criteria suggested by Bai and Ng (2002). Let $\widehat{\mathbf{V}}_m = [\widehat{\mathbf{v}}_1, \ldots, \widehat{\mathbf{v}}_m]$ denote the matrix of m orthonormal eigenvectors associated with m largest eigenvalues of $\mathbf{\Omega}$. The vector of common factors are estimated as

$$\widehat{\mathbf{f}}_t = [\widehat{f}_{1t}, \ldots, \widehat{f}_{mt}]' = \widehat{\mathbf{V}}'_m\Delta\mathbf{y}_t \ .$$

As shown by Bai and Ng (2002), the principal component estimator $\widehat{\mathbf{f}}_t$ yields a consistent estimator of the factor space as $min(N,T) \to \infty$. Thus, the elements of the vector

$$\left(\mathbf{I}_N - \widehat{\mathbf{V}}_m\widehat{\mathbf{V}}'_m\right)\Delta\mathbf{y}_t \equiv \mathbf{Q}_{\widehat{V}_m}\Delta\mathbf{y}_t, \tag{36}$$

are consistent estimates of the idiosyncratic components ξ_{it} as $N \to \infty$. Therefore, by assuming that ξ_{it} is i.i.d., the pooled regression t-statistic

$$t^*_{MP} = \frac{\sum_{t=1}^{T}\Delta\widetilde{\mathbf{y}}'_t\mathbf{Q}_{\widehat{V}_m}\widetilde{\mathbf{y}}_{t-1}}{\sqrt{\sum_{t=1}^{T}\widetilde{\mathbf{y}}'_{t-1}\mathbf{Q}_{\widehat{V}_m}\widetilde{\mathbf{y}}_{t-1}}}.$$

has a standard normal limiting distribution as $(N, T \to \infty)$ and $\lim \inf_{N,T\to\infty} \log N / \log T \to 0$ (cf. Moon and Perron 2004).

As argued by Banerjee, Marcellino and Osbat (2005) panel unit root tests may be severely biased if the panel units are cross-cointegrated, namely if under the null hypothesis (of unit roots) one or more linear combinations of $\mathbf{y}_t$ are stationary. This needs to be distinguished from the case where the errors are cross correlated without necessarily involving cointegration across the cross section units. Under the former two or more cross section units must share at least one common stochastic trend. Such a situation is likely to occur if the PPP hypothesis is examined (cf. Lyhagen, 2000 and Banerjee et al. 2005). The tests proposed by Moon and Perron (2004) and Pesaran (2005) are not appropriate if $\mathbf{y}_t$ contains *common* stochastic trends and their use could lead to misleading conclusions. For example, in the presence of cross-section cointegration the common trends are identified as common factors and by performing the transformation (36) these common factors are removed from the data. If the remaining idiosyncratic component is stationary, the test tends to indicate that the time series are stationary, although all panel units are nonstationary.

To overcome this difficulty Bai and Ng (2004a) have suggested a general approach for analyzing the common factors and idiosyncratic components separately. Whereas the tests of Moon and Perron (2004) and Pesaran (2005) assume that only the idiosyncractic components have unit roots, the test procedure of Bai and Ng (2004a) allows for the possibility of unit roots in idiosyncratic and/or the common stochastic components and attempts to ascertain the unit root properties of both components from the data. For example, if the idiosyncratic processes are independent random walks, there will be no cointegration among the panel units. Nevertheless it is possible to determine the number of common trends and to estimate them consistently as N and T tends to infinity (see also Bai 2004). However, testing for unit roots in the common components are likely to require particularly large panels; with the power of the test only favourably affected when T is increased.

The common factors are estimated by principal components and cointegration tests are used to determine the number of common trends. Furthermore, panel unit root tests are applied to the idiosyncratic components. The null hypothesis that the time series have a unit root is rejected if either the test of the common factors or the test for the idiosyncratic component cannot reject the null hypothesis of nonstationary components. A similar test procedure based on KPSS test statistics are proposed by Bai and Ng (2004b).[7]

[7] An alternative factor extraction method is suggested by Kapetanios (2005) who also

To allow for short-run and long-run (cross-cointegration) dependencies, Chang and Song (2005) suggest a nonlinear instrument variable test procedure. As the nonlinear instruments suggested by Chang (2002) are invalid in the case of cross-cointegration panel specific instruments based on the Hermit function of different order are used as nonlinear instruments. Chang and Song (2005) showed that the t-statistic computed from the nonlinear IV statistic are asymptotically standard normally distributed and, therefore, a panel unit statistics against the heterogeneous alternative H_{1b} can be constructed that has an standard normal limiting distribution.

Choi and Chue (2004) employ a subsampling procedure to obtain tests that are robust against a wide range of cross-section dependence such as weak and strong correlation as well as cross-cointegration. To this end the sample is grouped into a number of overlapping blocks of b time periods. Using all $(T-b+1)$ possible overlapping blocks, the critical value of the test is estimated by the respective quantile of the empirical distribution of the $(T-b+1)$ test statistics computed. The advantage of this approach is that the null distribution of the test statistic may depend on unknown nuisance parameters. Whenever the test statistics converge in distribution to some limiting null distribution as $T \to \infty$ and N fixed, the subsample critical values converge in probability to the true critical values. Using Monte Carlo simulations Choi and Chue (2004) demonstrate that the size of the subsample test is indeed very robust against various forms of cross-section dependence.

4 Finite Sample Properties of Panel Unit Root Tests

It has become standard to distinguish first generation panel unit root tests that are based on the assumption of independent cross section units and second generation tests that allow for some kind of cross-section dependence. Maddala and Wu (1999) compared several first generation tests. For the heterogeneous alternative under consideration they found that in most cases the Fisher test (18) performs similar or slightly better than the IPS statistic with respect to size and power. The Levin and Lin statistic (in the version of the 1993 paper) performs substantially worse. Similar results are obtained by Choi (2001). Madsen (2003) derived the local power function against homogeneous alternatives under different detrending procedures. Her Monte

provides detailed Monte Carlo results on the small sample performance of panel unit root tests based on a number of alternative estimates of the unobserved common factors. He shows that the factor-based panel unit root tests tend to perform rather poorly when the unobserved common factor is serially correlated.

Carlo simulations support the theoretical findings that the test based on estimating the mean under the null hypothesis (i.e. the initial observation is subtracted from the time series) outperforms tests based on alternative demeaning procedures. Similar findings are obtained by Bond et al. (2002).

Hlouskova and Wagner (2005) compare a large number of first generation panel unit root tests applied to processes with MA(1) errors. Not surprisingly, all tests are severely biased as the root of the MA process approaches unity. Overall, the tests of Levin et al. (2002) and Breitung (2000) have the smallest size distortions. These tests also perform best against the homogenous alternative, where the autoregressive coefficient is the same for all panel units. Of course this is not surprising as these tests are optimal under homogeneous alternatives. Furthermore, it turns out that the stationarity tests of Hadri (2000) perform very poorly in small samples. This may be due to the fact that asymptotic values for the mean and variances of the KPSS statistics are used, whereas Levin et al. (2002) and IPS (2003) provide values for small T as well.

The relative performance of several second generation tests have been studied by Gutierrez (2003), and Gengenbach, Palm and Urbain (2004), where the cross-section dependence is assumed to follow a factor structure. The results very much depend on the underlying model. The simulations carried out by Gengenbach, Palm and Urbain (2004) show that in general, the mean CADF test has better size properties than the test of Moon and Perron (2004), which tends to be conservative in small samples. However the latter test appears to have more power against stationary idiosyncratic components. Since these tests remove the common factors, they will eventually indicate stationary time series in cases where the series are actually nonstationary due to a common stochastic trend. The results of Gengenbach et al. (2004) also suggest that the approach of Bai and Ng (2004a) is able to cope with this possibility although the power of the unit test applied to the nonstationary component is not very high.

In general, the application of factor models in the case of weak correlation does not yield valid test procedures. Alternative unit root tests that allow for weak cross section dependence are considered in Breitung and Das (2005a). They found that the GLS t-statistic may have a severe size bias if T is only slightly larger than N. In these cases Chang's (2004) bootstrap procedure is able to improve the size properties substantially. The robust OLS t-statistic performs slightly worse but outperforms the nonlinear IV test of Chang (2002). However, Monte Carlo simulations recently carried out by Baltagi, Bresson and Pirotte (2005) show that there can be considerable size distortions even in panel unit root tests that allow for weak dependence. Interestingly enough Pesaran's test, which is not designed for weak cross section

dependence, tends to be the most robust to spatial type dependence.

5 Panel Cointegration: General Considerations

The estimation of long-run relationships has been the focus of extensive research in time series econometrics. In the case of variables on a single cross section unit the existence and the nature of long-run relations are investigated using cointegration techniques developed by Engle and Granger (1987), Johansen (1991,1995) and Phillips (1991). In this literature residual-based and system approaches to cointegration are advanced. In this section we review the panel counter part of this literature. But before considering the problem of cointegration in a panel a brief overview of the cointegration literature would be helpful.

Consider the n_i time series variables $\mathbf{z}_{it} = (z_{i1t},\ z_{i2t},\ \ldots,\ z_{in_it})'$ observed on the ith cross section unit over the period $t = 1, 2, ..., T$, and suppose that for each i

$$z_{ijt} \sim I(1), \qquad j = 1, 2, \ldots, n_i.$$

Then $\mathbf{z}_{it}$ is said to form one or more cointegrating relations if there are linear combinations of z_{ijt}'s for $j = 1, 2, ..., n_i$ that are $I(0)$ i.e. if there exists an $n_i \times r_i$ matrix ($r_i \geq 1$) such that

$$\underset{r_i \times n_i}{\boldsymbol{\beta}_i'} \quad \underset{n_i \times 1}{\mathbf{z}_{it}} \quad = \quad \underset{r_i \times 1}{\boldsymbol{\xi}_{it}} \quad \sim I(0).$$

r_i denotes the number of cointegrating (or long-run) relations. The residual-based tests are appropriate when $r_i = 1$, and $\mathbf{z}_{it}$ can be partitioned such that $\mathbf{z}_{it} = (y_{it}, \mathbf{x}_{it}')'$ with no cointegration amongst the $k_i \times 1$ ($k_i = n_i - 1$) variables, $\mathbf{x}_{it}$. The system cointegration approaches are much more generally applicable and allow for $r_i > 1$ and do not require any particular partitioning of the variables in $\mathbf{z}_{it}$.[8] Another main difference between the two approaches is the way the stationary component of $\boldsymbol{\xi}_{it}$ is treated in the analysis. Most of the residual-based techniques employ non-parametric (spectral density) procedures to model the residual serial correlation in the error correction terms, $\boldsymbol{\xi}_{it}$, whilst vector autoregressions (VAR) are utilized in the development of system approaches.

