EUROPEAN UNIVERSITY INSTITUTE
DEPARTMENT OF ECONOMICS

EUI Working Paper **ECO** No. 2003/24

On the Use of Panel Unit Root Tests on Cross-Sectionally Dependent Data: An Application to PPP

FABIAN BORNHORST

BADIA FIESOLANA, SAN DOMENICO (FI)

Published in Italy in November 2003
European University Institute
Badia Fiesolana
I-50016 San Domenico (FI)
Italy

On the use of panel unit root tests on cross-sectionally dependent data: an application to PPP*

Fabian Bornhorst¶

November 13, 2003

A Monte Carlo exercise demonstrates the different size distortions that two of the most commonly used panel unit root tests have when the sections of the panel are affected by correlated errors, when they are cointegrated, or both. For a specific form of sectional correlation, the limiting distribution is derived and asymptotic normality of the test statistic is established. To determine the nature of contemporaneous cross-sectional correlation in real data, covariance matrix estimation techniques are discussed and an appropriate bootstrap method for the estimation of standard errors is suggested. In an application to a panel of real exchange rates it is found that both aforementioned dependencies are present, and therefore the results of panel unit root tests – if applied at all – should be interpreted accordingly.

JEL classification: F31, C15, C23
Keywords: panel data, nonstationarity, cross-sectional dependence, PPP

*I would like to thank Anindya Banerjee, Helmut Luetkepohl, Juergen Maurer, Hannah Nielsen, Roland Straub and seminar participants at the European University Institute for helpful comments and suggestions. Thanks also to Christopher Baum from Boston College for co-authoring STATA-programs used in this paper. Needless to say, I take the sole responsibility for all remaining errors and shortcomings.
¶Department of Economics, European University Institute, Villa San Paolo, Via della Piazzuola 43, I–50133 Florence, Italy. Email: Fabian.Bornhorst@iue.it

1. Introduction

Panel unit root tests are becoming a standard tool in the analysis of mostly macroeconomic panels. Two procedures, the Levin, Lin and Chu (2002)[1] and the Im, Pesaran and Shin (1997) test for unit roots are among the most popular. The tests have been applied to a range of macroeconomic problems, *e.g.* to the question whether real exchange rates are random walk processes or not (*e.g.* O'Connell (1998), Papell (1997)) or to investigate the mean reversion properties of the current account (Wu, 2000). Evans and Karras (1996) use panel unit root tests to analyze the convergence of regions in the US using a modified Levin et al. (2002) test procedure, while Strauss (2000) addresses the question of permanent components in regional GDP using these panel unit root tests.

However, relatively little is known about the size and power properties of these tests when any of the distributional assumptions underlying their construction is violated. The asymptotic distribution of both test statistics relies on the independence of the sections of the panel. This assumption might often be violated in real data, especially in a macroeconomic context. Given their widespread use, it is important to know more about the reliability of the test results. The impact of such dependence on the performance of the tests is studied in this paper. Two different forms of sectional dependence are considered. In the short run (Section 3.1), positive cross-sectional dependence of the error terms is analyzed. It is found that in the case of *common* shocks, eliminating common time effects is remedy enough to restore the size properties reasonably well. In fact, the test statistic does converge to a standard normal distribution. In this respect this paper contrasts the finding of O'Connell (1998), who attests severe size distortions to the Levin et al. (2002) test in the presence of common contemporaneous correlation. When the contemporaneous correlation takes different forms, however, severe size distortions do occur. Long-run sectional dependence might be present if the series of the panel are cointegrated (Section 3.2). In this case, the series are nonstationary but share a common stochastic trend. Early work on the study of this effect on panel unit root tests has been done by Crowder (1997) in a simple cointegration framework. The effect cointegration has on unit root test is analytically studied in Lyhagen (2000). However, data generating process considered in this paper resembles more the one considered in Banerjee et al. (2000). In line with the results of these studies, it is found that the tests are oversized as a consequence of cointegration, as long as the errors are kept independent. In Section 3.3 cointegration is combined with sectional correlation, *i.e.* long and short run dependence are brought together. This seems natural as there is no prior reason to believe that these phenom-

[1] A previous version of this test was known as Levin and Lin (1993). See also Levin and Lin (1992).

ena should be mutually exclusive. The result is surprising. Considered separately, long and short run dependencies tend to yield oversized test results. If brought together, under some parameter configurations the size distortions go in the opposite direction: the over–rejection of the null hypothesis of a unit root vanishes and the tests become undersized. As a result, without further knowledge about the data generating process, panel unit root tests in presence of sectional dependence are inconclusive.

