Cross-Sectional Correlation Robust Tests for Panel Cointegration*

Christoph Hanck†

This version: November 13, 2006

Abstract

We use meta analytic combination procedures to develop new tests for panel cointegration. The main idea consists in combining p-values from time series cointegration tests on the different units of the panel. The tests are robust to heterogeneity as well as to cross-sectional dependence between the different units of the panel. To achieve the latter, we employ a sieve bootstrap procedure with joint resampling of the residuals of the different units. A simulation study shows that the suggested bootstrap tests can have substantially smaller error-in-rejection probabilities than tests ignoring the presence of cross-sectional dependence while preserving high power. We apply the tests to a panel of Post-Bretton Woods data to test for weak Purchasing Power Parity (PPP).

Keywords: panel cointegration tests, cross-sectional dependence, sieve bootstrap

*Research supported by DFG under Sonderforschungsbereich 475. I would like to thank conference participants at the FEMES 2006, Beijing and the Statistische Woche 2006, Dresden, and Pavel Stoimenov for helpful comments.

†Research supported by Ruhr Graduate School in Economics. Contact Information: Tel. (+49) 0231-7553127, christoph.hanck@uni-dortmund.de.

1 Introduction

The application of unit root and cointegration tests in macroeconometric practice is often hampered by the lack of large sample sizes. The (first-order) asymptotic approximation used for deriving the distribution of the test statistics may then be rather inaccurate. One solution to improve the power and reduce the error-in-rejection probability (size distortion) of these tests is to pool several time series into a panel data set and develop panel unit root and cointegration tests.

Pedroni [2004] and Kao [1999] generalize the residual-based tests of Engle and Granger [1987] and Phillips and Ouliaris [1990], Larsson et al. [2001] extend the Johansen [1988] tests to panel data while McCoskey and Kao [1998] propose a test for the null of panel cointegration in the spirit of Shin [1994]. All these tests, however, rely on the assumption that the different cross-sectional units of the panel are independent or, at most, exhibit dependence of a rather simple form. This assumption, characterizing the so-called first generation tests, greatly simplifies the derivation of limiting distributions of the panel test statistics, but may not hold in practice. Phillips and Sul [2003], Moon and Perron [2004] and Bai and Ng [2004] put forward factor approaches to deal with cross-sectional correlation in panel unit root and cointegration testing. Their approach may, however, require the validity of the factor structure assumption modelling the correlation structure of the panel units. As argued by Breitung and Das [2005], size distortions may result if this assumption, which is hard to verify, is not met.

The main contribution of the present paper therefore is to suggest new tests for panel cointegration that are robust to cross-sectional dependence (or, synonymously, cross-sectional correlation) of an arbitrary form. The main idea of the testing principle has been used in meta analytic studies for a long time [see Fisher, 1970; Hedges and Olkin, 1985] and was introduced into the panel literature by Maddala and Wu [1999] and Choi [2001], who propose meta analytic panel unit root tests: Consider the testing problem on the panel as consisting of N testing problems for each unit of the panel. Conduct N separate time series tests and obtain the corresponding p-values of the test statistics. Then, combine the p-values of the N tests (in a sense to be made precise below) into a single panel test statistic. Hanck [2005] extends their framework to the panel cointegration setting.

To robustify the panel cointegration tests against cross-sectional correlation of an arbitrary form, we use a bootstrap scheme that jointly resamples entire cross-sections of residuals to preserve the cross-sectional correlation structure in the panel. We provide some simulation

evidence to demonstrate the effectiveness of the suggested procedure. In particular, the bootstrap tests can have dramatically smaller error-in-rejection probabilities than tests ignoring the presence of cross-sectional dependence. At the same time, the bootstrap tests preserve high power. We use the tests to investigate the weak PPP hypothesis for a panel of Post-Bretton Woods exchange rate data. Our main result is that using cross-sectional correlation corrected critical values may make an important difference in econometric practice.

The remainder of the paper is organized as follows. The next section establishes notation and reviews the meta analytic p-value combination tests for panel cointegration under the assumption of cross-sectional independence. Section 3 discusses the bootstrap algorithm used to robustify the tests against general forms of cross-sectional dependence. Section 4 summarizes the simulation evidence on the effectiveness of the bootstrap tests under cross-sectional dependence. Section 5 illustrates the use of the tests. The final section concludes.

2 *P*-Value Combination Tests for Panel Cointegration

Consider the multivariate time series regression

$$y_{i,t} = \alpha_i + \kappa_i t + \vartheta_i t^2 + \boldsymbol{\beta}_i \mathbf{x}_{i,t} + u_{i,t} \qquad (i \in \mathbb{N}_N,\ t \in \mathbb{N}_{T_i}) \quad (1)$$

for each of the N units of a possibly unbalanced panel. ($a \in \mathbb{N}_c^b$ is shorthand for $a = b, \ldots, c$, omitting b if $b = 1$.) The $(K \times 1)$ column vector $\mathbf{x}_{i,t} = (x_{i1t}, \ldots, x_{iKt})^\top$ $(i \in \mathbb{N}_N,\ t \in \mathbb{N}_{T_i})$ collects the observations on the K regressors for given i and t. $y_{i,t}$ and $\mathbf{x}_{i,t}$ are integrated of order one, $I(1)$ $(i \in \mathbb{N}_N)$. The row vector $\boldsymbol{\beta}_i$, α_i, κ_i and ϑ_i may vary across i, allowing for heterogeneous cointegrating vectors and time polynomials of order up to two, i.e. constants, trend and squared trend terms. We make the following Functional Central Limit Assumption on the Data Generating Process (DGP) of the variables.

ASSUMPTION 1 *(Invariance Principle).*
Let $z_{i,t} := (y_{i,t}, \mathbf{x}_{i,t}^\top)^\top$ and $\xi_{i,t} := (\xi_{iyt}, \xi_{i1t}, \ldots, \xi_{iKt})^\top$. The true process $z_{i,t}$ is generated as $z_{i,t} = z_{i,t-1} + \xi_{i,t}$, $(i \in \mathbb{N}_N, t \in \mathbb{N}_{T_i})$.
$\xi_{i,t}$ satisfies $T_i^{-\frac{1}{2}} \sum_{t=1}^{[T_i r]} \xi_{i,t} \Rightarrow B_{i,\Omega_i}(r)$ $(i \in \mathbb{N}_N)$ as $T_i \to \infty$, where $r \in [0,1]$ and $\Rightarrow$ denotes weak convergence. $[x]$ is the integer part of x and $B_{i,\Omega_i}(r)$

is vector Brownian motion with asymptotic covariance matrix Ω_i. Also, the $K \times K$ lower right submatrix of Ω_i, $\Omega_{\mathbf{xx},i}$, has full rank.