In panel data models the analysis of cointegration is further complicated by heterogeneity, unbalanced panels, cross section dependence, cross unit

[8] System approaches to cointegration analysis that allow for weakly exogenous (or long-run forcing) variables has been considered in Pesaran, Shin and Smith (2000).

cointegration and the N and T asymptotics. But in cases where n_i and N are small such that $\Sigma_{i=1}^{N} n_i$ is less than 10, and T is relatively large ($T > 100$), as noted by Banerjee, Marcellino and Osbat (2004), many of these problems can be avoided by applying the system cointegration techniques to the pooled vector, $\mathbf{z}_t = (\mathbf{z}_{1t}', \mathbf{z}_{2t}', ..., \mathbf{z}_{Nt}')'$. In this setting cointegration will be defined by the relationships $\boldsymbol{\beta}'\mathbf{z}_t$ that could contain cointegration between variables from different cross section units as well as cointegration amongst the different variables specific to a particular cross section unit. This framework can also deal with residual cross section dependence since it allows for a general error covariance matrix that covers all the variables in the panel.

Despite its attractive theoretical features, the 'full' system approach to panel cointegration is not feasible even in the case of panels with moderate values of N and n_i. See Section 9 below for further details. In practice, cross section cointegration can be accommodated using common factors as in the work of Bai and Ng (2004), Pesaran (2004b), Pesaran, Schuermann and Weiner (2004, PSW) and its subsequent developments in Dees, di Mauro, Pesaran and Smith (2005, DdPS). Bai and Ng (2004a) consider the simple case where $n_i = 1$ but allow N and T to be large. But their set up can be readily generalized so that cointegration within each cross section unit as well as across the units can be considered. Following DdPS suppose that[9]

$$\mathbf{z}_{it} = \boldsymbol{\Gamma}_{id}\mathbf{d}_t + \boldsymbol{\Gamma}_{if}\mathbf{f}_t + \boldsymbol{\xi}_{it}, \tag{37}$$

for $i = 1, ..., N$; $t = 1, 2, ..., T$, and to simplify the exposition assume that $n_i = n$, where as before $\mathbf{d}_t$ is the $s \times 1$ vector of deterministics $(1, t)$ or observed common factors such as oil prices, $\mathbf{f}_t$ is a $m \times 1$ vector of unobserved common factors, $\boldsymbol{\Gamma}_{id}$ and $\boldsymbol{\Gamma}_{if}$ are $n \times s$ and $n \times m$ associated unknown coefficient matrices, $\boldsymbol{\xi}_{it}$ is an $n \times 1$ vector of error terms.

Unit root and cointegration properties of $\mathbf{z}_{it}$, $i = 1, 2, ..., N$, can be analyzed by allowing the common factors, $\mathbf{f}_t$, and/or the country-specific factors, $\boldsymbol{\xi}_{it}$, to have unit roots. To see this suppose

$$\Delta\mathbf{f}_t = \boldsymbol{\Lambda}(L)\boldsymbol{\eta}_t,\ \boldsymbol{\eta}_t \sim IID(\mathbf{0}, \mathbf{I}_m), \tag{38}$$

$$\Delta\boldsymbol{\xi}_{it} = \boldsymbol{\Psi}_i(L)\mathbf{v}_{it}, \quad \mathbf{v}_{it} \sim IID(\mathbf{0}, \mathbf{I}_n), \tag{39}$$

where L is the lag operator and

$$\boldsymbol{\Lambda}(L) = \sum_{\ell=0}^{\infty} \underset{m\times m}{\boldsymbol{\Lambda}_\ell} L^\ell, \ \ \boldsymbol{\Psi}_i(L) = \sum_{\ell=0}^{\infty} \underset{n\times n}{\boldsymbol{\Psi}_{i\ell}} L^\ell. \tag{40}$$

[9] DdPS also allow for common observed macro factors (such as oil prices), but they are not included to simplify the exposition.

The coefficient matrices, $\mathbf{\Lambda}_\ell$ and $\mathbf{\Psi}_{i\ell}$, $i = 1, 2, ..., N$, are absolute summable, so that $Var(\Delta \mathbf{f}_t)$ and $Var(\Delta \boldsymbol{\xi}_{it})$ are bounded and positive definite, and $[\mathbf{\Psi}_i(L)]^{-1}$ exists. In particular we require that

$$||\sum_{\ell=0}^{\infty} \mathbf{\Psi}_{i\ell}\mathbf{\Psi}'_{i\ell}|| \leq K < \infty, \tag{41}$$

where K is a fixed constant.

Using the familiar decomposition

$$\mathbf{\Lambda}(L) = \mathbf{\Lambda}(1) + (1-L)\mathbf{\Lambda}^*(L), \text{ and } \mathbf{\Psi}_i(L) = \mathbf{\Psi}_i(1) + (1-L)\mathbf{\Psi}_i^*(L),$$

the common stochastic trend representations of (38) and (39) can now be written as

$$\mathbf{f}_t = \mathbf{f}_0 + \mathbf{\Lambda}(1)\mathbf{s}_t + \mathbf{\Lambda}^*(L)(\boldsymbol{\eta}_t - \boldsymbol{\eta}_0),$$

and

$$\boldsymbol{\xi}_{it} = \boldsymbol{\xi}_{i0} + \mathbf{\Psi}_i(1)\mathbf{s}_{it} + \mathbf{\Psi}_i^*(L)(\mathbf{v}_{it} - \mathbf{v}_{i0}),$$

where

$$\mathbf{s}_t = \sum_{j=1}^{t} \boldsymbol{\eta}_j \text{ and } \mathbf{s}_{it} = \sum_{j=1}^{t} \mathbf{v}_{ij}.$$

Using the above results in (37) now yields

$$\begin{aligned} \mathbf{z}_{it} &= \mathbf{a}_i + \mathbf{\Gamma}_{id}\mathbf{d}_t + \mathbf{\Gamma}_{if}\mathbf{\Lambda}(1)\mathbf{s}_t + \mathbf{\Psi}_i(1)\mathbf{s}_{it} \\ &\quad + \mathbf{\Gamma}_{if}\mathbf{\Lambda}^*(L)\boldsymbol{\eta}_t + \mathbf{\Psi}_i^*(L)\mathbf{v}_{it}, \end{aligned}$$

where[10]

$$\mathbf{a}_i = \mathbf{\Gamma}_{if}[\mathbf{f}_0 - \mathbf{\Lambda}^*(L)\boldsymbol{\eta}_0] + \boldsymbol{\xi}_{i0} - \mathbf{\Psi}_i^*(L)\mathbf{v}_{i0}.$$

In this representation $\mathbf{\Lambda}(1)\mathbf{s}_t$ and $\mathbf{\Psi}_i(1)\mathbf{s}_{it}$ can be viewed as common global and individual-specific stochastic trends, respectively; whilst $\mathbf{\Lambda}^*(L)\boldsymbol{\eta}_t$ and $\mathbf{\Psi}_i^*(L)\mathbf{v}_{it}$ are the common and individual-specific stationary components. From this result it is clear that, in general, it will not be possible to simultaneously eliminate the two types of the common stochastic trends (global and individual-specific) in $\mathbf{z}_{it}$.

Specific cases of interest where it would be possible for $\mathbf{z}_{it}$ to form a cointegrating vector are when $\mathbf{\Lambda}(1) = \mathbf{0}$ or $\mathbf{\Psi}_i(1) = \mathbf{0}$. Under the former panel cointegration exists if $\mathbf{\Psi}_i(1)$ is rank deficient. The number of cointegrating relations could differ across i and is given by $r_i = n - Rank[\mathbf{\Psi}_i(1)]$. Note that

[10] In usual case where $\mathbf{d}_t$ is specified to include an intercept, 1, $\mathbf{a}_i$ can be absorbed into the deterministics.

even in this case $\mathbf{z}_{it}$ can be cross-sectionally correlated through the common stationary components, $\mathbf{\Lambda}^{*}(L)\boldsymbol{\eta}_t$. Under $\mathbf{\Psi}_i(1)=\mathbf{0}$ for all i with $\mathbf{\Lambda}(1)\neq\mathbf{0}$, we will have panel cointegration if there exists $n\times r_i$ matrices $\boldsymbol{\beta}_i$ such that $\boldsymbol{\beta}_i'\mathbf{\Gamma}_{if}\mathbf{\Lambda}(1)=\mathbf{0}$. Notice that this does not require $\mathbf{\Lambda}(1)$ to be rank deficient.

Turning to the case where $\mathbf{\Lambda}(1)$ and $\mathbf{\Psi}_i(1)$ are both non-zero, panel cointegration could still exist but must involve both $\mathbf{z}_{it}$ and $\mathbf{f}_t$. But since $\mathbf{f}_t$ is unobserved it must be replaced by a suitable estimate. The global VAR (GVAR) approach of Pesaran et al. (2004) and Dees et al. (2005) implements this idea by replacing $\mathbf{f}_t$ with the (weighted) cross section averages of $\mathbf{z}_{it}$. To see how this can be justified first differencing (37) and using (39) note that

$$[\mathbf{\Psi}_i(L)]^{-1}(1-L)(\mathbf{z}_{it}-\mathbf{\Gamma}_{id}\mathbf{d}_t-\mathbf{\Gamma}_{if}\mathbf{f}_t)=\mathbf{v}_{it}.$$

Using the approximation

$$(1-L)[\mathbf{\Psi}_i(L)]^{-1}\approx\sum_{\ell=0}^{p}\mathbf{\Phi}_{i\ell}L^{\ell}=\mathbf{\Phi}_i(L,p),$$

we obtain the following approximate VAR(p) model

$$\mathbf{\Phi}_i(L,p)(\mathbf{z}_{it}-\mathbf{\Gamma}_{id}\mathbf{d}_t-\mathbf{\Gamma}_{if}\mathbf{f}_t)\approx\mathbf{v}_{it}. \tag{42}$$

When the common factors, $\mathbf{f}_t$, are observed the model for the ith cross-section unit decouples from the rest of the units and can be estimated using the econometric techniques developed in Pesaran, Shin and Smith (2000) with $\mathbf{f}_t$ treated as weakly exogenous. But in general where the common factors are unobserved appropriate proxies for the common factors can be used. There are two possible approaches, one could either use the principal components of the observables, $\mathbf{z}_{it}$, or alternatively following Pesaran (2004b) $\mathbf{f}_t$ can be approximated in terms of $\bar{\mathbf{z}}_t=N^{-1}\Sigma_{i=1}^{N}\mathbf{z}_{it}$, the cross section averages of the observables. To see how this procedure could be justified in the present context, average the individual equations given by (37) over i to obtain

$$\bar{\mathbf{z}}_t=\bar{\mathbf{\Gamma}}_d\mathbf{d}_t+\bar{\mathbf{\Gamma}}_f\mathbf{f}_t+\bar{\boldsymbol{\xi}}_t, \tag{43}$$

where $\bar{\mathbf{\Gamma}}_d=N^{-1}\Sigma_{i=1}^{N}\mathbf{\Gamma}_{id}$, $\bar{\mathbf{\Gamma}}_f=N^{-1}\Sigma_{i=1}^{N}\mathbf{\Gamma}_{if}$, and $\bar{\boldsymbol{\xi}}_t=N^{-1}\Sigma_{i=1}^{N}\boldsymbol{\xi}_{it}$. Also, note from (39) that

$$\bar{\boldsymbol{\xi}}_t-\bar{\boldsymbol{\xi}}_{t-1}=N^{-1}\sum_{j=1}^{N}\mathbf{\Psi}_j(L)\mathbf{v}_{jt}. \tag{44}$$