The application in Section 4 contributes to the purchasing power parity debate by addressing the question of mean reversion in a panel of real exchange rates. A panel of 18 exchange rates is first analyzed by estimating the contemporaneous covariance matrix of the error terms and corresponding standard errors. Different ways of estimating covariance matrices in the presence of heteroscedasticity and serial correlation are discussed and a bootstrap algorithm developed by Politis and Romano (1994) is suggested as a way of obtaining standard errors for these estimates. To know whether long-run sectional dependence is present in the data, a cointegration analysis following Johansen (1995) on a subset of exchange rates is conducted. Together with the simulation results obtained earlier, the existence of both dependencies in the data puts a big caveat on the use of panel unit root tests in this context in particular, and on cross-sectionally dependent data in general.

2. Panel Unit Root Tests

The test developed by Levin et al. (2002) (henceforth LLC) can be seen as a natural extension of the Dickey and Fuller (1981) test for a unit root to a set of time series. It builds on the method previously suggested by Quah (1990) and Breitung and Meyer (1991). In the light of the criticism by Pesaran and Smith (1995) of the use of pooled regressions of the LLC type, Im et al. (1997) (henceforth IPS) allow for heterogeneity of the series under the alternative and do not make use of traditional panel estimation techniques. They propose instead a group-mean Lagrange multiplier test and a group mean t–test based on the individual ADF test statistics. The asymptotic properties for both tests are derived by assuming a diagonal path limit. The behaviour of the cross-section dimension (N) and the time dimension (T) are functionally tied, *i.e.* $(T(N), N \to \infty)$. For LLC, as both go to infinity, T increases faster than N, such that $N/T \to 0$, whereas IPS only require $\sqrt{N}/T \to 0$.

This section presents the framework for the analysis of panel unit root tests. As in the univariate case, three forms of deterministics are considered starting from the following data generating process (DGP) that yields nonstationary series if the autoregressive coefficient ρ_i is equal to one:

$$\Delta x_{it} = (\rho_i - 1)x_{it-1} + \mu_i + \beta_i \cdot t + \epsilon_{it}.$$

The index i indicates the section of the panel ($i = 1, ..., N$) and the time index t ranges from 1 to T. The constant of each section is denoted by μ_i and $\beta_i \cdot t$ represents a time trend in the data. The assumptions on the error term ϵ_{it} are discussed further below. Table 1 summarizes the *a priori* restrictions and the hypothesis to be tested in each of the three models. The most general specification, model $m = 3$ in the classification of LLC, is designed to discriminate between a set of I(1) processes with drift under the null and a set of trendstationary processes under the alternative. In model 2, the trend parameter is restricted to zero *a priori*. It is

Table 1: Different models and hypothesis

model	*a priori*	Null-hypothesis and alternative
m= 3		$H_0^{(3)} : \rho_i = 1 \ \forall\, i \ (\Rightarrow \mu_i \neq 0, \beta_i = 0)$
		$H_1^{(3)} :\mid \rho_i \mid< 1 \ \forall\, i \ (\Rightarrow \beta_i \neq 0)$
m= 2	$\beta_i = 0$	$H_0^{(2)} : \rho_i = 1 \,\forall\, i \ (\Rightarrow \mu_i = 0)$
		$H_1^{(2)} :\mid \rho_i \mid< 0 \,\forall\, i \ (\Rightarrow \mu_i \neq 0)$
m= 1	$\beta_i = 0$	$H_0^{(1)} : \rho_i = 0 \ \forall\, i$
	$\mu_i = 0$	$H_1^{(1)} :\mid \rho_i \mid< 0 \ \forall\, i$

used to discriminate between a set of I(1) processes without drift under the null and allows stationary processes with an expected value different from zero under the alternative. This model will be used throughout the Monte Carlo study. In the simplest model, under the null hypothesis of a unit root, x_{it} is a set of I(1) processes without drift, while under the alternative it is a set of stationary processes all with an expected value of zero.

2.1. Levin, Lin and Chu (2002)

The LLC test is implemented in four steps.

Step 1: Elimination of time specific effects. The cross-section average at t is subtracted from the data, *i.e.* $x_{it} = \tilde{x}_{it} - \frac{1}{N}\Sigma_{i=1}^{N}\tilde{x}_{it}$, which is equivalent to the introduction of time specific dummy variables. This step will play a crucial role in the simulation exercise in Section 3.