This assumption ensures, among other things, that there are no cointegrating relationships among the regressors in (1).[1] p_i denotes the p-value of a time series cointegration test applied to the ith unit of the panel. Let θ_{i,T_i} be a time series cointegration test statistic on unit i for a sample size of T_i. Let F_{T_i} denote the null cumulative distribution function (cdf) of θ_{i,T_i}. Since the tests considered here are one-sided, $p_i = F_{T_i}(\theta_{i,T_i})$ if the test rejects for small values of θ_{i,T_i} and $p_i = 1 - F_{T_i}(\theta_{i,T_i})$ if the test rejects for large values of θ_{i,T_i}. We only consider time series tests with the null of no cointegration.

We test the following null hypothesis:

$$H_0: \text{ There is no cointegration for any } i, \qquad (i \in \mathbb{N}_N)$$

against the alternative

$$H_1: \text{ There is cointegration for at least one } i.$$

Under H_0, $\{u_{i,t}\}_t$ in (1) is an $I(1)$ stochastic process ($i \in \mathbb{N}_N$). The alternative H_1 states that there are 1 to N cointegrated units in the panel. That is, a rejection neither allows to conclude that the entire panel is cointegrated nor does it provide information about the number of units of the panel that exhibit cointegrating relationships.

We make the following assumptions [see Pedroni, 2004]:

Assumption 2 *(Continuity).*
Under H_0, θ_{i,T_i} has a continuous distribution function ($i \in \mathbb{N}_N$).

Assumption 3 *(Cross-Sectional Uncorrelatedness).*
$\mathsf{E}[\xi_{i,t}\xi_{l,s}^\top] = \mathbf{0}$ $(s,t \in \mathbb{N}_{T_i}, i \neq l)$. The error process $\xi_{i,t}$ is generated as a linear vector process $\xi_{i,t} = C_i(L)\eta_{i,t}$, where L is the lag operator and $\eta_{i,t}$ is vector white noise.

Remarks

- Assumption 2 is a regularity condition which ensures a uniform distribution of the p-values of the time series test statistics under H_0: $p_i \sim \mathcal{U}[0,1]$ $(i \in \mathbb{N}_N)$ [see, e.g., Bickel and Doksum, 2001, Sec. 4.1]. It is satisfied by the time series tests considered in this paper.

[1] See Pedroni [2004] for further discussion.

- Assumption 3 is strong [see, e.g., Banerjee et al., 2005]. It implies that the different units of a panel must not be linked to each other beyond relatively simple forms of correlation such as common time effects. These can be eliminated by demeaning across the cross-sectional dimension. This assumption is likely to be violated in many typical macroeconomic panel data sets. For instance, consider a panel data set consisting of exchange rates vis-à-vis the U.S. dollar. The exchange rates of, say, the Euro and the Mexican Peso generally do not react identically to a macroeconomic shock in the U.S., given the very different structure of financial and trade links with the U.S.

- We emphasize that Assumption 3 is sufficient, but not necessary. Section 3 presents an approach that allows to dispense with this assumption.

We now present the test statistics for panel cointegration put forward in this paper. Combine the N p-values of the individual time series cointegration tests, p_i ($i \in \mathbb{N}_N$), as follows:

$$P_{\chi^2} = -2\sum_{i=1}^{N} \ln(p_i) \tag{2a}$$

$$P_{\Phi^{-1}} = N^{-\frac{1}{2}} \sum_{i=1}^{N} \Phi^{-1}(p_i), \tag{2b}$$

where Φ^{-1} denotes the inverse of the cdf of the standard normal distribution.[2] When considered as a group we refer to Eqs. (2a) and (2b) as P tests. Furthermore, we refer to P tests relying on Assumption 3 as "simple" P tests. The P tests, via pooling p-values, provide convenient tests for panel cointegration by imposing minimal homogeneity restrictions on the panel. For instance, the different units of the panel can be unbalanced. Furthermore, the evidence for (non-)cointegration is first investigated for each unit of the panel and then compactly expressed with the p-value of the time series cointegration test. Hence, the coefficients describing the relationship between the different variables for each unit of the panel can be heterogeneous across i. Thus, the availability of large-T time series allows for pooling the data into a panel without having to impose strong homogeneity restrictions on $\boldsymbol{\beta}$ as in traditional panel data analysis.[3]

We now turn to the asymptotic distributions of the tests.

[2] See also Maddala and Wu [1999] and Choi [2001].

[3] For an overview of panel data models relying on $N \to \infty$ asymptotics see Hsiao [2003].

THEOREM 1.

Under the null of no panel cointegration and Assumptions 1, 2 and 3, as $T_i \to \infty$ ($i \in \mathbb{N}_N$), the P tests are asymptotically distributed as

$$(i)\ \ P_{\chi^2} \to_d \chi^2_{2N}$$
$$(ii)\ \ P_{\Phi^{-1}} \to_d \mathcal{N}(0,1).$$

PROOF. (i) The proof is an application of the transformation theorem for absolutely continuous random variables (r.v.s) [see, e.g., Bierens, 2005, Thm. 4.2]. Under H_0, $p_i \sim \mathcal{U}[0,1]$. Let $y = g(p_i) := -2\ln(p_i)$. Then, $p_i = g^{-1}(y) = e^{-\frac{1}{2}y}$ and the density of $-2\ln(p_i)$ is given by

$$f_{-2\ln(p_i)}(y) = f_{p_i}(g^{-1}(y)) \left| \frac{\partial}{\partial y} g^{-1}(y) \right|.$$

Hence, $\frac{\partial}{\partial y} g^{-1}(y) = -\frac{1}{2}e^{-\frac{1}{2}y}$ and $|\frac{\partial}{\partial y} g^{-1}(y)| = \frac{1}{2}e^{-\frac{1}{2}y}$. We have $f_{p_i}(g^{-1}(y)) = 1 \ \forall\ g^{-1}(y) \in [0,1]$. This implies $f_{-2\ln(p_i)}(y) = \frac{1}{2}e^{-\frac{1}{2}y}$. The density of a χ^2_R r.v. is $f_{\chi^2_R}(y) = \frac{1}{2^{R/2}\Gamma(R/2)} y^{\frac{R}{2}-1} e^{-\frac{y}{2}}$. With $R = 2$, we get $f_{\chi^2_2}(y) = \frac{1}{2\Gamma(1)} e^{-\frac{y}{2}}$. Recall that $\Gamma(1) = \int_0^\infty t^{1-1} e^{-t}\, dt = 1$. So,

$$f_{\chi^2_2}(y) = \frac{1}{2} e^{-\frac{y}{2}}.$$

We have shown that $f_{-2\ln(p_i)}(y) = f_{\chi^2_2}(y)$. The proof is complete since the sum of N *independent* χ^2_R r.v.s is distributed as χ^2_{NR}.