But using results in Pesaran (2004b), for each t and as $N\to\infty$ we have $\bar{\boldsymbol{\xi}}_t-\bar{\boldsymbol{\xi}}_{t-1}\overset{q.m.}{\to}0$, and hence $\bar{\boldsymbol{\xi}}_t\overset{q.m.}{\to}\bar{\boldsymbol{\xi}}$, where $\bar{\boldsymbol{\xi}}$ is a time-invariant random

variable. Using this result in (43) and assuming that the $n \times m$ average factor loading coefficient matrix, $\bar{\mathbf{\Gamma}}_f$, has full column rank (with $n \geq m$) we obtain

$$\mathbf{f}_t \overset{q.m.}{\rightarrow} \left(\bar{\mathbf{\Gamma}}_f' \bar{\mathbf{\Gamma}}_f\right)^{-1} \bar{\mathbf{\Gamma}}_f \left(\bar{\mathbf{z}}_t - \bar{\mathbf{\Gamma}}_d \mathbf{d}_t - \bar{\boldsymbol{\xi}}\right),$$

which justifies using the observable vector $\{\mathbf{d}_t, \bar{\mathbf{z}}_t\}$ as proxies for the unobserved common factors.

The various contributions to the panel cointegration literature will now be reviewed in the context of the above general set up. First generation literature on panel cointegration tends to ignore the possible effects of global unobserved common factors, or attempts to account for them either by cross-section de-meaning or by using observable common effects such as oil prices or U.S. output. This literature also focusses on residual based approaches where it is often assumed that there exists at most one cointegrating relation in the individual specific models. Notable contributions to this strand of the literature include Kao (1999), Pedroni (1999, 2001, 2004), and more recently Westerlund (2005a). System approaches to panel cointegration that allow for more than one cointegrating relations include the work of Larsson, Lyhagen and Lothgren (2001), Groen and Kleibergen (2003) and Breitung (2005) who generalized the likelihood approach introduced in Pesaran, Shin and Smith (1999). Like the second generation panel unit root tests, recent contributions to the analysis of panel cointegration have also emphasized the importance of allowing for cross section dependence which, as we have noted above, could be due to the presence of common stationary or non-stationary components or both. The importance of allowing for the latter has been emphasized in Banerjee, Marcellino and Osbat (2004) through the use of Monte Carlo experiments in the case of panels where N is very small, at most 8 in their analysis. But to date a general approach that is capable of addressing all the various issues involved does not exist if N is relatively large.

We now consider in some further detail the main contributions, beginning with a brief discussion of the spurious regression problem in panels.

6 Residual-Based Approaches to Panel Cointegration

Under this approach $\mathbf{z}_{it}$ is partitioned as $\mathbf{z}_{it} = (y_{it}, \mathbf{x}_{it}')'$ and the following regressions

$$y_{it} = \boldsymbol{\delta}_i' \mathbf{d}_{it} + \mathbf{x}_{it}' \boldsymbol{\beta} + u_{it}, \; i = 1, 2, ..., N, \tag{45}$$

are considered, where as before $\boldsymbol{\delta}_i' \mathbf{d}_{it}$ represent the deterministics and the $k \times 1$ vector of regressors, $\mathbf{x}_{it}$, are assumed to be $I(1)$ and not cointegrated. How-

ever, the innovations in $\Delta\mathbf{x}_{it}$, denoted by $\boldsymbol{\varepsilon}_{it} = \Delta\mathbf{x}_{it} - E(\Delta\mathbf{x}_{it})$, are allowed to be correlated with u_{it}. Residual-based approaches to panel cointegration focus on testing for unit roots in OLS or panel estimates of u_{it}.

6.1 Spurious regression

Let $\mathbf{w}_{it} = (u_{it}, \boldsymbol{\varepsilon}'_{it})'$ and assume that the conditions for the functional central limit theorem are satisfied such that

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[\cdot T]} \mathbf{w}_{it} \xrightarrow{d} \boldsymbol{\Sigma}_i^{1/2} \mathbf{W}_i(\cdot),$$

where $\mathbf{W}_i$ is a $(k+1) \times 1$ vector of standard Brownian motions, $\xrightarrow{d}$ denotes weak convergence on $D[0,1]$ and

$$\boldsymbol{\Sigma}_i = \begin{bmatrix} \sigma_{i,u}^2 & \boldsymbol{\sigma}_{i,u\varepsilon} \\ \boldsymbol{\sigma}'_{i,u\varepsilon} & \boldsymbol{\Sigma}_{i,\varepsilon\varepsilon} \end{bmatrix}.$$

Kao (1999) showed that in the homogeneous case with $\boldsymbol{\Sigma}_i = \boldsymbol{\Sigma}$, $i = 1, \ldots, N$, and abstracting from the deterministics, the OLS estimator $\widehat{\boldsymbol{\beta}}$ converges in probability to the limit $\boldsymbol{\Sigma}_{\varepsilon\varepsilon}^{-1} \boldsymbol{\sigma}_{\varepsilon u}$, where it is assumed that $\mathbf{w}_{it}$ is i.i.d. across i. In the heterogeneous case $\boldsymbol{\Sigma}_{\varepsilon\varepsilon}$ and $\boldsymbol{\sigma}_{\varepsilon u}$ are replaced by the means $\overline{\boldsymbol{\Sigma}}_{\varepsilon\varepsilon} = N^{-1} \sum_{i=1}^N \boldsymbol{\Sigma}_{i,\varepsilon\varepsilon}$ and $\overline{\boldsymbol{\sigma}}_{\varepsilon u} = N^{-1} \sum_{i=1}^N \boldsymbol{\sigma}_{i,\varepsilon u}$, respectively (cf. Pedroni 2000). In contrast, the OLS estimator of $\boldsymbol{\beta}$ fails to converge within a pure time series framework (cf. Phillips 1987). On the other hand, if $\mathbf{x}_{it}$ and y_{it} are independent random walks, then the t-statistics for the hypothesis that one component of $\boldsymbol{\beta}$ is zero is $O_p(T^{1/2})$ and, therefore, the t-statistic has similar properties as in the time series case. As demonstrated by Entorf (1997) and Kao (1999), the tendency for spuriously indicating a relationship among y_{it} and $\mathbf{x}_{it}$ may even be stronger in panel data regressions as in the pure time series case. Therefore, it is important to test whether the errors in a panel data regression like (45) are stationary.

6.2 Tests of Panel Cointegration

As in the pure time series framework, the variables in a regression function can be tested against cointegration by applying unit roots tests of the sort suggested in the previous sections to the residuals of the estimated regression. Unfortunately, panel unit root tests cannot be applied to the residuals in (45) if $\mathbf{x}_{it}$ is (long-run) endogenous, that is, if $\boldsymbol{\sigma}_{\varepsilon u} \neq 0$. Letting $T \to \infty$ be

followed by $N \to \infty$, Kao (1999) show that the limiting distribution of the DF t-statistic applied to the residuals of a pooled OLS regression of (45) is

$$(t_\phi - \sqrt{N}\,\mu_K)/\sigma_K \;\xrightarrow{d}\; \mathcal{N}(0,1), \tag{46}$$

where the values of μ_K and σ_K depend on the kind of deterministics included in the regression, the contemporaneous covariance matrix $E(\mathbf{w}_{it}\mathbf{w}_{it}')$ and the long-run covariance matrix $\mathbf{\Sigma}_i$. Kao (1999) proposed adjusting t_ϕ by using consistent estimates of μ_K and σ_K, where he assumes that the nuisance parameters are the same for all panel units (homogenous short-run dynamics).

Pedroni (2004) suggest two different test statistics for the models with heterogeneous cointegration vectors. Let $\widehat{u}_{it} = y_{it} - \widehat{\boldsymbol{\delta}}_i'\mathbf{d}_{it} - \widehat{\boldsymbol{\beta}}_i'\mathbf{x}_{it}$ denote the OLS residual of the cointegration regression. Pedroni considers two different classes of test statistics: (i) the "panel statistic" that is equivalent to the unit root statistic against homogeneous alternatives and (ii) the "Group Mean statistic" that is analogous to the panel unit root tests against heterogeneous alternatives. The two versions of the t statistic are defined as

$$\text{panel} \qquad Z_{Pt} = \left(\tilde{\sigma}^2_{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{u}^2_{i,t-1}\right)^{-1/2}\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{u}_{i,t-1}\hat{u}_{it} - T\sum_{i=1}^{N}\hat{\lambda}_i\right)$$

$$\text{group-mean} \qquad \widetilde{Z}_{Pt} = \sum_{i=1}^{N}\left(\hat{\sigma}^2_{ie}\sum_{t=1}^{T}\hat{u}^2_{i,t-1}\right)^{-1/2}\left(\sum_{t=1}^{T}\hat{u}_{i,t-1}\hat{u}_{it} - T\hat{\lambda}_i\right)$$

where $\hat{\lambda}_i$ is a consistent estimator of the one-sided long run variance $\lambda_i = \sum_{j=1}^{\infty} E(e_{it}e_{i,t-j})$, $e_{it} = u_{it} - \delta_i u_{i,t-1}$, $\delta_i = E(u_{it}u_{i,t-1})/E(u^2_{i,t-1})$, $\hat{\sigma}^2_{ie}$ denotes the estimated variance of e_{it} and $\tilde{\sigma}^2_{NT} = N^{-1}\sum_{i=1}^{N}\hat{\sigma}^2_{ie}$. Pedroni presents values of μ_p, σ^2_p and $\tilde{\mu}_p, \tilde{\sigma}^2_p$ such that $(Z_{Pt} - \mu_p\sqrt{N})/\sigma_p$ and $(\widetilde{Z}_{Pt} - \tilde{\mu}_p\sqrt{N})/\tilde{\sigma}_p$ have standard normal limiting distributions under the null hypothesis.

Other residual-based panel cointegration tests include the recent contribution of Westerlund (2005a) that are based on variance ratio statistics and do not require corrections for the residual serial correlations.