Step 2: Computation of ADF-statistics and normalized residuals. The choice of the lags L_i to be included should be based on a common information criterion (*e.g.* Akaike or Schwartz) and done *after* the elimination of time specific effects. Instead of the usual equation:

$$\Delta x_{it} = \delta_i x_{it-1} + \sum_{j=1}^{L_i} \theta_{ij} \Delta x_{it-j} + \mu_i + \beta_i t + \epsilon_{it},$$

the coefficient of interest, δ_i, is estimated by partitioning the regression using the Frisch–Waugh theorem to obtain residuals from each step:

$$\Delta x_{it} = \sum_{j=1}^{L_i} \theta_{ij}^{(1)} \Delta x_{it-j} + \mu_i^{(1)} + \beta_i^{(1)} t + e_{it} \Rightarrow \hat{e}_{it}$$
$$x_{it-1} = \sum_{j=1}^{L_i} \theta_{ij}^{(2)} \Delta x_{it-j} + \mu_i^{(2)} + \beta_i^{(2)} t + v_{it-1} \Rightarrow \hat{v}_{it-1}.$$

The regression of the residuals gives an estimator for δ_i:

$$\hat{e}_{it} = \delta_i \hat{v}_{it-1} + \epsilon_{it}. \tag{1}$$

In order to control for heterogeneity in the variances of the series, the residuals are normalized by the standard error σ_{ei} of regression (1), estimated by:

$$\hat{\sigma}_{ei}^2 = \frac{1}{T - L_i - 1} \sum_{t=L_i+2}^{T} \left(\hat{e}_{it} - \hat{\delta}_i \hat{v}_{it-1} \right)^2,$$

and the normalization is done as follows:

$$\tilde{e}_{it} = \frac{\hat{e}_{it}}{\hat{\sigma}_{ei}} \qquad \text{and} \qquad \tilde{v}_{it-1} = \frac{\hat{v}_{it-1}}{\hat{\sigma}_{ei}}.$$

Step 3: Computation of the long-run variance. For each series the long-run variance is computed using the first differences:

$$\hat{\sigma}_{xi}^2 = \frac{1}{T-1} \sum_{t=2}^{T} \Delta x_{it}^2 + 2 \sum_{\tau=1}^{K} w_{K\tau} \left(\frac{1}{T-1} \sum_{t=2+\tau}^{T} \Delta x_{it} \Delta x_{it-\tau} \right). \tag{2}$$

The choice of covariance weights ensures positive estimates of the long-run variances. LLC suggest the Bartlett weights, $w_{K\tau} = 1 - \tau/(K+1)$. The estimate is consistent if the truncation parameter K grows exponentially at a rate less than T, LLC suggest $K = 3.21T^{1/3}$. The ratio of the estimated long-run variation and the standard deviation is computed, which under the null approaches one. For the adjustment, the average of this ratio across sections is also needed:[2]

$$\hat{s}_i = \frac{\hat{\sigma}_{xi}}{\hat{\sigma}_{ei}} \qquad \text{and} \qquad \hat{S}_N = \frac{1}{N}\sum_{i=1}^{N} \hat{s}_i.$$

Step 4: Computation of the test statistic. Under the null hypothesis the normalized residuals $\tilde{e}_{it}$ are independent of the normalized lagged residuals $\tilde{v}_{it-1}$. This is estimated using OLS:

$$\tilde{e}_{it} = \delta \tilde{v}_{it-1} + \tilde{\epsilon}_{it}. \tag{3}$$

Under the null hypothesis and in model 1, the regression t–statistic t_δ is asymptotically normal, but has to be adjusted in models 2 and 3, so that, in general:

$$t_\delta^* = \frac{t_\delta - N\,\tilde{T}\,\hat{S}_N\,\hat{\sigma}_\epsilon^{-2}\,\mathrm{SE}(\hat{\delta})\,\mu^*_{m\tilde{T}}}{\sigma^*_{m\tilde{T}}} \overset{H_0^{(m)}}{\sim} N(0,1),$$

where $\mathrm{SE}(\hat{\delta})$ is the standard error of $\hat{\delta}$, $\hat{\sigma}_\epsilon$ is the standard error of the regression (3), $\mu^*_{m\tilde{T}}$ and $\sigma^*_{m\tilde{T}}$ are necessary adjustments for the mean and the standard deviation. These vary according to m, the model chosen and $\tilde{T}$, the average number of observations per section in the panel adjusting for lagged differences, $\tilde{T} = T - \frac{1}{N}\Sigma_{i=1}^{N} L_i$ (see Table 2 in LLC).