(ii) Since, under H_0, $\mathsf{P}(\Phi^{-1}(p_i) \leqslant x) = \mathsf{P}(p_i \leqslant \Phi(x)) = \Phi(x)$, we have $\Phi^{-1}(p_i) \sim \mathcal{N}(0,1)$. By the convolution theorem [see, e.g., Spanos, 1986, pp. 99], the sum of N *independent* $\mathcal{N}(0,1)$ r.v.s has a $\mathcal{N}(0,N)$ distribution. Hence, $P_{\Phi^{-1}}$ is also normal with

$$\mathsf{E}[P_{\Phi^{-1}}] = 0 \ \text{ and } \ \mathsf{Var}[P_{\Phi^{-1}}] = \mathsf{Var}\left[N^{-\frac{1}{2}} \sum\nolimits_{i=1}^{N} \Phi^{-1}(p_i)\right] = \frac{1}{N}\mathsf{Var}\left[\sum\nolimits_{i=1}^{N} \Phi^{-1}(p_i)\right] = 1. \qquad \square$$

Using consistent (as $T_i \to \infty$) time series cointegration tests, $p_i \to_p 0$ under the alternative of cointegration. Hence, quite intuitively, the smaller p_i, the more it contributes towards rejecting the null of no panel cointegration. The decision rule therefore is to reject the null of no panel cointegration when the realized test statistic P_{χ^2} exceeds the critical value from a χ^2_{2N} distribution at the desired significance level. For (2b) one would reject for large negative values of the panel test statistic $P_{\Phi^{-1}}$. We also see from the proof that the tests have a well-defined asymptotic distribution for any finite N. This feature is attractive because in many applications of panel cointegration analysis like the above example, the assumption of N, the number of units in the panel, going to infinity may not be a natural one. To rationalize the alternative hypothesis H_1, note that a small fraction of the units of the panel exhibiting strong evidence of time series cointegration, thus yielding low p-values, might lead to a rejection of the null hypothesis. Thus, as stated above, rejection of H_0 should not be taken as evidence of the entire panel being cointegrated.

We now discuss how to obtain the p-values required for computation of the P test statistics. The null distributions of residual based cointegration tests generally converge to functionals of Brownian motion. Hence, analytic expressions of the distribution functions are not available and p-values of the tests cannot simply be obtained by evaluating the corresponding cdf. In the time series case, it is now fairly standard practice to obtain p-values of unit root and cointegration tests using response surface regressions. We use p-values of the Augmented Dickey-Fuller (ADF) cointegration tests [Engle and Granger, 1987] as provided by MacKinnon [1996].[4] Suppressing deterministic trend terms for brevity, the p-values are derived from the t-statistic of $\varrho_i - 1$ in the OLS regression

$$\Delta\hat{u}_{i,t} = (\varrho_i - 1)\hat{u}_{i,t-1} + \sum_{j=1}^{J_i} \nu_j \Delta\hat{u}_{i,t-j} + \epsilon_{J_i,i,t}, \tag{3}$$

where $\Delta := 1 - L$ and J_i is the number of lagged differences required to render $\epsilon_{J_i,i,t}$ white noise. Here, $\hat{u}_{i,t}$ is the usual OLS residual from the first stage Engle and Granger [1987] regression (1). However, as should be clear from the above discussion, the P tests are general enough to accommodate any time series cointegration test for which p-values are available. Alternatively, one could, for instance, capture serial correlation by the semiparametric approach of Phillips and Ouliaris [1990].[5] We focus on the Engle and Granger [1987] ADF cointegration test because of its popularity and widespread availability.

3 Allowing for cross-sectional error dependence

We now relax Assumption 3. As can be seen from the proof of Theorem 1, this assumption guarantees the correct null distributions of the test statistics. Theorem 1 no longer holds under general forms of cross-sectional dependence. We suggest a bootstrap approach to capture the dependence structure in the panel with the aim to construct panel cointegration tests robust to the presence of cross-sectional correlation. We employ the sieve

[4] MacKinnon improves upon his prior work by using a heteroskedasticity and serial correlation robust technique to approximate between the estimated quantiles of the response surface regressions.

[5] In Hanck [2005], we also employ the p-values for the Johansen [1988] λ_{trace} and λ_{max} tests. The corresponding P tests perform poorly, however, as the Johansen [1988] time series cointegration tests severely overreject for the time series lengths usually available in macroeconometric practice. Accordingly, the corresponding p-values are not distributed as $\mathcal{U}[0,1]$ under H_0 and hence not suitable for the P tests. Hanck [2005] also discusses other p-value combination tests and provides more extensive Monte Carlo evidence on the simple P tests.

bootstrap.[6] The sieve bootstrap approximates $u_{i,t}$ with a finite order autoregressive process, where the order increases with sample size, and resamples from the residuals. Under the following assumption, the sieve bootstrap yields an accurate approximation [Chang and Park, 2003]:

> Assumption 4 *(Linearity).*
> The first differences of the equilibrium errors are generated as (possibly heterogeneous) linear processes, $\Delta u_{i,t} = \phi_i(L)\epsilon_{i,t}$, where $\phi_i(z) := \sum_{\ell=0}^{\infty} \phi_{i,\ell} z^\ell$.

More precisely, the bootstrap algorithm proceeds as follows.