7 Tests for Multiple Cointegration

It is also possible to adapt Johansen's (1995) multivariate test based on a VAR representation of the variables. Let $\Lambda_i(r)$ denote the cross-section specific likelihood-ratio ("trace") statistic of the hypothesis that there are (at most) r stationary linear combinations in the cointegrated VAR system

given by $\mathbf{z}_{it} = (y_{it}, \mathbf{x}'_{it})'$. Following the unit root test proposed in IPS (2003), Larsson et al. (2001) suggested the standardized LR-bar statistic

$$\widetilde{\Lambda}(r) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{\Lambda_i(r) - E[\lambda_i(r)]}{\sqrt{Var[\lambda_i(r)]}},$$

to test the null hypothesis that $r = 0$ against the alternative that at most $r = r^0 \geq 1$. Using a sequential limit theory it can be shown that $\widetilde{\Lambda}(r)$ is asymptotically standard normally distributed. Asymptotic values of $E[\lambda_i(r)]$ and $Var[\lambda_i(r)]$ are tabulated in Larsson et al. (2001) for the model without deterministic terms and Breitung (2005) for models with a constant and a linear time trend. Unlike the residual-based tests, the LR-bar test allows for the possibility of multiple cointegration relations in the panel.

It is also possible to test the null hypothesis that the errors of the cointegration regression are stationary. That is, under the null hypothesis it is assumed that $y_{it}, \mathbf{x}_{it}$ are cointegrated with cointegration rank $r = 1$. McCoskey and Kao (1998) suggest a panel version of Shin's (1994) cointegration test based on the residuals of a fully modified OLS regression. Westerlund (2005b) suggests a related test procedure based on the CUSUM statistic.

8 Estimation of Cointegrating Relations in Panels

8.1 Single Equation Estimators

First, we consider a single-equation framework where it is assumed that y_{it} and the $k \times 1$ vector of regressors $\mathbf{x}_{it}$ are $I(1)$ with at most one cointegrating relations amongst them, namely that there exists a linear relationship of the form (45) such that the error u_{it} is stationary. As before it is assumed that $\mathbf{z}_{it} = (y_{it}, \mathbf{x}'_{it})'$ is i.i.d. across i, and the regressors, $\mathbf{x}_{it}$, are not cointegrated. We do not explicitly consider deterministic terms like individual specific constants or trends as the asymptotic theory applies to mean- or trend-adjusted variables as well.

It is assumed that the vector of coefficients, $\boldsymbol{\beta}$, is the same for all cross-section units, that is, a homogeneous cointegration relationship is assumed. Alternatively, it may be assumed that the cointegration parameters are cross section specific (heterogenous cointegration).

By applying a sequential limit theory it can be shown that the OLS estimator of $\boldsymbol{\beta}$ is $T\sqrt{N}$ consistent and, therefore, the time series dimension

is more informative on the long-run coefficients than the cross-section dimension. Furthermore, is important to notice that – as in the time series framework – the OLS estimator is consistent but inefficient in the model with endogenous regressors.

Pedroni (1995) and Phillips and Moon (1999, p. 1085) proposed a "fully-modified OLS" (FM-OLS) approach to obtain an asymptotically efficient estimator for homogenous cointegration vectors. This estimator adjusts for the effects of endogenous regressors and short-run dynamics of the errors (cf. Phillips and Hansen 1990). To correct for the effect of (long-run) endogeneity of the regressors, the dependent variable is adjusted for the part of the error that is correlated with the regressor

$$y_{it}^{+} = y_{it} - \boldsymbol{\sigma}'_{i,\varepsilon u}\boldsymbol{\Sigma}_{i,\varepsilon\varepsilon}^{-1}\Delta\mathbf{x}_{it}. \tag{47}$$

A second correction is necessary when computing the OLS estimator

$$\widehat{\boldsymbol{\beta}}_{FM} = \left[\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbf{x}_{it}\mathbf{x}'_{it}\right]^{-1}\left[\sum_{i=1}^{N}\sum_{t=1}^{T}(\mathbf{x}_{it}y_{it}^{+} - \boldsymbol{\lambda}_{i,\boldsymbol{\varepsilon}u})\right], \tag{48}$$

where

$$\boldsymbol{\lambda}_{i,\varepsilon u} = E\left(\sum_{j=0}^{\infty}\boldsymbol{\varepsilon}_{i,t-j}u_{it}\right).$$

The nuisance parameters can be estimated consistently using familiar non-parametric procedures.

An alternative approach is the "Dynamic OLS" (DOLS) estimator suggested by Saikkonen (1991). This estimator is based on the error decomposition

$$u_{it} = \sum_{k=-\infty}^{\infty}\boldsymbol{\gamma}'_k\Delta\mathbf{x}_{i,t+k} + v_{it}\ , \tag{49}$$

where v_{it} is orthogonal to all leads and lags of $\Delta\mathbf{x}_{it}$. Inserting (49) in the regression (45) yields

$$y_{it} = \boldsymbol{\beta}'\mathbf{x}_{it} + \sum_{k=-\infty}^{\infty}\boldsymbol{\gamma}'_k\Delta\mathbf{x}_{i,t+k} + v_{it}\ . \tag{50}$$

In practice the infinite sums are truncated at some small numbers of leads and lags (cf. Kao and Chiang 2000, Mark and Sul 2003). Westerlund (2005c) considers data dependent choices of the truncation lags. Kao and Chiang (2000) show that in the homogeneous case with $\boldsymbol{\Sigma}_i = \boldsymbol{\Sigma}$ and individual

specific intercepts the limiting distribution of the DOLS estimator $\widehat{\boldsymbol{\beta}}_{DOLS}$ is given by

$$T\sqrt{N}(\widehat{\boldsymbol{\beta}}_{DOLS} - \boldsymbol{\beta}) \xrightarrow{d} \mathcal{N}(0, 6\,\sigma^2_{u|\varepsilon}\boldsymbol{\Sigma}^{-1}_{\varepsilon\varepsilon}),$$

where

$$\sigma^2_{u|\varepsilon} = \sigma^2_u - \sigma'_{\varepsilon u}\boldsymbol{\Sigma}^{-1}_{\varepsilon\varepsilon}\sigma_{\varepsilon u}.$$

Furthermore, the FM-OLS estimator possesses the same asymptotic distribution as the DOLS estimator. In the heterogeneous case $\boldsymbol{\Sigma}_{\varepsilon\varepsilon}$ and $\sigma^2_{u|\varepsilon}$ are replaced by $\overline{\boldsymbol{\Sigma}}_{\varepsilon\varepsilon} = N^{-1}\sum_{i=1}^N \boldsymbol{\Sigma}_{i,\varepsilon\varepsilon}$ and $\overline{\sigma}^2_{u|\varepsilon} = N^{-1}\sum_{i=1}^N \sigma^2_{i,u|\varepsilon}$, respectively (cf. Phillips and Moon 1999). Again, the matrix $\boldsymbol{\Sigma}_i$ can be estimated consistently (for $T \to \infty$) by using a nonparametric approach.

In many applications the number of time periods is smaller than 20 and, therefore, the kernel based estimators of the nuisance parameters may perform poorly in such small samples. Pesaran, Shin and Smith (1999) adapted a parametric model to estimate the cointegration vector based on the error correction format

$$\Delta y_{it} = \phi_i y_{i,t-1} + \boldsymbol{\gamma}'_i \mathbf{x}_{it} + v_{it}\ , \tag{51}$$

where for simplicity of exposition we have abstracted from deterministics and lagged changes in y_{it} and $\mathbf{x}_{it}$.[11]. It is assumed that the long-run parameters are identical across the cross section units, i.e., $\boldsymbol{\beta}_i = -\boldsymbol{\gamma}_i/\phi_i = \boldsymbol{\beta}$ for $i = 1, \ldots, N$. Economic theory often predicts the same cointegration relation(s) across the cross section units, although is often silent on the magnitude of short-run dynamics, ϕ_i, across i. For example, the long-run relationships predicted by the PPP, the uncovered interest parity, or the Fisher equation are the same across countries, although the speed of convergence to these long-run relations could differ markedly over countries due to differences in economic and political institutions.[12] For further discussions see, for example, Pesaran (1997).

Letting $\zeta_{it}(\boldsymbol{\beta}) = y_{i,t-1} - \boldsymbol{\beta}'\mathbf{x}_{it}$, the model is rewritten as

$$\Delta y_{it} = \phi_i \zeta_{it}(\boldsymbol{\beta}) + v_{it}\ . \tag{52}$$

Pesaran et al. (1999) have suggested an ML estimation method based on the concentrated likelihood function

$$L_c(\boldsymbol{\beta}) = c - \sum_{i=1}^N \frac{T}{2}\log|\widehat{\sigma}^2_{i,v}(\boldsymbol{\beta})|\ , \tag{53}$$

[11] Since there are no restrictions on the additional variables they can be concentrated out from the likelihood function by replacing Δy_{it}, $y_{i,t-1}$ and $\mathbf{x}_{it}$ by residuals obtained from regressions on the deterministic terms and lagged differences of the variables, $\Delta\mathbf{z}_{i,t-1}, \Delta\mathbf{z}_{i,t-2}, \ldots$

[12] The problem of testing the slope homogeneity hypothesis in panels is reviewed in Hsiao and Pesaran (2005).

where c is a constant, and

$$\begin{aligned}
\widehat{\sigma}_{i,v}^2(\boldsymbol{\beta}) &= \frac{1}{T}\sum_{t=1}^{T}\widetilde{v}_{it}(\boldsymbol{\beta})^2, \\
\widetilde{v}_{it}(\boldsymbol{\beta}) &= \Delta y_{it} - \left(\frac{\sum_{t=1}^{T}\Delta y_{it}\zeta_{it}(\boldsymbol{\beta})}{\sum_{t=1}^{T}\zeta_{it}(\boldsymbol{\beta})^2}\right)\zeta_{it}(\boldsymbol{\beta}).
\end{aligned}$$

Pesaran et al. (1999) suggested a Gauss-Newton algorithm to maximize (53). The means of the error correction coefficients are estimated by the simple average of the individual coefficients (or the ML estimates) of ϕ_i, $i = 1, \ldots, N$. This estimator is called the *pooled mean group estimator.*

8.2 System Estimators

The single equation estimators have several drawbacks that can be avoided by using a system approach. First, these estimators assume that all regressors are I(1) and not cointegrated. If there are more than one cointegration relationships, then the matrix $\boldsymbol{\Sigma}_{\varepsilon\varepsilon}$ is singular and the asymptotic results are no longer valid. Second, the cointegration relationship has to be normalized such that the variable y_{it} enters with unit coefficient. As has been argued by Boswijk (1995), this normalization is problematical if the original coefficient of the variable y_{it} tends to zero.