The asymptotic properties are derived in Levin and Lin (1993, Section 4).[3] In model specifications 2 and 3 the estimator $\hat{\delta}$ has a downward bias, which is due to the dynamic specification of the panel, especially for small T and N (Nickell, 1981). This makes the mean adjustments necessary. Furthermore, under the null, the variance of the estimator $\hat{\delta}$ falls at the rate $\frac{1}{NT^2}$, reflecting super–consistency. As N grows large, the variance of $\hat{\delta}$ gets smaller and smaller, which makes the variance adjustment necessary. If not adjusted, mean and variance bias would force the t–value to negative infinity in models 2 and 3. Under the alternative, x_{it} is already stationary, so Δx_{it} has asymptotically zero variation at zero frequency, meaning that each standard deviation ratio s_i as well as the average ratio $\hat{S}_N$ becomes small. In this case the mean adjustment does not influence the t–value adjustment, so that the adjusted value diverges to negative infinity. This shows the advantage of using an estimate of the long-run variance to discriminate between stationary and nonstationary processes.

[2] In the case of a trend the steps above should be implemented after demeaning the differenced series.
[3] See also page 26 in the Appendix for a detailed treatment of the asymptotic properties in the case $m = 1$.

2.2. Im, Pesaran and Shin (1997)

The IPS test extends the LLC framework by allowing for a mixture of stationary and nonstationary series under the alternative hypothesis. The test is defined for models 2 and 3, and the alternative is modified to:

$$H_1^{(IPS)} = \rho_i < 0, \; \forall\, i = 1, 2, ...N_1, \rho_i = 0, \; \forall\, i = N_1 + 1, ..., N.$$

IPS suggest a group mean lagrange multiplier (LM) test and a group mean t–test based on the individual ADF t–values. In simulations done by the authors the t–test outperforms the LM test slightly. According to the ADF lag order chosen in each section and the length T, adjustments are necessary to the mean and variance. The test statistics becomes:

$$\Psi_{\bar{t}} = \frac{\sqrt{N}\Big\{\bar{t}_{N,T} - \frac{1}{N}\sum_{i=1}^{N} E[t_{i,T}(L_i, 0) \mid \rho_i = 0]\Big\}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N} Var[t_{i,T}(L_i, 0) \mid \rho_i = 0]}} \overset{H_0^{(IPS)}}{\sim} N(0, 1).$$

The adjustments $E[...]$ and $Var[...]$ are tabulated in the paper. The expression $\bar{t}_{N,T} = \frac{1}{N}\sum_{i=1}^{N} t_{i,T}(L_i, \theta_i)$ is the mean of the actual ADF test statistics. IPS also suggest the inclusion of time specific effects in the regression or, alternatively, the demeaning of the panel at each t. Note, however, that in contrast to LLC, the IPS–test uses an average of t–statistics and not a single estimated t-value from the pooled series.

3. The effect of cross-sectional dependence

The model considered in this paper is designed to discriminate between a set of I(1) series without drift and a set of AR(ρ) series with expectation different from zero. In terms of standard macroeconomic time series, this configuration refers to, for example, interest rates, exchange rates and possibly price indices. The DGP takes the following form:

$$\begin{pmatrix} \Delta \mathbf{x}_t \\ \Delta \mathbf{y}_t \end{pmatrix} = \mathbf{A}\mathbf{B}' \begin{pmatrix} \mathbf{x}_{t-1} \\ \mathbf{y}_{t-1} \end{pmatrix} + \boldsymbol{\mu} + \boldsymbol{\epsilon}_t. \tag{4}$$

Both $\mathbf{x}_t$ and $\mathbf{y}_t$ are $(N \times 1)$ vectors, $\boldsymbol{\mu}$ and $\boldsymbol{\epsilon}_t$ are $(2N \times 1)$ vectors. The vector of interest is always $\mathbf{x}_t$. The $\mathbf{y}_t$ are used to simulate potentially shared stochastic trends if desired. The matrices $\mathbf{A}$ and $\mathbf{B}$ determine the long-run relation between the variables and will be defined according to the set of experiments. For example, if

$$\mathbf{A} = \alpha \begin{pmatrix} -\mathbf{I} & 0 \\ 0 & -\mathbf{I} \end{pmatrix}, \mathbf{B}' = \begin{pmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{I} \end{pmatrix} \text{ and } \boldsymbol{\mu} = \mathbf{0},$$

and $\alpha = 0$, then $\mathbf{x}_t$ will be a set of independent I(1) variables without drift. The short run correlation is modeled through the error structure:

$$\boldsymbol{\epsilon}_t \sim N(0, \sigma^2 \boldsymbol{\Omega}) \text{ and } \boldsymbol{\Omega} = \begin{pmatrix} \boldsymbol{\Sigma} & 0 \\ 0 & \mathbf{I} \end{pmatrix}.$$