1. Compute the P test statistic(s) according to (2a) and (2b). Denote the realizations by $\tilde{P}_{\chi^2}$ and $\tilde{P}_{\Phi^{-1}}$.
2. Suppressing deterministic trend terms and denoting estimates by a $\hat{}$, estimate equation (1) by OLS:
$$y_{i,t} = \hat{\alpha}_i + \hat{\boldsymbol{\beta}}_i \mathbf{x}_{i,t} + \hat{u}_{i,t}. \qquad (i \in \mathbb{N}_N,\ t \in \mathbb{N}_{T_i})$$
3. Fit an autoregressive process to $\Delta\hat{u}_{i,t}$ $(i \in \mathbb{N}_N, t \in \mathbb{N}_{T_i}^2)$. It is natural to use the Yule-Walker procedure because it always yields an invertible representation [Brockwell and Davis, 1991, Secs. 8.1–2]. Letting $\overline{\Delta\hat{u}_i} := (T_i - 1)^{-1} \sum_{t=2}^{T_i} \Delta\hat{u}_{i,t}$, compute
$$\hat{\gamma}_i(j) := \frac{1}{T_i - 1 - j} \sum_{t=2}^{T_i - j} (\Delta\hat{u}_{i,t} - \overline{\Delta\hat{u}_i})(\Delta\hat{u}_{i,t+j} - \overline{\Delta\hat{u}_i}), \qquad (i \in \mathbb{N}_N,\ j \in \mathbb{N}_q)$$
the empirical autocovariances of $\Delta\hat{u}_{i,t}$ up to order q. Defining
$$\hat{\Gamma}_{i,q} := \begin{pmatrix} \hat{\gamma}_i(0) & \cdots & \hat{\gamma}_i(q-1) \\ \vdots & \ddots & \vdots \\ \hat{\gamma}_i(q-1) & \cdots & \hat{\gamma}_i(0) \end{pmatrix}$$
and $\hat{\gamma}_i := (\hat{\gamma}_i(1), \ldots, \hat{\gamma}_i(q))^\top$, obtain the AR coefficient vector as
$$(\hat{\phi}_{q,i,1}, \ldots, \hat{\phi}_{q,i,q})^\top := \hat{\Gamma}_{i,q}^{-1} \hat{\gamma}_i. \qquad (i \in \mathbb{N}_N)$$
4. The residuals are, as usual, given by
$$\hat{\epsilon}_{q,i,t} := \Delta\hat{u}_{i,t} - \sum_{\ell=1}^{q} \hat{\phi}_{q,i,\ell} \Delta\hat{u}_{i,t-\ell}. \qquad (i \in \mathbb{N}_N, t \in \mathbb{N}_{T_i}^{q+2})$$
Center $\hat{\epsilon}_{q,i,t}$ to obtain
$$\tilde{\epsilon}_{q,i,t} := \hat{\epsilon}_{q,i,t} - \frac{1}{T_i - q - 1} \sum_{g=q+2}^{T_i} \hat{\epsilon}_{q,i,g}. \qquad (i \in \mathbb{N}_N, t \in \mathbb{N}_{T_i}^{q+2})$$

[6] For related approaches, see Maddala and Wu [1999] and Swensen [2003].

5. Resample nonparametrically from $\tilde{\epsilon}_{q,i,t}$ to get $\epsilon^*_{q,i,t}$. To preserve the empirical cross-sectional dependence structure, jointly resample residual vectors

$$\tilde{\epsilon}_{q,\bullet,t} := (\tilde{\epsilon}_{q,1,t}, \ldots, \tilde{\epsilon}_{q,N,t}). \qquad (t \in \mathbb{N}_{T_i}^{q+2})$$

6. Recursively construct the bootstrap samples as[7]

$$\Delta u^*_{q,i,t} = \sum_{\ell=1}^{q} \hat{\phi}_{q,i,\ell} \Delta u^*_{q,i,t-\ell} + \epsilon^*_{q,i,t}. \qquad (i \in \mathbb{N}_N, t \in \mathbb{N}_{T_i}^{q+2})$$

7. It is necessary to impose the null of a unit root when generating the artificial data in bootstrap unit root tests to achieve consistency [Basawa et al., 1991]. Accordingly, impose the null of non-cointegration by integrating $\Delta u^*_{i,t}$ to obtain $u^*_{i,t}$ and form

$$y^*_{q,i,t} = \hat{\alpha}_i + \hat{\boldsymbol{\beta}}_i \mathbf{x}_{i,t} + u^*_{q,i,t}. \qquad (i \in \mathbb{N}_N, t \in \mathbb{N}_{T_i})$$

8. Perform the P tests using the artificial data set $(y^*_{q,i,t}, \mathbf{x}_{i,t}^\top)^\top$. Denote the realizations of the test statistics by, e.g., $P^{b*}_{\chi^2}$.
9. Repeat steps 4 to 8 many, say B, times.
10. Denote the indicator function by $\mathbf{1}\{\ \}$ and choose a significance level α. Then, reject H_0 of the P_{χ^2} or the $P_{\Phi^{-1}}$ test if

$$\frac{1}{B}\sum_{b=1}^{B} \mathbf{1}\{P^{b*}_{\chi^2} > \tilde{P}_{\chi^2}\} < \alpha \qquad \text{or} \qquad \frac{1}{B}\sum_{b=1}^{B} \mathbf{1}\{P^{b*}_{\Phi^{-1}} < \tilde{P}_{\Phi^{-1}}\} < \alpha, \qquad (4)$$

respectively.

Remarks

- We provide no formal proof of the consistency of this bootstrap procedure. It might be conjectured from the proofs of bootstrap consistency for unit root tests [Swensen, 2003; Chang and Park, 2003] and for inference in cointegrating regressions [Chang et al., 2006]. The latter authors argue that their results may hold more generally, e.g., for panel cointegration models.

- Steps 4 and 6 respectively "prewhiten" (or "sieve") and "recolour" the residuals using the sieve bootstrap. Thus, we attempt to generate a valid bootstrap distribution of the data across the time series dimension using a parametric AR approximation to the true DGP. There is, however, no plausible parametric approximation of the

[7] We run the recursion for 30 initial observations before using the $\Delta u^*_{q,i,t}$ to mitigate the effect of initial conditions.

dependence structure across the cross-sectional dimension. We therefore resample entire cross-sections of residuals to preserve the cross-sectional dependence structure of the data. The resampling scheme across the cross-sectional dimension is thus similar to block bootstrap procedures [Künsch, 1989].

- The selection of the lag order q in step 3 can be based on any of the well-known selection criteria such as the Akaike Information Criterion or a top-down procedure. The goal of the sieve bootstrap is to prewhiten the residuals across t to obtain random resamples from $\tilde{\epsilon}_{q,i,t}$. Hence, a selection scheme based on the whiteness of $\tilde{\epsilon}_{q,i,t}$ is also an appealing choice.

- It is possible to let q vary over i, $q_i \neq q$, to capture heterogeneity in the error processes. We do not make this possibility explicit in the notation.