In the case of short panels with T fixed and N large, Binder, Hsiao and Pesaran (2005) consider estimation and inference in panel vector autoregressions (PVARs) with homogeneous slopes where (i) the individual effects are either random or fixed, (ii) the time-series properties of the model variables are unknown *a priori* and may feature unit roots and cointegrating relations. Generalized Method of Moments (GMM) and Quasi Maximum Likelihood (QML) estimators are obtained and compared in terms of their asymptotic and finite sample properties. It is shown that the asymptotic variances of the GMM estimators that are based on levels as well as first-differences of the model variables depend on the variance of the individual effects; whereas by construction the fixed effects QML estimator is not subject to this problem. Monte Carlo evidence is provided showing that the fixed effects QML estimator tends to outperform the various GMM estimators in finite sample under both normal and non-normal errors. The paper also shows how the fixed effects QML estimator can be successfully used for unit root and cointegration tests in short panels.

In the case of panels with large N and T, Larsson and Lyhagen (1999), Groen and Kleibergen (2003) and Breitung (2005) consider the vector error correction model (VECM) for the $k+1$ dimensional vector $\mathbf{z}_{it} = (y_{it}, \mathbf{x}_{it}')'$ given by

$$\Delta \mathbf{z}_{it} = \boldsymbol{\alpha}_i \boldsymbol{\beta}_i' \mathbf{z}_{i,t-1} + \mathbf{w}_{it}, \tag{54}$$

where $\mathbf{w}_{it} = (u_{it}, \boldsymbol{\varepsilon}_{it}')'$ and once again we leave out deterministic terms and lagged differences. To be consistent with the approaches considered above, we confine ourselves to the case of homogenous cointegration, that is, we let $\boldsymbol{\beta}_i = \boldsymbol{\beta}$ for $i = 1, \ldots, N$. Larsson and Lyhagen (1999) propose a ML estimator, whereas the estimator of Groen and Kleibergen (2003) is based on a nonlinear GMM approach.

It is well known that the ML estimator of the cointegration parameters for a single series may behave poorly in small samples. Phillips (1994) has shown that the finite sample moments of the estimator do not exist. Using Monte Carlo simulations Hansen, Kim and Mittnik (1998) and Brüggemann and Lütkepohl (2004) found that the ML estimator may produce implausible estimates far away from the true parameter values. Furthermore the asymptotic χ^2 distribution of the likelihood ratio test for restrictions on the cointegration parameters may be poor guide for small sample inference (e.g. Gredenhoff and Jacobson, 2001).

To overcome these problems, Breitung (2005) proposed a computationally convenient two-step estimator, which is adopted from Ahn and Reinsel (1990). This estimator is based on the fact that the Fisher information is block-diagonal with respect to the short and long-run parameters. Accordingly, an asymptotically efficient estimator can be constructed by estimating the short- and long-run parameters in separate steps. Suppose that the $n \times r$ matrix of cointegrating vectors is 'normalized' as $\boldsymbol{\beta} = (\mathbf{I}_r, \mathbf{B})'$, where $\mathbf{I}_r$ is the identity matrix of order r and $\mathbf{B}$ is the $(n-r) \times r$ matrix of unknown coefficients.[13] Then $\boldsymbol{\beta}$ is exactly identified and the Gaussian ML estimator of $\mathbf{B}$ is equivalent to the OLS estimator of $\mathbf{B}$ in

$$\mathbf{z}_{it}^* = \mathbf{B}\mathbf{z}_{i,t-1}^{(2)} + \mathbf{v}_{it} \ ,$$

where $\mathbf{z}_{it}^{(2)}$ is the $r \times 1$ vector defined by $\mathbf{z}_{it} = [\mathbf{z}_{it}^{(1)'}, \mathbf{z}_{it}^{(2)'}]'$, and

$$\mathbf{z}_{it}^* = (\boldsymbol{\alpha}_i' \boldsymbol{\Sigma}_i^{-1} \boldsymbol{\alpha}_i)^{-1} \boldsymbol{\alpha}_i' \boldsymbol{\Sigma}_i^{-1} \Delta \mathbf{z}_{it} - \mathbf{z}_{i,t-1}^{(1)}.$$

[13] The analysis can be readily modified to take account of other types of exact identifying restrictions on $\boldsymbol{\beta}$ that might be more appropriate from the view-point of long-run economic theory. See Pesaran and Shin (2002) for a general discussion of identification and testing of cointegrating relations in the context of a single cross section units.

The matrices $\boldsymbol{\alpha}_i$ and $\boldsymbol{\Sigma}_i$ can be replaced by $\sqrt{T}$-consistent estimates without affecting the limiting distribution. Accordingly, these matrices can be estimated for each panel unit separately, e.g., by using Johansen's (1991) ML estimator. To obtain the same normalization as in (**??**) the estimator for $\boldsymbol{\alpha}_i$ is multiplied with the $r \times r$ upper block of the ML estimator of $\boldsymbol{\beta}$.

Breitung (2005) showed that the limiting distribution of the OLS estimator of $\mathbf{B}$ is asymptotically normal. Therefore, tests of restrictions on the cointegration parameters have the standard limiting distributions (i.e. a χ^2 distribution for the usual Wald tests).

Some Monte Carlo experiments were performed by Breitung (2005) to compare the small sample properties of the two-step estimator with the FM-OLS and DOLS estimators. The results suggest that the latter two tests may be severely biased in small samples, whereas the bias of the two-step estimator is relatively small. Furthermore, the standard errors (and hence the size properties of the t-statistics) of the two-step procedure are more reliable than the ones of the semi-parametric estimation procedures.

9 Cross-section Dependence and the Global VAR

As pointed out earlier an important limitation of the econometric approaches discussed so far is that they assume that all cross-section units are independent. In many applications based on multi-country data sets this assumption is clearly unrealistic. To accommodate cross-dependence among panel units Mark, Ogaki and Sul (2004) and Moon and Perron (2005) proposed a Dynamic Seemingly Unrelated Regression (DSUR) estimator. Their approach is based on a GLS estimator of the dynamic representation (50) when there exists a single cointegrating relation between y_{it} and $\mathbf{x}_{it}$, and does not allow for the possibility of cross unit cointegration. Let $\mathbf{h}_{it}(p) = \left(\Delta\mathbf{x}'_{i,t-p}, \ldots, \Delta\mathbf{x}'_{i,t+p}\right)'$ and $\mathbf{h}_{pt} = \left(\mathbf{h}'_{p1t}, \ldots, \mathbf{h}'_{pNt}\right)'$. To correct for endogeneity of the regressors, first y_{it} and $\mathbf{x}_{it}$ are regressed on $\mathbf{h}_{pt}$. Let $\tilde{y}_{it}$ and $\tilde{\mathbf{x}}_{it}$ denote the resulting regression residuals. Furthermore, define $\tilde{\mathbf{y}}_t = (\tilde{y}_{1t}, \ldots, \tilde{y}_{Nt})'$ and $\widetilde{\mathbf{X}}_t = (\tilde{\mathbf{x}}_{1t}, \ldots, \tilde{\mathbf{x}}_{Nt})'$. The DSUR estimator of the (homogeneous) cointegration vector is

$$\widehat{\boldsymbol{\beta}}_{dsur} = \left(\sum_{t=p+1}^{T-p} \widetilde{\mathbf{X}}'_t \boldsymbol{\Sigma}_{uu}^{-1} \widetilde{\mathbf{X}}_t\right) \sum_{t=p+1}^{T-p} \widetilde{\mathbf{X}}'_t \boldsymbol{\Sigma}_{uu}^{-1} \tilde{\mathbf{y}}_t \tag{55}$$

where $\mathbf{\Sigma}_{uu}$ denotes the long-run covariance matrix of $\mathbf{u}_t = (u_{1t}, \ldots, u_{Nt})'$, namely

$$\mathbf{\Sigma}_{uu} = \lim_{T\to\infty} \frac{1}{T} E\left[\left(\sum_{t=1}^{T} \mathbf{u}_t\right)\left(\sum_{t=1}^{T} \mathbf{u}_t'\right)\right],$$

for a fixed N. This matrix is estimated by using an autoregressive representation of $\mathbf{u}_t$. See also (49).

An alternative approach is suggested by Breitung (2005), where a SUR procedure is applied in the second step of the two-step estimator. Bai and Kao (2004) consider a factor structure of the errors as in (31). They proposed a two-step FMOLS procedure, where in the common factors are estimated from the residuals of an initial FMOLS estimation. The endogeneity correction at the second stage is

$$\hat{y}_{it}^{+} = y_{it} - (\boldsymbol{\gamma}_i' \widehat{\mathbf{\Sigma}}_{i,f\xi} + \widehat{\mathbf{\Sigma}}_{i,\varepsilon\xi}) \widehat{\mathbf{\Sigma}}_{i,\varepsilon}^{-1} \Delta \mathbf{x}_{it} \ ,$$

where $\widehat{\mathbf{\Sigma}}_{i,f\xi}$ denotes the estimated long-run covariance matrix for $\mathbf{f}_t, \xi_{it}$ and $\widehat{\mathbf{\Sigma}}_{i,\varepsilon\xi}$ is the estimated long-run covariance matrix of the idiosyncratic errors ξ_{it} and the innovations $\boldsymbol{\varepsilon}_{it}$. The two-step FMOLS estimator results as

$$\widehat{\boldsymbol{\beta}}_{2FM} = \left(\sum_{i=1}^{N}\sum_{t=1}^{T} \mathbf{x}_{it}\mathbf{x}_{it}'\right)^{-1} \left[\sum_{i=1}^{N}\left(\sum_{t=1}^{T} \mathbf{x}_{it}\hat{y}_{it}^{+} - T(\widehat{\mathbf{\Lambda}}_{i,\varepsilon f}\hat{\boldsymbol{\gamma}}_i + \widehat{\mathbf{\Lambda}}_{i,\varepsilon\xi})\right)\right]$$

where $\widehat{\mathbf{\Lambda}}_{i,\varepsilon f}$ and $\widehat{\mathbf{\Lambda}}_{i,\varepsilon\xi}$ are the respective one-sided long-run covariance matrices.

A common feature of these approaches is that cross-section dependence can be represented by a contemporaneous correlation of the errors, and do not allow for the possibility of cross unit cointegration. In many applications it is more realistic to allow for some form of dynamic cross-section dependence. A general model to accommodate cross-section cointegration and dynamic links between panel units is the panel VECM model considered by Groen and Kleibergen (2003) and Larsson and Lyhagen (1999). As in Section 5, let $\mathbf{z}_{it}$ denote a n-dimensional vector of times series on the ith cross section unit. Consider the $nN \times 1$ vector $\mathbf{z}_t = (\mathbf{z}_{1t}', \ldots, \mathbf{z}_{Nt}')'$ of all available time series in the panel data set. The VECM representation of this time series vector is

$$\Delta \mathbf{z}_t = \Pi \mathbf{z}_{t-1} + \Gamma_1 \Delta \mathbf{z}_{t-1} + \cdots + \Gamma_p \Delta \mathbf{z}_{t-p} + \mathbf{u}_t \ . \tag{56}$$

For cointegrated systems $rank(\Pi) < nN$. It is obvious that such systems typically involve a large number of parameters as the number of parameters increases with N^2. Therefore, to obtain reliable estimates of the parameters

T must be considerably larger than N. In many macroeconomic applications, however, the number of time periods is roughly as large as the number of cross-section units. Therefore, a simple structure must be imposed on the matrices $\mathbf{\Pi}, \mathbf{\Gamma}_1, \ldots, \mathbf{\Gamma}_p$ that yields a reasonable approximation to the underlying dynamic system.