Note that contemporaneous correlation only affects the vector $\mathbf{x}_t$, not $\mathbf{y}_t$. The innovation variance is chosen to be $\sigma^2 = 1$ throughout the paper. In general, the correlation matrix $\boldsymbol{\Sigma}$ takes the following form:[4]

$$E[\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t] = \underset{N \times N}{\boldsymbol{\Sigma}} = \begin{pmatrix} 1 & & & & \\ \omega_{21} & 1 & & & \\ \omega_{31} & \omega_{32} & 1 & & \\ & & & \cdots & \\ & & & & 1 \end{pmatrix}, \tag{5}$$

where the correlations are $| \omega_{ij} | < 1$.

3.1. Cross-sectional Correlation

The first set of experiments is designed to measure the impact of cross-sectional correlation. The absence of error correlation ($\omega_{ij} = 0, \ \forall i, j$) produces the desired size properties, see Table A.1 in the Appendix. Once a common, positive sectional correlation is introduced ($\omega_{ij} = 0.7, \ \forall i, j$), the tests appears to be slightly oversized (Table 2), especially for small N. This contrasts sharply the findings of O'Connell (1998) who finds size distortions of as much as 50 % for the 5 % size. Such distortions can be reproduced if step one of the LLC test, *i.e.* the elimination of common time effects, is not carried out. The results for different values of ω are presented in Tables 3 and 4 for the LLC and IPS test respectively. The power of the LLC test and the IPS test was analyzed for the two alternatives $\rho = .9$ and $\rho = .95$, where $\rho = 1 - \alpha$. This exercise was repeated for varying covariance structures, $\omega = \{0, 0.7, 0.8, 0.9\}$, $N = 25$, $T = \{60, 100\}$, and $\boldsymbol{\mu} = \mathbf{1}$. The results of this analysis are reported in Table 3 for the LLC test and in Table 4 for the IPS test. They show that once common time effects are eliminated, the power of the tests is not severely affected by cross-sectional correlation. More interestingly, the distortions in power and size are independent of the degree of cross-sectional dependence. The natural question that arises is why demeaning, or, equivalently, the inclusion of time dummies, seems to be such an effective instrument if errors are correlated in the way studied here. The expected value of the outer product of the

[4]Considering only the first N elements of $\boldsymbol{\epsilon}_t$.

Table 2: Size properties with common shocks, eliminating common time effects

$\omega = 0.7$	LLC			IPS		
	nominal size 10%					
	$N = 5$	$N = 10$	$N = 20$	$N = 5$	$N = 10$	$N = 20$
$T = 25$	.128	.115	.111	.116	.100	.104
$T = 50$	.127	.112	.106	.117	.108	.102
$T = 100$	.115	.110	.106	.110	.108	.101
	nominal size 5%					
	$N = 5$	$N = 10$	$N = 20$	$N = 5$	$N = 10$	$N = 20$
$T = 25$	.065	.052	.059	.064	.052	.058
$T = 50$	.067	.054	.049	.068	.053	.050
$T = 100$	.060	.056	.055	.054	.058	.055
	nominal size 1%					
	$N = 5$	$N = 10$	$N = 20$	$N = 5$	$N = 10$	$N = 20$
$T = 25$	.014	.009	.013	.016	.010	.015
$T = 50$	.015	.011	.011	.016	.012	.013
$T = 100$	.011	.012	.014	.014	.015	.011

Note: Based on 4,000 replications. The values reported are the percentage of rejections using the indicated nominal level. Ideally, real and nominal size should be equal.

error terms is (considering the relevant first N elements) $E[\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t'] = \boldsymbol{\Sigma}$. The elimination of time effects can be rewritten as:

$$(\boldsymbol{\epsilon}_t - \bar{\boldsymbol{\epsilon}}_t)(\boldsymbol{\epsilon}_t - \bar{\boldsymbol{\epsilon}}_t)' = \left[\left(\mathbf{I} - \frac{\mathbf{ll}'}{N}\right)\boldsymbol{\epsilon}_t\right]\left[\left(\mathbf{I} - \frac{\mathbf{ll}'}{N}\right)\boldsymbol{\epsilon}_t\right]' \text{ where } \mathbf{l} = (1, ..., 1)$$
$$\mathbf{Q}\boldsymbol{\epsilon}_t(\mathbf{Q}\boldsymbol{\epsilon}_t)' = \mathbf{Q}\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t'\mathbf{Q}' = \mathbf{Q}\boldsymbol{\Sigma}\mathbf{Q}'.$$

where $\mathbf{Q}$ is:

$$\underset{N\times N}{\mathbf{Q}} = \begin{pmatrix} 1-\frac{1}{N} & -\frac{1}{N} & ... & -\frac{1}{N} \\ & 1-\frac{1}{N} & ... & -\frac{1}{N} \\ & & ... & \\ & & & 1-\frac{1}{N} \end{pmatrix} = \frac{1}{N}\begin{pmatrix} N-1 & -1 & ... & -1 \\ & N-1 & ... & -1 \\ & & ... & \\ & & & N-1 \end{pmatrix}.$$