- It is also possible to compute the bootstrap P tests based on resamples from prewhitened residuals of a VAR regression of $(\Delta\hat{u}_{i,t}, \Delta\mathbf{x}_{i,t}^{\top})^{\top}$. For a related purpose, Chang et al. [2006] advocate a similar scheme in order to capture endogeneity of the regressors $\mathbf{x}_{i,t}$. Our simulation results are however similar to the ones to be presented for the sieve bootstrap in the next section. The VAR approach is computationally considerably more expensive and is therefore not discussed in detail. Results are available upon request.

4 Monte Carlo study

We perform a Monte Carlo study of the tests proposed in the previous sections. The main results are as follows. The simple P tests can have high errors in rejection probabilities (ERPs) when the units of the panel exhibit cross-sectional correlation of a general form.[8] Compared to the simple P tests, bootstrapping the P tests as outlined in the previous section can strongly reduce the ERP.

The DGP used here is an extension of a design which has been used in many Monte Carlo studies of (panel) cointegration tests. See Engle and Granger [1987] and, for the extension to the panel data setting, Kao [1999]. For simplicity, consider the bivariate

[8]Suppose the DGP is in the null hypothesis set, the tests are performed at a nominal level α and the rejection frequency, or type I error rate, of the test is $R(\alpha)$. Then, $ERP := |R(\alpha) - \alpha|$. The term size distortion is often used synonymously [see Davidson and MacKinnon, 1999].

case, i.e. $K = 1$.

$$y_{i,t} - \alpha_i - \beta_i x_{i,t} = v_{i,t}, \qquad a_1 y_{i,t} - a_2 x_{i,t} = w_{i,t} \qquad (i \in \mathbb{N}_N) \tag{5}$$

where

$$\begin{aligned} v_{i,t} &= \rho_i v_{i,t-1} + \breve{e}_{zi,t}, & \Delta w_{i,t} &= \breve{e}_{wi,t}, \\ \breve{e}_{zi,t} &= e_{zi,t} + \lambda_i \varepsilon_t & \breve{e}_{wi,t} &= e_{wi,t} + \pi_i e_{wi,t-1} \end{aligned}$$

and

$$\begin{pmatrix} e_{zi,t} \\ e_{wi,t} \\ \varepsilon_t \end{pmatrix} \overset{iid}{\sim} \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \psi\sigma & 0 \\ \psi\sigma & \sigma^2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right)$$

Remarks

- When $|\rho_i| < 1$ the equilibrium error $v_{i,t}$ in (5) is stationary such that $y_{i,t}$ and $x_{i,t}$ are cointegrated with $(1 \;\; \alpha_i \;\; \beta_i)^\top$. Further, ρ_i need not be constant across i. We do, however, choose a common ρ in the simulations to limit the number of experiments and to facilitate the interpretation of the results.

- Solving the system of equations (5) for $x_{i,t}$, we can write

  $$x_{i,t} = \frac{a_1\alpha_i + a_1 v_{i,t} - w_{i,t}}{a_2 - a_1\beta_i}.$$

  $x_{i,t}$ is weakly exogenous when $a_1 = 0$.

- The panel is cross-sectionally dependent because of the common factor ε_t and the idiosyncratic factor loadings $\lambda_i \sim \mathcal{U}[\zeta_1, \zeta_2]$, where $\mathcal{U}[\zeta_1, \zeta_2]$ denotes the uniform distribution with lower bound ζ_1 and upper bound ζ_2.[9] Similar factor structures have been employed by Bai and Ng [2004] and Phillips and Sul [2003]. Even though we generate cross-sectional dependence via a single factor, the validity of the suggested bootstrap tests does not depend on knowledge of the latent factor structure.

- Since $\lambda_i \neq \lambda = \text{cst.}$ it is not possible to remove the cross-sectional dependence by subtracting time-specific means $N^{-1}\sum_i z_{i,t}$, as would be possible under the stronger assumptions of, e.g., Westerlund [2005].

- If $\pi_i \neq 0$ there is a moving-average component in the errors. In particular, values $-1 < \pi_i < 0$ are well-known to have a potentially severe size-distorting effect on unit root and cointegration tests [Schwert, 1989].

[9] Uniform random numbers are generated using the KM algorithm from which Normal variates are created with the fast acceptance-rejection algorithm, both implemented in GAUSS. Part of the calculations are performed with COINT 2.0 by Peter Phillips and Sam Ouliaris.

The dimensions of the panel are $N \in \{10, 30\}$ and, after having discarded 75 initial observations, $T \in \{50, 100\}$. These are representative for the dimensions of data sets often encountered in macroeconometric applications. For a given cross-sectional dimension we draw the unit specific intercepts as $\alpha_i \sim \mathcal{U}[0, 5]$ and keep them fixed for both T. We choose $\beta_i = \mathcal{U}[1, 2]$, $a_2 = -1$ and $\sigma = 1$ and investigate all combinations of the following values for the parameters of the model: $a_1 \in \{0, 1\}$, $\psi \in \{0, 0.5\}$, $\pi \in \{-0.5, 0\}$ and $\rho \in \{0.8, 0.9, 1\}$. Further, there are three different degrees of cross-sectional dependence: $(\zeta_1, \zeta_2) = (0, 0)$, $(\zeta_1, \zeta_2) = (0, 1)$ and $(\zeta_1, \zeta_2) = (1, 4)$, corresponding to no, "weak" and "strong" cross-sectional dependence [see Mark et al., 2005]. That is, 72 experiments are conducted for each of the 4 panel dimensions.

To limit the computational burden, we use $M = 1,000$ replications with $B = 1,000$ bootstrap resamples in each.[10] The p-values are obtained from the Engle and Granger [1987] ADF test, selecting the number of lagged differences J_i in the second stage ADF regressions (3) according to the automatic procedure suggested by Ng and Perron [2001]. We record a rejection if (i) the P_{χ^2} test statistic exceeds the 5% critical value of the χ^2_{2N} distribution, the $P_{\Phi^{-1}}$ test statistic falls below the 5% quantile of the cdf of the standard normal distribution, -1.645, or (ii) for the bootstrap version of the tests, if equation (4) applies.

To summarize, the DGP used here simultaneously addresses several issues that have proved both relevant for empirical work and challenging for unit root and cointegration tests. Hence, while cautioning that interpretation of Monte Carlo results should strictly speaking be confined to the DGP at hand [see Horowitz and Savin, 2000], we are optimistic that tests which perform well under this experimental design may be useful for applied studies.