The Global VAR (GVAR) introduced by Pesaran, Schuermann and Weiner (2004) and further developed in Dees et al. (2005) can be seen as a theory guided reduction of the general dynamic model given by (37), (38) and (39). In the context of this set up the individual cross section models in the GVAR can be approximated by the $VARZ^*(p_i, p_i)$ in $\mathbf{z}_{it}$ and $\mathbf{z}_{it}^*$:[14]

$$\mathbf{\Phi}_i (L, p_i) (\mathbf{z}_{it} - \mathbf{\Gamma}_{id}\mathbf{d}_t - \mathbf{\Gamma}_{if}\mathbf{z}_{it}^*) = \mathbf{v}_{it}. \tag{57}$$

for $i = 1, 2, ..., N$, where $\mathbf{d}_t$ are the observed common effects (such as intercepts, time trends or oil prices), and $\mathbf{z}_{it}^*$ is defined by

$$\mathbf{z}_{it}^* = \sum_{j=1}^{N} w_{ij}\mathbf{z}_{jt}.$$

The weights, w_{ij}, $j = 1, 2, ..., N$ must satisfy the following conditions

$$w_{ii} = 0, \ \sum_{j=1}^{N} w_{ij} = 1, \text{ and } \sum_{j=1}^{N} w_{ij}^2 \to 0, \text{ as } N \to \infty,$$

and could be time varying. Typical examples of such weights are $w_{ij} = 1/(N-1)$, for $i \neq j$, trade weights or other measures of economic distance between the cross section units. The estimation of (57) can proceed by treating the cross section averages as weakly exogenous $I(1)$ variables using standard time series cointegration techniques developed, for example, in Pesaran , Shin and Smith (2000). The assumption that $\mathbf{z}_{it}^*$ are weakly exogenous $I(1)$, or long-run forcing, for $\mathbf{z}_{it}$, can be tested. For further details see Pesaran, Schuermann and Weiner (2004). It turns out that this is a reasonable assumption for all countries except for the U.S. where most of the variables should be treated as endogenous.

10 Concluding Remarks

As this review shows the literature on panel unit roots and cointegration has been expanding very rapidly; in part responding to the complex nature of

[14] VARZ* represents a VAR model augmented with $\mathbf{z}_{it}^*$ as weakly exogenous variables.

the interactions and dependencies that generally exist over time and across the individual units in the panel. Observations on firms, industries, regions and countries tend to be cross correlated as well as serially dependent. The problem of cross section dependence is particularly difficult to deal with since it could arise for a variety of reasons; spatial spill over effects, common unobserved shocks, social interactions or a combination of these factors. Parameter heterogeneity and deterministics also pose additional difficulties and how they are treated under the null and the alternative hypothesis can affect the outcome of the empirical analysis.

Initially, the panel unit root and cointegration tests were developed assuming that the errors of the individual equations are cross sectionally independent. These, referred to as the first generation tests, continue to form an important part of the literature, providing a theoretical basis for the more recent (second generation) developments that attempt to take account of the residual cross section dependence in panels in the case of panels where the time dimension and the cross section dimension are both relatively large. In the analysis of cointegration the hypothesis testing and estimation problems are also further complicated by the possibility of cross section cointegration. These and other issues are currently the subject of extensive research.

References

Ahn, S.K. and G.C. Reinsel (1990), Estimation for Partially Nonstationary Multivariate Autoregressive Models, Journal of the American Statistical Association, 85, 813–823.

Andrews, D.W.K. (1998), Hypothesis Testing with a Restricted Parameter Space, Journal of Econometrics, 84, 155-199.

Arellano, M. (2003), Panel Data Econometrics, Oxford: Oxford University Press.

Bai, J. (2004), Estimating Cross-section Common Stochastic Trends in Nonstationary Panel Data, Journal of Econometrics, 122, 137–183.

Bai, J. and C. Kao (2004), On the Estimation and Inference of a Panel Cointegration Model with Cross-Sectional Dependence, New York University, Unpublished.

Bai, J. and S. Ng (2002), Determining the Number of Factors in Approximate Factor Models, Econometrica, 70, 191–221.

Bai, J. and S. Ng (2004a), A Panic Attack on Unit Roots and Cointegration, Econometrica, 72, 1127–1177.

Bai, J. and S. Ng (2004b), A New Look at Panel Testing of Stationarity and the PPP Hypothesis, New York University, unpublished.

Baltagi, B.H. and C. Kao (2000), Nonstationary Panels, Cointegration in Panels and Dynamic Panels, A Survey, in: B. Baltagi (ed.), Nonstationary Panels, Panel Cointegration, and Dynamic Panels, Advances in Econometrics, Vol. 15, Amsterdam: JAI Press, 7–52.

Baltagi, B.H., G. Bresson, and A. Pirotte, (2005), Panel Unit Root Tests and Spatial Dependence, Department of Economics, Texas A&M University, unpublished.

Banerjee, A. (1999), Panel Data Unit Roots and Cointegration: An Overview, Oxford Bulletin of Economics and Statistics, 61, 607–629.

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

Banerjee, A., M. Marcellino and C. Osbat (2005), Testing for PPP: Should we use Panel Methods?, forthcoming in Empirical Economics.

Binder, M., Hsiao, C., and M. H. Pesaran (2005), Estimation and Inference in Short Panel Vector Autoregressions with Unit Roots and Cointegration, Econometric Theory 21, 795-837.

Bond, S., C. Nauges and F. Windmeijer (2002), Unit Roots and Identification in Autoregressive Panel Data Models: A Comparison of Alternative Tests, unpublished.

Boswijk, H. P. (1995). Efficient Inference on Cointegration Parameters in Structural Error Correction Models, Journal of Econometrics, 69, 133-158.

Breitung, J. (2000), The Local Power of Some Unit Root Tests for Panel Data, in B. Baltagi (ed.), Nonstationary Panels, Panel Cointegration, and Dynamic Panels, Advances in Econometrics, Vol. 15, JAI, Amsterdam, 161–178.

Breitung, J. (2002), Nonparametric Tests for Unit Roots and Cointegration, Journal of Econometrics, 108, 343–363.

Breitung, J. (2005), A Parametric Approach to the Estimation of Cointegration Vectors in Panel Data, Econometric Reviews, 151-174.

Breitung J. and B. Candelon (2005), Purchasing Power Parity During Currency Crises: A panel unit root test under structural breaks, World Economic Review, 141, 124–140.

Breitung, J. and W. Meyer (1994), Testing for Unit Roots in Panel Data: Are Wages on Different Bargaining Levels Cointegrated?, *Applied Economics*, 26, 353–361.

Breitung, J. and S. Das (2005a), Panel Unit Root Tests Under Cross Sectional Dependence, forthcoming in: Statistica Neerlandica.

Breitung, J. and S. Das (2005b), Testing for Unit Roots in Panels With a Factor Structure, University of Bonn, unpublished.

Brüggemann, R. and H. Lütkepohl (2004), Practical Problems with Reduced Rank ML Estimators for Cointegration Parameters and a Simple Alternative, EUI Florence, unpublished.

Carrion-i-Sevestre, J.L., T. Del Barrio and E. Lopez-Bazo (2004), Breaking the Panels. An application to the GDP per capita, University of Barcelona, unpublished.

Chang, Y. (2002), Nonlinear IV Unit Root Tests in Panels with Cross-Sectional Dependency, Journal of Econometrics, 110, 261–292.

Chang, Y. (2004), Bootstrap Unit root Tests in Panels with Cross-sectional Dependency, Journal of Econometrics, 120, 263-293.

Chang, Y., J.Y. Park and P.C.B. Phillips (2001), Nonlinear Econometric Models with Cointegrated and Deterministically Trending Regressors, Econometrics Journal, 4, 1–36.

Chang, Y. and W. Song (2005), Unit Root Tests for Panels in the Presence of Short-run and Long-run Dependencies: Nonlinear IV Approach with Fixed N and Large T, Rice University, *unpublished*.

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

Choi, I. (2004), Nonstationary Panels, forthcoming in: Palgrave Handbooks of Econometrics, Vol. 1.

Choi, I. and T.K. Chue (2004), Subsampling Hypothesis Tests for Nonstationary Panels with Applications to the PPP Hypothesis, Hong Kong University of Science and Technology, unpublished.

Coakley, J. and A.M. Fuertes (1997), New Panel Unit Root Tests of PPP, Economics Letters, 57, 17–22.

Hsiao, C., and M. H. Pesaran (2005), Random Coefficient Panel Data Models, in L. Matyas and P. Sevestre, (eds.), The Econometrics of Panel Data, Kluwer Academic Publishers (forthcoming).

Coakley, J., N. Kellard and S. Smaith (2005), The PPP debate: Price Matters!, Economic Letters, 88, 209–213.

Dees, S., F. di Mauro, M. H. Pesaran and L. V. Smith (2005), Exploring the International Linkages of the Euro Area: A Global VAR Analysis , CESifo Working Paper No. 1425. Presented at CESifo Area Conference on Macro, Money and International Finance, February 2005.

Dickey, D.A. and W.A. Fuller (1979), Distribution of the Estimates for Autoregressive Time Series With a Unit Root, Journal of the American Statistical Association 74, 427–431.

Elliott, G., T. Rothenberg, and J. Stock (1996), Efficient Tests for an Autoregressive Unit Root, Econometrica, 64, 813–836.

Engle, R.F., and C.W.J. Granger (1987), Co-integration and Error Correction: Representation, Estimation, and Testing, Econometrica, 55, 251–276.

Entorf, H. (1997), Random Walks With Drifts: Nonsense Regression and Spurious Fixed-effects Estimation, Journal of Econometrics, 80, 287-296.

Fisher, R.A. (1932), Statistical Methods for Research Workers, Oliver and Bond, Edinburgh, 4th ed.

Flores, R., P. Jorion, P.Y. Preumont and A. Szarfarz (1999), Multivariate Unit Root Tests of the PPP Hypothesis, Journal of Empirical Finance, 6, 335–353.

Fuller, W.A., (1996), Introduction to Statistical Time Series, Second Edition, Wiley.

Gengenbach, C. F.C. Palm and J.-P. Urbain (2004), Panel Unit Root Tests in the Presence of Cross-Sectional Dependencies: Comparison and Implications for Modelling, Universiteit Maastricht, unpublished.

Gredenhoff, M. and T. Jacobson (2001), Bootstrap Testing Linear Restrictions on Cointegrating Vectors, Journal of Business and Economic Statistics, 19, 63–72.