If $\boldsymbol{\Sigma}$ takes the form where all off–diagonal elements are equal to ω, the above expression further simplifies to:

$$\mathbf{Q}\boldsymbol{\Sigma}\mathbf{Q}' = \frac{1-\omega}{N}\begin{pmatrix} N-1 & -1 & ... & -1 \\ & N-1 & ... & -1 \\ & & ... & \\ & & & N-1 \end{pmatrix} =$$

$$= (1-\omega)\frac{N-1}{N}\begin{pmatrix} 1 & \frac{-1}{N-1} & ... & \frac{-1}{N-1} \\ & 1 & ... & \frac{-1}{N-1} \\ & & ... & \\ & & & 1 \end{pmatrix} \underset{N\to\infty}{=} (1-\omega)\mathbf{I}. \quad (6)$$

Table 3: Size and power properties of LLC for varying ω and ρ

$N = 25$	nom. size			power					
				$\rho = .9$			$\rho = .95$		
$\omega = 0$	10%	5%	1%	10%	5%	1%	10%	5%	1%
$T = 60$	.09	.04	.005	.99	.98	.85	.56	.38	.13
$T = 100$	.01	.04	.002	1	1	1	.83	.68	.31
$\omega = 0.7$	10%	5%	1%	10%	5%	1%	10%	5%	1%
$T = 60$	.10	.04	.005	.99	.98	.84	.54	.37	.12
$T = 100$	.11	.04	.009	1	1	.99	.79	.64	.29
$\omega = 0.8$	10%	5%	1%	10%	5%	1%	10%	5%	1%
$T = 60$	.10	.04	.005	.98	.96	.80	.54	.36	.10
$T = 100$	.11	.05	.011	1	1	1	.80	.60	.24
$\omega = 0.9$	10%	5%	1%	10%	5%	1%	10%	5%	1%
$T = 60$	.10	.04	.005	.99	.97	.82	.54	.36	.12
$T = 100$	.11	.04	.008	1	1	1	.79	.64	.28

Note: Based on 2,000 replications. One minus the power is the probability that the test fails to reject the null if it is false for a given significance level.

After demeaning, the degree of cross-sectional correlation (the value of ω) leaves the relation of the off–diagonal to the diagonal elements unchanged, but it is this relation which determines the degree to which independence is violated. It is therefore not surprising that the LLC and IPS test do not show significant differences in power and size for varying ω. Moreover, for reasonable large N, the off–diagonal entries are small, *e.g.* with $N = 20$ the remaining 'effective' correlation is -0.05. For large N this approaches zero, just as it is in the absence of any cross correlation. In fact, as shown in Appendix A.1, the test statistic approaches a standard normal distribution.

This argument is limited, however, to the special form of $\mathbf{\Sigma}$ where all $\omega_{ij} = \omega$. Because this might not always be the case, the correlation matrix is now chosen to be a band matrix, where the correlation coefficient decreases with the distance from the main diagonal. The idea behind this specification is that there might be some natural ordering of the sections, reflecting *e.g.* the geographical distribution of units in a spatial model. Errors are more correlated the closer two sections are:

$$E[\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t] = \underset{N\times N}{\mathbf{\Sigma}} = \begin{pmatrix} 1 & \omega^1 & \omega^2 & \omega^3 & \dots & \omega^{N-1} \\ & 1 & \omega^1 & \omega^2 & \dots & \omega^{N-2} \\ & & 1 & \omega^1 & \dots & \omega^{N-3} \\ & & & 1 & \dots & \\ & & & & \dots & \omega^1 \\ & & & & & 1 \end{pmatrix}. \tag{7}$$