Table I reports the rejection rates from the Monte Carlo study of the bootstrap and the simple P tests. For brevity, we only give the (representative) results for the serially correlated case ($\pi = -0.5$) and $\psi = 0.5$.[11] In order to illustrate the importance of a suitable lag order q when applying the sieve bootstrap we choose $q = 2$ (columns BS_2) and the data dependent rule $q = [4 \cdot (T/100)^{.25}]$ (columns BS_s). The columns "Sim" refer to the simple P tests.

The main findings may be summarized as follows. The simple P tests overreject in

[10] To gauge the sampling variability of the experiments we perform selected experiments several times. The rejection rates never differed by more than 2 percentage points and usually by less. This suggests the number of replications is sufficient.

[11] The full set of results of the study is available upon request.

TABLE I—REJECTION RATES OF THE P TESTS.

		P_{χ^2}						$P_{\Phi^{-1}}$					
	N	10			30			10			30		
	T	BS_2	BS_s	Sim	BS_2	BS_s	Sim	BS_2	BS_s	Sim	BS_2	BS_s	Sim
(a) $\lambda_i \sim \mathcal{U}[1,4]$							(i) $a_1 = 0$						
$\rho = 1$	50	.068	.049	.090	.092	.072	.141	.111	.081	.143	.127	.113	.193
	100	.099	.056	.102	.114	.082	.166	.125	.091	.142	.156	.119	.233
$\rho = .9$	50	.138	.092	.228	.156	.133	.313	.208	.157	.319	.235	.216	.452
	100	.433	.294	.548	.477	.341	.691	.521	.393	.679	.691	.437	.815
$\rho = .8$	50	.332	.267	.483	.396	.279	.632	.438	.379	.606	.502	.394	.753
	100	.842	.736	.913	.879	.753	.954	.890	.805	.950	.912	.825	.981
							(ii) $a_1 = 1$						
$\rho = 1$	50	.200	.084	.272	.387	.148	.470	.237	.150	.214	.451	.253	.379
	100	.256	.082	.162	.471	.128	.293	.289	.124	.180	.536	.248	.346
$\rho = .9$	50	.185	.101	.165	.369	.160	.272	.224	.152	.096	.443	.307	.191
	100	.273	.104	.078	.536	.143	.142	.327	.152	.076	.658	.334	.149
$\rho = .8$	50	.229	.116	.155	.408	.211	.206	.303	.202	.093	.546	.415	.106
	100	.452	.150	.138	.749	.300	.251	.550	.288	.162	.839	.593	.295
(b) $\lambda_i \sim \mathcal{U}[0,1]$							(i) $a_1 = 0$						
$\rho = 1$	50	.047	.021	.024	.044	.011	.008	.077	.043	.021	.086	.055	.009
	100	.103	.034	.035	.126	.036	.033	.127	.055	.033	.195	.087	.039
$\rho = .9$	50	.157	.090	.090	.328	.208	.174	.326	.250	.131	.627	.540	.297
	100	.763	.484	.599	.968	.873	.924	.903	.786	.774	.995	.980	.978
$\rho = .8$	50	.569	.423	.486	.922	.773	.862	.804	.732	.630	.987	.974	.948
	100	.999	.987	.998	1.00	1.00	1.00	.999	1.00	1.00	1.00	1.00	1.00
							(ii) $a_1 = 1$						
$\rho = 1$	50	.159	.057	.216	.289	.080	.322	.201	.119	.131	.383	.255	.182
	100	.237	.068	.114	.477	.099	.191	.289	.118	.127	.559	.239	.185
$\rho = .9$	50	.219	.113	.170	.404	.192	.275	.315	.229	.132	.590	.450	.188
	100	.469	.232	.244	.905	.551	.537	.626	.455	.308	.974	.867	.695
$\rho = .8$	50	.418	.227	.346	.790	.569	.625	.618	.470	.346	.942	.889	.630
	100	.972	.791	.908	1.00	.989	1.00	.993	.946	.963	1.00	1.00	1.00
(c) $\lambda_i \equiv 0$							(i) $a_1 = 0$						
$\rho = 1$	50	.030	.018	.015	.044	.006	.005	.059	.041	.011	.074	.033	.005
	100	.081	.030	.019	.098	.015	.013	.092	.049	.016	.172	.063	.010
$\rho = .9$	50	.145	.091	.068	.349	.165	.111	.332	.255	.091	.737	.605	.186
	100	.789	.512	.567	.996	.940	.958	.927	.797	.784	1.00	.999	.995
$\rho = .8$	50	.577	.367	.444	.953	.853	.848	.843	.742	.592	1.00	.998	.964
	100	1.00	.995	.998	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
							(ii) $a_1 = 1$						
$\rho = 1$	50	.157	.062	.196	.271	.085	.308	.198	.129	.132	.370	.153	.170
	100	.265	.071	.142	.496	.090	.209	.304	.110	.124	.564	.215	.199
$\rho = .9$	50	.241	.127	.198	.425	.205	.305	.327	.237	.143	.645	.510	.250
	100	.583	.319	.345	.946	.664	.709	.744	.551	.450	.990	.940	.853
$\rho = .8$	50	.481	.317	.405	.863	.649	.752	.680	.535	.435	.983	.947	.784
	100	.994	.890	.969	1.00	1.00	1.00	.999	.990	.995	1.00	1.00	1.00

Note: $\psi = 0.5$, $\pi_i = -0.5$. $M, B = 1,000$ replications and resamples.

Panels (a), (b) and (c) correspond to "strong", "weak" and no correlation.

Panels (i) and (ii) correspond to the exogenous and endogenous case.

BS_2, BS_s and Sim are the sieve bootstrap with $q = 2$, q data dependent, and simple, resp.

the presence of cross-sectional correlation. The strength of the dependence—see the rows $\rho = 1$ in panels (a) and (b)—does matter. The ERPs are particularly severe in the presence of endogenous regressors (see panels (ii) vs. (i)). Under "strong" correlation (panel (a)) and endogeneity ($a_1 = 1$), the simple P tests even seem to be biased as they reject more frequently for samples generated under H_1 than for samples generated under H_0. On the other hand, the difference in the performance of the tests between the uncorrelated case (c) and the weakly correlated case (b) is small, suggesting that correlation robust tests may be most expedient when one suspects strong forms of dependence in the data. Generally, the sieve bootstrap is capable of removing or at least substantially reducing the ERP. Size control usually is more effective in columns BS_s than in columns BS_2. Thus, as expected, it is necessary to suitably prewhiten the residuals with a data-dependent lag order selection scheme for q. This is in line with many other studies in the nonstationary panel literature. For instance, Hlouskova and Wagner [2006] find that selection of the lag length in panel unit root ADF regressions plays an important role for the behavior of many popular tests.