Groen, J.J.J. and F. Kleibergen (2003), Likelihood-Based Cointegration Analysis in Panels of Vector Error-Correction Models, Journal of Business and Economic Statistics 21: 295–318.

Gutierrez, L. (2003), Panel Unit Roots Tests for Cross-Sectionally Correlated Panels: A Monte Carlo Comparison, *Economics Letters*, 80, 105-111.

Hadri, K. (2000), Testing for Stationarity in Heterogeneous Panel Data, Econometrics Journal, 3, 148–161.

Hadri, K. and R. Larsson (2005), Testing for Stationarity in Heterogeneous panel data where the time dimension is Fixed, Econometrics Journal, 8, 55–69.

Hansen, G., J.R. Kim and S. Mittnik (1998), Testing Cointegrating Coefficients in Vector Autoregressive Error Correction Models, Economics Letters, 58, 1–5.

Harris, D., S. Leybourne and B. McCabe (2004), Panel Stationarity Tests for Purchasing Power Parity with Cross-sectional Dependence, University of Nottingham, unpublished.

Harris, R.D.F. and E. Tzavalis (1999), Inference for Unit Roots in Dynamic Panels where the Time Dimension is Fixed, Journal of Econometrics, 91, 201-226.

Harvey, A. and D. Bates (2003), Multivariate Unit Root Tests, Stability and Convergence, University of Cambridge, DAE Working Paper No. 301, University of Cambridge, England.

Hassler, U., M. Demetrescu and A. Tarcolea (2004), Asymptotically Normal Unit Root Testing: Extensions to Panel Data and Structural Breaks, Goethe University of Frankfurt, unpublished.

Hlouskova, J. and M. Wagner (2005), The Performance of Panel Unit Root and Stationarity Tests: Results from a Large Scale Simulation Study, University of Bern, unpublished.

Hurlin, C. and V. Mignon, (2004), Second Generation Panel Unit Root Tests, THEMA-CNRS, University of Paris X, memo.

Im, K.S., J. Lee and M. Tieslau (2005), Panel LM Unit Root Tests with Level Shifts, Oxford Bulletin of Economics and Statistics, 63, 393-419.

Im, K.S., and M. H. Pesaran (2003), On the Panel Unit Root Tests Using Nonlinear Instrumental Variables, Cambridge Working Papers in Economics, number 0347, University of Cambridge.

Im, K.S., M.H. Pesaran, and Y. Shin (1995), Testing for Unit Roots in Heterogenous Panels, DAE Working Papers Amalgamated Series No. 9526, University of Cambridge.

Im, K.S., M.H. Pesaran, and Y. Shin (2003), Testing for Unit Roots in Heterogenous Panels, Journal of Econometrics, 115, 53–74.

Johansen, S. (1991), Estimation and Hypothesis Testing of Cointegrating Vectors in Gaussian Vector Autoregressive Models, Econometrica, 59, 1551–1580.

Johansen, S. (1995), Likelihood-based Inference in Cointegrated Vector Autoregressive Models, Oxford: Oxford University Press.

Jönsson, K. (2005), Cross-Sectional Dependency and Size Distortion in a Small-sample Homogeneous Panel-data Unit Root Test, Oxford Bulletin of Economics and Statistics, 63, 369-392.

Kao, C. (1999), Spurious Regression and Residual-based Tests for Cointegration in Panel Data, Journal of Econometrics, 90, 1-44.

Kao, C. and M.-H. Chiang (2000), On the Estimation and Inference of a Cointegrated Regression in Panel Data, in: Baltagi B. (ed.), Nonstationary Panels, Panel Cointegration, and Dynamic Panels, Advances in Econometrics, Vol. 15, Amsterdam: JAI Press, pp. 161–178.

Kapetanios, G. (2005), Dynamic Factor Extraction of Cross-Sectional Dependence in Panel Unit Root Tests, revised version of Working Paper Number 509, Queen Mary College, University of London.

Kiefer, N. and T. Vogelsang (2002), Heteroskedasticity-Autocorrelation Robust Standard Errors Using the Bartlett Kernel Without Truncation, Econometrica, 70, 2093–2095.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. and Shin, Y. (1992) Testing the Null Hypothesis of Stationary Against the Alternative of a Unit Root: How Sure are we that Economic Time Series Have a Unit Root? Journal of Econometrics 54, 159–178.

Larsson, R. and J. Lyhagen (1999). Likelihood-based inference in multivariate panel cointegration models. Working paper series in Economics and Finance, no. 331, Stockholm School of Economics.

Larsson, R., J. Lyhagen and M. Lothgren (2001), Likelihood-based Cointegration Tests in Heterogenous Panels, Econometrics Journal 4: 109–142.

Levin, A. and C.F. Lin (1993), Unit root tests in panel data: asymptotic and finite-sample properties. Unpublished manuscript, University of California, San Diego.

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

Leybourne, S.J. (1995), "Testing for Unit Roots using Forward and Reverse Dickey-Fuller Regressions", *Oxford Bulletin of Economics and Statistics*, 57, 559-571.

Lyhagen, J. (2000), Why not Use Standard Panel Unit Root Tests for Testing PPP, Stockholm School of Economics, unpublished.

Maddala, G.S. and Wu, S. (1999), A Comparative Study of Unit Root Tests with Panel Data and a new simple test, Oxford Bulletin of Economics and Statistics, 61, 631–652.

Madsen, E. (2003), Unit Root Inference in Panel Data Models Where the Time-series Dimension is fixed: A comparison of different test, CAM working paper No. 2003-13.

Mark, N.C. and D. Sul (2003), Cointegration Vector Estimation by Panel DOLS and Long-run Money Demand, Oxford Bulletin of Economics and Statistics, 65, 655–680.

Mark, N.C., M. Ogaki and D. Sul (2004), Dynamic Seemingly Unrelated Cointegration Regression, forthcoming in: Review of Economic Studies.

McCoskey, S. and C. Kao (1998), A Residual-Based Test of the Null of Cointegration in Panel Data, Econometric Reviews, 17, 57–84.

McCoskey, S. and C. Kao (1999), Testing the Stability of a Production Function with Urbanization as a Shift Factor, Oxford Bulletin of Economics and Statistics, 61, 671–90.

Moon, R. and B. Perron (2004), Testing for Unit Root in Panels with Dynamic Factors, Journal of Econometrics, 122, 81–126.

Moon, H.R., B. Perron, and P.C.B. Phillips (2003), Incidental Trends and the Power of Panel Unit Root Tests, Cowles Fundation, Discussion Paper No. 1435.

Moon, H.R. and B. Perron (2005), Efficient Estimation of the SUR Cointegration Regression Model and Testing for Purchasing Power Parity, forthcoming in: Econometric Reviews.

Murray, C. J. and D.H. Papell (2002), Testing for Unit Roots in Panels in the Presence of Structural Change With an Application to OECD Unemployment, in B. Baltagi (ed.), Nonstationary Panels, Panel Cointegration, and Dynamic Panels, Advances in Econometrics, Vol. 15, JAI, Amsterdam, 223–238.

Nabeya, S. (1999), Asymptotic Moments of Some Unit Root Test Statistics in the Null Case, Econometric Theory, 15, 139–149.

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

Park, H. and W. Fuller (1995), Alternative Estimators and Unit Root Tests for the Autoregressive Process, Journal of Time Series Analysis, 16, 415-429.

Pedroni, P. (1995), Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Test with an Application to the PPP Hypothesis, Indiana University Working Papers in Economics, No. 95-013.

Pedroni, P. (1999), Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors, Oxford Bulletin of Economics and Statistics, 61, 653–670.

Pedroni, P. (2000), Fully Modified OLS for Heterogenous Cointegrated Panels, in: Baltagi B. (ed.), Nonstationary Panels, Panel Cointegration, and Dynamic Panels, Advances in Econometrics, Vol. 15, Amsterdam: JAI Press, pp. 93–130.

Pedroni, P. (2001), Purchasing Power Parity Tests in Cointegrated Panels, Review of Economics and Statistics, 83, 727–731.

Pedroni, P. (2004), Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests With an Application to the PPP Hypothesis, Econometric Theory, 20, 597–625.

Pedroni, P. and T. Vogelsang (2005), Robust Tests for Unit Roots in Heterogeneous Panels, Williams College, unpublished.

Pesaran, M.H. (1997), The Role of Economic Theory in Modelling the Long Run, Economic Journal 107, 178-91.

Pesaran M.H. (2004a), General Diagnostic Tests for Cross Section Dependence in Panels, *Cambridge Working Papers in Economics, No. 435*, University of Cambridge, and CESifo Working Paper Series No. 1229.

Pesaran, M.H., (2004b), Estimation and Inference in Large Heterogeneous Panels with Cross Section Dependence, revised version of CESifo Working Paper Series No. 869, October 2002.

Pesaran, M.H. (2005), A Simple Panel Unit Root Test in the Presence of Cross Section Dependence, Cambridge Working Papers in Economics 0346, revised Version, University of Cambridge.

Pesaran, M.H. and Y. Shin (2002), Long-Run Structural Modelling, *Econometric Reviews*, 21, 49-87.

Pesaran, M.H., T. Schuermann, and S.M. Weiner (2004), Modelling Regional Interdependencies Using a Global Error-Correcting Macroeconometric Model, Journal of Business and Economic Statistics, 22, 129–162.

Pesaran, M.H., Y. Shin and R.P. Smith (1999), Pooled Mean Group Estimation of Dynamic Heterogeneous Panels, Journal of the American Statistical Association, 94, 621–624.

Pesaran, M.H., Y. Shin and R.J. Smith (2000), Structural Analysis of Vector Error Correction Models with Exogenous I(1) Variables, Journal of Econometrics, 97, 293-343.

Pesaran, M.H., Smith, R. (1995), Estimating Long-run Relationships from Dynamic Heterogeneous Panels, Journal of Econometrics, 68, 79-113.

Pesaran, M.H., R. Smith, and K.S. Im (1996), Dynamic linear models for heterogenous panels. In: Matyas, L., Sevestre, P. (eds), The Econometrics of Panel Data: A Handbook of the Theory with Applications, second revised edition, pp. 145-195.

Pesaran, M.H. and T. Yamagat (2005), Testing Slope Homogeneity in Large Panels, Cambridge Working Papers in Economics, March 2005, number 0513.

Phillips, P.C.B. (1991), Optimal Inference in Co-integrated Systems, Econometrica, 59, 282–306.

Phillips, P.C.B. (1994), Some Exact Distribution Theory for Maximum Likelihood Estimators of Cointegrating Coefficients in Error Correction Models, Econometrica, 62, 73–93.

Phillips, P.C.B. and B.E. Hansen (1990), Statistical Inference in Instrumental Variable Regression with I(1) Processes, Review of Economic Studies 57: 99–125.

Phillips, P.C.B. and H.R. Moon (1999), Linear Regression Limit Theory for Nonstationary Panel Data, Econometrica, 67, 1057–1111.