Table 4: Size and power properties of IPS for varying ω and ρ

$N = 25$	nom. size			power					
				$\rho = .9$			$\rho = .95$		
$\omega = 0$	10%	5%	1%	10%	5%	1%	10%	5%	1%
$T = 60$	.11	.05	.013	1	1	.99	.86	.76	.46
$T = 100$	.11	.05	.010	1	1	1	1	.99	.95
$\omega = 0.7$	10%	5%	1%	10%	5%	1%	10%	5%	1%
$T = 60$	.12	.07	.010	1	1	.99	.87	.77	.47
$T = 100$	.12	.05	.008	1	1	1	1	.99	.93
$\omega = 0.8$	10%	5%	1%	10%	5%	1%	10%	5%	1%
$T = 60$	.12	.07	010	1	1	.99	.89	.78	.47
$T = 100$	.12	.05	.006	1	1	1	1	.99	.93
$\omega = 0.9$	10%	5%	1%	10%	5%	1%	10%	5%	1%
$T = 60$	.12	.07	.010	1	1	.99	.87	.77	.46
$T = 100$	.11	.05	.009	1	1	1	1	.99	.94

Note: See Table 3.

Table 5 reports the effect of this disturbance has on the performance of the LLC test, given $\omega = 0.7$ and varying N and T. It shows that the test performs quite poorly. Increasing N seem to worsen the results.

Table 5: Size properties of LLC with errors as in (7)

	nom. size 10%		nom. size 5%		nom. size 1%	
$\omega^i = 0.7^i$	$N = 10$	$N = 25$	$N = 10$	$N = 25$	$N = 10$	$N = 25$
$T = 20$	.227	.235	.156	.170	.066	.080
$T = 60$	.249	.250	.170	.180	.064	.078
$T = 100$	.258	.252	.177	.180	.068	.084

Note: based on 10,000 replications.

Short run correlation of this type does affect the size properties, no matter if common time effects are eliminated or not. In the case of common effects, the distortions are far less worrisome than previously claimed.

3.2. Cross-Sectional Cointegration

There are several parameters that influence the specific form of cointegration that one can observe in a vector of time series. One aspect is the number of cointegrating vectors (CIVs) in a system, or, complementarily, the number of stochastic trends driving it. Another set of parameters are the values of the loading matrix. In the extreme case, all variables are just linear combinations of one stochastic trend, and the 'long-run' equilibrium is realized almost immediately after a shock. The performance of the tests might depend on how strongly the variables are tied to the long-run relation. In a set of experiments not reported here, where cointegration takes that form, both the LLC and the IPS test were very badly oversized. Since in that setup a common stochastic trend is a time specific effect common to all series, step one of LLC just eliminates it and transforms all series into stationary processes. Lower values of the loading matrix $\mathbf{A}$ may seem more realistic and loosen the tightness of the long-run relation. Here the DGP takes the following form:

$$\mathbf{A} = 0.1 \begin{pmatrix} -\mathbf{I} & 0 \\ 0 & -\mathbf{I} \end{pmatrix} \text{ and } \mathbf{B} = \begin{pmatrix} \mathbf{I} & -\mathbf{I} \\ 0 & \mathbf{C} \end{pmatrix}$$

the cointegrating matrix $\mathbf{C}$ is:

$$\mathbf{C} = \begin{pmatrix} 1 & -1 & 0 & 0 & ... & 0 \\ 0 & 1 & -1 & 0 & ... & 0 \\ 0 & 0 & 1 & -1 & ... & 0 \\ & & & & ... & \\ 0 & 0 & 0 & 0 & ... & 0 \\ 0 & 0 & 0 & 0 & ... & 0 \end{pmatrix}$$

The number of zero rows (b) determines the number of common trends driving the system. There are $N - b$ cointegrating relationships. The following number of cointegrating vectors were considered: $N - 1$, $N/2$ and $N/4$, in case of a fraction the integer part of it is chosen.

Table 6: Size properties with cointegration, $b = N - 1$

	LLC			IPS		
$b = N - 1$	nominal size 10%					
	$N = 5$	$N = 10$	$N = 20$	$N = 5$	$N = 10$	$N = 20$
$T = 25$	.167	.157	.158	.164	.150	.164
$T = 50$	.192	.167	.173	.261	.255	.274
$T = 100$	.249	.227	.203	.520	.479	.449
	nominal size 5%					
	$N = 5$	$N = 10$	$N = 20$	$N = 5$	$N = 10$	$N = 20$
$T = 25$	.094	.078	.088	.096	.086	.096
$T = 50$	.105	.081	.088	.153	.147	.174
$T = 100$	.139	.132	.117	.383	.378	.356
	nominal size 1%					
	$N = 5$	$N = 10$	$N = 20$	$N = 5$	$N = 10$	$N = 20$
$T = 25$	.024	.019	.024	.030	.024	.026
$T = 50$	.024	.017	.017	.044	.044	.055
$T = 100$	.031	.029	.034	.159	.182	.198

Note: Based on 4,000 replications.