Concerning the power of the tests, consider rows $\rho = 0.9, 0.8$. Power increases with T, as expected. The increase in power with growing N is more pronounced when the dependence in the data is smaller. This is intuitive because under strong dependence, the amount of independent information in the panel is smaller for a given N.

5 An Empirical Application to the PPP Hypothesis

In this section, we reconsider the Purchasing Power Parity (PPP) hypothesis that has attracted wide attention in the literature [see Taylor and Taylor, 2004, for a recent survey]. Assuming the law of one price to hold at least in the long run, its absolute version implies that the ratio of domestic price level $P_{i,t}$ and foreign price level $P^*_{i,t}$ should be close to the exchange rate $S_{i,t}$. Equivalently, the real exchange rate $R_{i,t}$ should be near one. Empirically, denoting natural logs by lowercase letters, the PPP hypothesis therefore postulates that

$$r_{i,t} = p_{i,t} - p^*_{i,t} - s_{i,t} \qquad (i \in \mathbb{N}_N)$$

is a stationary process. Due to factors such as transportation costs, one often allows for non-unitary coefficients to obtain an equation of the form

$$p^*_{i,t} = a_i + \beta_{i1} p_{i,t} + \beta_{i2} s_{i,t} + e_{i,t}, \qquad (i \in \mathbb{N}_N) \tag{6}$$

TABLE II—TESTS FOR WEAK PURCHASING POWER PARITY

Countries	$\hat{a}_i$	$\hat{\beta}_{1i}$	$\hat{\beta}_{2i}$	J_i	t_{ϱ_i-1}
Australia	16.848	0.833	3.982	2	-1.912
Austria	-23.109	1.216	-0.115	2	-2.412
Belgium	-3.907	1.114	-0.205	1	-1.626
Canada	-5.292	0.943	10.491	1	-0.809
Denmark	12.461	0.974	-1.005	0	-0.751
Finland	10.224	0.915	0.455	3	-1.841
France	18.154	0.959	-2.014	1	-0.446
Germany	-40.631	1.425	-3.421	3	-2.778
Greece	38.940	-0.036	0.352	0	-2.273
Ireland	30.309	0.875	-23.705	0	-1.082
Italy	28.505	0.659	0.003	3	-2.815
Japan	6.847	1.098	-0.119	0	-1.222
Netherlands	-21.331	1.290	-6.165	0	-3.010
New Zealand	26.845	0.734	3.613	2	-1.727
Norway	11.707	0.888	0.489	1	-1.500
Portugal	38.443	0.383	0.155	3	-2.821
Spain	27.919	0.657	0.046	1	-2.340
Sweden	14.692	0.766	1.426	3	-2.423
Switzerland	-30.012	1.313	-2.622	0	-1.203
United Kingdom	16.231	0.819	5.693	1	-2.002
United States	numeraire country				

Panel Results		simple	bootstrap critical values for $q =$				
		cr. val	2	4	6	8	10
$\tilde{P}_{\chi^2}$	16.751	55.758	46.018	33.160	38.908	37.408	41.151
$\tilde{P}_{\Phi^{-1}}$	3.483	-1.645	-0.522	1.190	0.383	0.432	-0.052

Note: First 3 columns are estimates of (6), J_i are number of lagged differences chosen for unit root tests on $\hat{e}_{i,t}$. t_{ϱ_i-1} is t-statistic from (3). Bottom part gives realized test statistics (left), simple (middle) and bootstrap (right) 5% critical values for different AR orders q in the sieve bootstrap procedure.

referred to as the weak PPP hypothesis. Since the variables typically are $I(1)$, the weak PPP hypothesis implies that $e_{i,t}$ is stationary, or, equivalently, that $p^*_{i,t}$, $p_{i,t}$ and $s_{i,t}$ are cointegrated.

We use quarterly Post-Bretton Woods data for a panel of OECD countries, observing the spot exchange rate and the consumer price index ranging from 1973:1 until 1998:2.[12] The numeraire country (with respect to which the spot exchange rate is sampled) is the United States. As has been pointed by, inter alia, O'Connell [1998], the choice of a numeraire country naturally induces a common component in the set of regressions (6). It is therefore

[12]The data is from the IMF Financial Statistics CD-ROM. The list of included countries can be found in Table II.

interesting to investigate the effect of using cross-sectional correlation robust tests of the PPP hypothesis. Table II reports the estimation results. The first three columns give the estimates of a_i, β_{i1} and β_{i2}. The results differ across countries, making it attractive to use a panel procedure that allows for heterogeneity. The number of lagged differences J_i chosen by the procedure of Ng and Perron [2001] is reported in column 4. The fifth column gives the results for the t-statistics of the individual ADF cointegration regressions (3). All statistics fail to produce evidence in favor of PPP, clearly exceeding the critical value of -3.837. The Netherlands have the smallest statistic $t_{\varrho_{\text{Neth}-1}} = -3.010$, implying a p-value of 0.251 [MacKinnon, 1996]. The estimates of the panel tests are reported in the bottom of Table II. In view of the individual outcomes, it is not surprising to see that the simple and bootstrap p-value combination tests also do not reject the null hypothesis of no PPP.

Even though neither tests rejects the null, the critical values differ substantially. In particular, the bootstrap critical values are almost all closer to the realized test statistic, indicating that using the critical values of the simple tests would be overly conservative in the present application. We conclude that using cross-sectional correlation robust tests may well make an important difference in econometric practice.

6 Conclusion

We suggest new meta analytic p-value combination tests for panel cointegration. The tests are robust to cross-sectional dependence of an arbitrary form. To achieve robustness, a bootstrap procedure is used. Further advantages compared to other widely used panel cointegration tests are that they are flexible, intuitively appealing and comparatively easy to compute.

The simulation study reveals that the approach is generally effective in improving the reliability of the "simple" P tests. However, care is needed in properly selecting the lag order in the sieve bootstrap scheme. This finding is analogous to the well-known result that the size and power of time series unit root and cointegration tests are fairly sensitive to the selection of the correct number of lagged differences in the ADF regressions (or, selection of lag truncations in estimating long-run variances).