Phillips, P.C.B. and S. Ouliaris (1990), Asymptotic Properties of Residual Based Tests for Cointegration, Econometrica, 58, 165–193.

Phillips, P.C.B. and D. Sul (2003a), Dynamic Panel Estimation and Homogeneity Testing Under Cross Section Dependence, Econometrics Journal, 6, 217-259.

Phillips, P.C.B. and D. Sul (2003b), The Elusive Empirical Shadow of Growth Convergence, Cowles Foundation Discussion Paper 98, Yale University.

Ploberger, W. and P.C.B. Phillips (2002), Optimal Testing for Unit Roots in Panel Data, *unpublished*.

Robinson, P.M. (1994), Efficient Tests of Nonstationary Hypotheses, Journal of the American Statistical Association, 89, 1420-1437.

Saikkonen, P. (1991), Asymptotic Efficient Estimation of Cointegration Regressions, Econometric Theory, 7, 1–21.

Schmidt, P. and P.C.B. Phillips (1992), LM Test for a Unit Root in the Presence of Deterministic Trends, Oxford Bulletin of Economics and Statistics 54, 257–287.

Shin, Y. (1994), A Residual-based Test of the Null of Cointegration Against the Alternative of no Cointegration, Econometric Theory, 10, 91—-115.

Smith, V., S. Leybourne, T.-H. Kim and P. Newbold (2004), More Powerful Panel Data Unit Root Tests With an Application to Mean Reversion in Real Exchange Rates, Journal of Applied Econometrics, 19, 147–170.

Tanaka, K. (1990), Testing for a Moving Average Root, Econometric Theory, 6, 433–444.

Tzavalis, E. (2002), Structural Breaks and Unit Root Tests for Short Panels, Queen Mary University of London, unpublished.

Westerlund, J. (2005a), New Simple Tests for Panel Cointegration, forthcoming in Econometrics Reviews.

Westerlund, J. (2005b), A Panel CUSUM Test of the Null of Cointegration, Oxford Bulletin of Economics and Statistics, 62, 231–262.

Westerlund, J. (2005c), Data Dependent Endogeneity Correction in Cointegrated Panels, forthcoming in: Oxford Bulletin of Economics and Statistics.

CESifo Working Paper Series

(for full list see www.cesifo-group.de)

1505 Peter A. Zadrozny, Necessary and Sufficient Restrictions for Existence of a Unique Fourth Moment of a Univariate GARCH(p,q) Process, July 2005

1506 Rainer Niemann and Corinna Treisch, Group Taxation, Asymmetric Taxation and Cross-Border Investment Incentives in Austria, July 2005

1507 Thomas Christiaans, Thomas Eichner and Ruediger Pethig, Optimal Pest Control in Agriculture, July 2005

1508 Biswa N. Bhattacharyay and Prabir De, Promotion of Trade and Investments between China and India: The Case of Southwest China and East and Northeast India, July 2005

1509 Jean Hindriks and Ben Lockwood, Decentralization and Electoral Accountability: Incentives, Separation, and Voter Welfare, July 2005

1510 Michelle R. Garfinkel, Stergios Skaperdas and Constantinos Syropoulos, Globalization and Domestic Conflict, July 2005

1511 Jesús Crespo-Cuaresma, Balázs Égert and Ronald MacDonald, Non-Linear Exchange Rate Dynamics in Target Zones: A Bumpy Road towards a Honeymoon – Some Evidence from the ERM, ERM2 and Selected New EU Member States, July 2005

1512 David S. Evans and Michael Salinger, Curing Sinus Headaches and Tying Law: An Empirical Analysis of Bundling Decongestants and Pain Relievers, August 2005

1513 Christian Keuschnigg and Martin D. Dietz, A Growth Oriented Dual Income Tax, July 2005

1514 Fahad Khalil, David Martimort and Bruno Parigi, Monitoring a Common Agent: Implications for Financial Contracting, August 2005

1515 Volker Grossmann and Panu Poutvaara, Pareto-Improving Bequest Taxation, August 2005

1516 Lars P. Feld and Emmanuelle Reulier, Strategic Tax Competition in Switzerland: Evidence from a Panel of the Swiss Cantons, August 2005

1517 Kira Boerner and Silke Uebelmesser, Migration and the Welfare State: The Economic Power of the Non-Voter?, August 2005

1518 Gabriela Schütz, Heinrich W. Ursprung and Ludger Wößmann, Education Policy and Equality of Opportunity, August 2005

1519 David S. Evans and Michael A. Salinger, Curing Sinus Headaches and Tying Law: An Empirical Analysis of Bundling Decongestants and Pain Relievers, August 2005

1520 Michel Beine, Paul De Grauwe and Marianna Grimaldi, The Impact of FX Central Bank Intervention in a Noise Trading Framework, August 2005

1521 Volker Meier and Matthias Wrede, Pension, Fertility, and Education, August 2005

1522 Saku Aura and Thomas Davidoff, Optimal Commodity Taxation when Land and Structures must be Taxed at the Same Rate, August 2005

1523 Andreas Haufler and Søren Bo Nielsen, Merger Policy to Promote 'Global Players'? A Simple Model, August 2005

1524 Frederick van der Ploeg, The Making of Cultural Policy: A European Perspective, August 2005

1525 Alexander Kemnitz, Can Immigrant Employment Alleviate the Demographic Burden? The Role of Union Centralization, August 2005

1526 Baoline Chen and Peter A. Zadrozny, Estimated U.S. Manufacturing Production Capital and Technology Based on an Estimated Dynamic Economic Model, August 2005

1527 Marcel Gérard, Multijurisdictional Firms and Governments' Strategies under Alternative Tax Designs, August 2005

1528 Joerg Breitscheidel and Hans Gersbach, Self-Financing Environmental Mechanisms, August 2005

1529 Giorgio Fazio, Ronald MacDonald and Jacques Mélitz, Trade Costs, Trade Balances and Current Accounts: An Application of Gravity to Multilateral Trade, August 2005

1530 Thomas Christiaans, Thomas Eichner and Ruediger Pethig, A Micro-Level 'Consumer Approach' to Species Population Dynamics, August 2005

1531 Samuel Hanson, M. Hashem Pesaran and Til Schuermann, Firm Heterogeneity and Credit Risk Diversification, August 2005

1532 Mark Mink and Jakob de Haan, Has the Stability and Growth Pact Impeded Political Budget Cycles in the European Union?, September 2005

1533 Roberta Colavecchio, Declan Curran and Michael Funke, Drifting Together or Falling Apart? The Empirics of Regional Economic Growth in Post-Unification Germany, September 2005

1534 Kai A. Konrad and Stergios Skaperdas, Succession Rules and Leadership Rents, September 2005

1535 Robert Dur and Amihai Glazer, The Desire for Impact, September 2005

1536 Wolfgang Buchholz and Wolfgang Peters, Justifying the Lindahl Solution as an Outcome of Fair Cooperation, September 2005

1537 Pieter A. Gautier, Coen N. Teulings and Aico van Vuuren, On-the-Job Search and Sorting, September 2005

1538 Leif Danziger, Output Effects of Inflation with Fixed Price- and Quantity-Adjustment Costs, September 2005

1539 Gerhard Glomm, Juergen Jung, Changmin Lee and Chung Tran, Public Pensions and Capital Accumulation: The Case of Brazil, September 2005

1540 Yvonne Adema, Lex Meijdam and Harrie A. A. Verbon, The International Spillover Effects of Pension Reform, September 2005

1541 Richard Disney, Household Saving Rates and the Design of Social Security Programmes: Evidence from a Country Panel, September 2005

1542 David Dorn and Alfonso Sousa-Poza, Early Retirement: Free Choice or Forced Decision?, September 2005

1543 Clara Graziano and Annalisa Luporini, Ownership Concentration, Monitoring and Optimal Board Structure, September 2005

1544 Panu Poutvaara, Social Security Incentives, Human Capital Investment and Mobility of Labor, September 2005

1545 Kjell Erik Lommerud, Frode Meland and Odd Rune Straume, Can Deunionization Lead to International Outsourcing?, September 2005

1546 Robert Inklaar, Richard Jong-A-Pin and Jakob de Haan, Trade and Business Cycle Synchronization in OECD Countries: A Re-examination, September 2005

1547 Randall K. Filer and Marjorie Honig, Endogenous Pensions and Retirement Behavior, September 2005

1548 M. Hashem Pesaran, Til Schuermann and Bjoern-Jakob Treutler, Global Business Cycles and Credit Risk, September 2005

1549 Ruediger Pethig, Nonlinear Production, Abatement, Pollution and Materials Balance Reconsidered, September 2005

1550 Antonis Adam and Thomas Moutos, Turkish Delight for Some, Cold Turkey for Others?: The Effects of the EU-Turkey Customs Union, September 2005

1551 Peter Birch Sørensen, Dual Income Taxation: Why and how?, September 2005

1552 Kurt R. Brekke, Robert Nuscheler and Odd Rune Straume, Gatekeeping in Health Care, September 2005

1553 Maarten Bosker, Steven Brakman, Harry Garretsen and Marc Schramm, Looking for Multiple Equilibria when Geography Matters: German City Growth and the WWII Shock, September 2005

1554 Paul W. J. de Bijl, Structural Separation and Access in Telecommunications Markets, September 2005

1555 Ueli Grob and Stefan C. Wolter, Demographic Change and Public Education Spending: A Conflict between Young and Old?, October 2005

1556 Alberto Alesina and Guido Tabellini, Why is Fiscal Policy often Procyclical?, October 2005

1557 Piotr Wdowinski, Financial Markets and Economic Growth in Poland: Simulations with an Econometric Model, October 2005

1558 Peter Egger, Mario Larch, Michael Pfaffermayr and Janette Walde, Small Sample Properties of Maximum Likelihood Versus Generalized Method of Moments Based Tests for Spatially Autocorrelated Errors, October 2005

1559 Marie-Laure Breuillé and Robert J. Gary-Bobo, Sharing Budgetary Austerity under Free Mobility and Asymmetric Information: An Optimal Regulation Approach to Fiscal Federalism, October 2005

1560 Robert Dur and Amihai Glazer, Subsidizing Enjoyable Education, October 2005

1561 Carlo Altavilla and Paul De Grauwe, Non-Linearities in the Relation between the Exchange Rate and its Fundamentals, October 2005

1562 Josef Falkinger and Volker Grossmann, Distribution of Natural Resources, Entrepreneurship, and Economic Development: Growth Dynamics with Two Elites, October 2005

1563 Yu-Fu Chen and Michael Funke, Product Market Competition, Investment and Employment-Abundant versus Job-Poor Growth: A Real Options Perspective, October 2005

1564 Kai A. Konrad and Dan Kovenock, Equilibrium and Efficiency in the Tug-of-War, October 2005

1565 Joerg Breitung and M. Hashem Pesaran, Unit Roots and Cointegration in Panels, October 2005