A further interesting finding is that the power gains from using panel tests are the smaller the stronger the cross-sectional correlation in the data. We interpret this as reflecting

that combining correlated time series leads to less additional information than combining independent series. This could imply that the impressive power gains reported in simulations of first generation panel unit root tests vs. time series unit root tests are partially an artefact of the experimental design relying on independence of the units. Our empirical study of the weak PPP hypothesis reveals that cross-sectional correlation robust critical values for the P tests can differ substantially from their simple counterparts. It would be interesting to investigate how other commonly used panel cointegration tests such as those developed by Pedroni [2004] or Kao [1999] perform under general forms of cross-sectional correlation.

Another task for future research is to establish whether the bootstrap procedure suggested in this paper is consistent in the sense of leading to the same limiting distributions for the bootstrap statistics as for the original ones. It may be possible to establish bootstrap consistency along the lines of Chang and Park [2003], Chang et al. [2006] and Swensen [2003].

References

Bai, J. and S. Ng: 2004, 'A PANIC Attack on Unit Roots and Cointegration'. *Econometrica* **72**(4), 1127–1177.

Banerjee, A., M. Marcellino, and C. Osbat: 2005, 'Testing for PPP: Should We Use Panel Methods?'. *Empirical Economics* **30**(1), 77–91.

Basawa, I., A. Mallik, W. McCormick, J. Reeves, and R. Taylor: 1991, 'Bootstrapping Unstable First-Order Autoregressive Processes'. *Annals of Statistics* **19**(2), 1098–1101.

Bickel, P. J. and K. A. Doksum: 2001, *Mathematical Statistics*, Vol. 1. Upper Saddle River, New Jersey: Prentice Hall, 2nd edition.

Bierens, H. J.: 2005, *Introduction to the Mathematical and Statistical Foundations of Econometrics*. Cambridge: Cambridge University Press.

Breitung, J. and S. Das: 2005, 'Panel Unit Root Tests Under Cross Sectional Dependence'. *Statistica Neerlandica* **59**(4), 414–433.

Brockwell, P. J. and R. A. Davis: 1991, *Time Series: Theory and Methods*. New York: Springer, 2nd edition.

Chang, Y. and J. Y. Park: 2003, 'A Sieve Bootstrap for the Test of a Unit Root'. *Journal of Time Series Analysis* **24**(4), 379–400.

Chang, Y., J. Y. Park, and K. Song: 2006, 'Bootstrapping Cointegrating Regressions'. *Journal of Econometrics* (forthcoming).

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

Davidson, R. and J. G. MacKinnon: 1999, 'The Size Distortion of Bootstrap Tests'. *Econometric Theory* **15**, 361–376.

Engle, R. and C. Granger: 1987, 'Co-Integration and Error Correction: Representation, Estimation, and Testing'. *Econometrica* **55**(2), 251–76.

Fisher, R.: 1970, *Statistical Methods for Research Workers*. London: Oliver and Boyd, 14th edition.

Hanck, C.: 2005, 'A Meta Analytic Approach to Testing for Panel Cointegration'. Technical report, RGS.

Hedges, L. and I. Olkin: 1985, *Statistical Methods for Meta-Analysis*. San Diego: Academic Press.

Hlouskova, J. and M. Wagner: 2006, 'The Performance of Panel Unit Root and Stationarity Tests: Results from a Large Scale Simulation Study'. *Econometric Reviews* **25**(1), 85–116.

Horowitz, J. L. and N. Savin: 2000, 'Empirically Relevant Critical Values for Hypothesis Tests: A Bootstrap Approach'. *Journal of Econometrics* **95**, 375–389.

Hsiao, C.: 2003, *Analysis of Panel Data*. New York: Cambridge University Press, second edition.

Johansen, S.: 1988, 'Statistical Analysis of Cointegration Vectors'. *Journal of Economic Dynamics and Control* **12**(2–3), 231–254.

Kao, C.: 1999, 'Spurious Regression and Residual Based Tests for Cointegration in Panel Data'. *Journal of Econometrics* **90**(1), 1–44.

Künsch, H. R.: 1989, 'The Jackknife and the Bootstrap for General Stationary Observations'. *Annals of Statistics* **17**, 1217–1241.

Larsson, R., J. Lyhagen, and M. Löthgren: 2001, 'Likelihood-based Cointegration tests in heterogeneous panels'. *The Econometrics Journal* **4**(1), 109–142.

MacKinnon, J. G.: 1996, 'Numerical Distribution Functions for Unit Root and Cointegration Tests'. *Journal of Applied Econometrics* **11**(6), 601–618.

Maddala, G. and S. Wu: 1999, 'A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test'. *Oxford Bulletin of Economics and Statistics* **61**(S1), 631–652.

Mark, N. C., M. Ogaki, and D. Sul: 2005, 'Dynamic Seemingly Unrelated Cointegrating Regression'. *Review of Economic Studies* **72**(3), 797–820.

McCoskey, S. and C. Kao: 1998, 'A Residual Based Test of the Null of Cointegration in Panel Data'. *Econometric Reviews* **17**(1), 57–84.

Moon, H. R. and B. Perron: 2004, 'Testing for a unit root in panels with dynamic factors'. *Journal of Econometrics* **122**, 81–126.

Ng, S. and P. Perron: 2001, 'Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power'. *Econometrica* **69**(6), 1519–1554.

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

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**(3), 597–625.

Phillips, P. and D. Sul: 2003, 'Dynamic Panel Estimation and Homogeneity Testing under Cross Section Dependence'. *Econometrics Journal* **6**, 217–259.

Phillips, P. C. and S. Ouliaris: 1990, 'Asymptotic Properties of Residual Based Tests for Cointegration'. *Econometrica* **58**(1), 165–193.

Schwert, G.: 1989, 'Tests for Unit Roots: A Monte Carlo Investigation'. *Journal of Business & Economic Statistics* **7**(2), 147–159.

Shin, Y.: 1994, 'A Residual-Based Test of the Null of Cointegration against the Alternative of No Cointegration'. *Econometric Theory* **10**(1), 91–115.

Spanos, A.: 1986, *Statistical Foundations of Econometric Modelling*. New York: Cambridge University Press.

Swensen, A. R.: 2003, 'Bootstrapping Unit Root Tests for Integrated Processes'. *Journal of Time Series Analysis* **24**(1), 99–126.

Taylor, A. and M. P. Taylor: 2004, 'The Purchasing Power Debate'. *Journal of Economic Perspectives* **18**(4), 135–158.

Westerlund, J.: 2005, 'New Simple Tests for Panel Cointegration'. *Econometric Reviews* **24**(3), 297–316.