Table 6 reports the results for $CIV = N - 1$. The tests are oversized and the problem increases in T. Together with Tables A.2 and A.3 in the Appendix, where results for other values of CIV are presented, it becomes clear that the tests perform worse the more $CIVs$ are present.

An analytical treatment of the asymptotic behavior of the LLC test statistics for the cases considered in the simulation exercise would give insights into the origins of the size distortions. The interested reader is referred to Lyhagen (2000), who provides an analytical argument for the special case in which there are $N-1$ cointegrating relations and an instantaneous adjustment to the equilibrium takes place ($\alpha = 1$). He derives the limiting distributions for the t–statistic. The variety of parameters that can determine the cointegration among the sections (number of CIVs, α) makes a general analytical treatment of this bias rather complicated. Furthermore, the additional insight of an analytical treatment is limited as a potential correction of the size distortion would have to account for all possible cases.

3.3. Cross-Sectional Correlation and Cointegration

The two previous sections indicated that both kinds of dependencies have oversizing effects and therefore yield to an over–rejection of the null-hypothesis. Neither econometric nor economic theory gives any reason to believe that the two dependencies are mutually exclusive. In this sections the two are brought together. The cointegration is chosen to be as in the above section, and, in addition, the errors are correlated in the way specified in Section 3.1.

The results reported in Table 7 are surprising. The distribution of the test statistic is shifted to the right. The bias increases with T and yields a considerable distortion in the opposite direction. This of course causes the power of the test to come close to unity.

In theory, corrections to the test statistics are possible. The variety of cases (N, T, number of CIVs, α), however, limits the practicability of such an approach. Hence, in practice, a careful assessment of the dependencies present in the data is necessary before applying any unit root test.

Table 7: Size properties with cointegration and correlation

$CIV = N/2$	LLC			IPS		
$\omega = 0.7$	nominal size 10%					
	$N=5$	$N=10$	$N=20$	$N=5$	$N=10$	$N=20$
$T=25$	.092	.055	.029	.065	.036	.015
$T=50$	.094	.058	.029	.054	.035	.007
$T=100$	.111	.064	.025	.061	.024	.004
	nominal size 5%					
	$N=5$	$N=10$	$N=20$	$N=5$	$N=10$	$N=20$
$T=25$	.049	.030	.012	.038	.019	.007
$T=50$	.055	.033	.014	.030	.010	.004
$T=100$	.067	.039	.012	.035	.010	.002
	nominal size 1%					
	$N=5$	$N=10$	$N=20$	$N=5$	$N=10$	$N=20$
$T=25$	.011	.007	.002	.011	.003	.001
$T=50$	.018	.001	.003	.007	.001	.001
$T=100$	.032	.013	.002	.011	.004	.000

Note: Based on 4,000 replications.

4. Should panel unit root tests be applied to real exchange rates?

With the growth of the panel unit root methodology, the debate over the validity of the purchasing power parity (PPP) has experienced a revival. While previous research could hardly find any empirical evidence for PPP, one could expect more insight from the application of panel methods.[5] In a non technical way, PPP means that once different currencies are controlled for, the same basket of goods should cost the same amount of money no matter in which country it is purchased. The existence of permanent deviations from such an equilibrium seems implausible as it would allow arbitrage gains, which in turn would push the real exchange rate back to the equilibrium. Although nobody believes in arbitrage possibilities with fast food, a popular application of PPP is the Economist's Big Mac index. Assuming PPP holds, actual exchange rates are expressed as the deviation from the McParity, hinting on the current under– or overvaluation of currencies.[6] Although many arguments have been put forward in the theoretical literature why PPP might fail, PPP is still a very popular concept and something many economists like to believe in. However, one cannot reject the impression that much of the debate centres on the applied methods.

[5] For a survey of empirical results before the panel era, see *e.g.* Froot and Rogoff (1995).

[6] The fall of the Euro after its introduction was predictable if one had believed in Burgernomics. For more on the issue, see Economist (2001, April 21st).

4.1. PPP - revisited

If PPP holds, in the long-run the real exchange rate between two countries is stable and deviations from equilibrium are not permanent. Let E_{it} denote the nominal exchange rage between country i and a $base$ country at time t. Then, multiplying a basket of goods (normalized to one) with the ratio of the prices in country i, P_{it}, and in the country of the base currency, P_t^{base}, defines the real exchange rate Q_{it}:

$$Q_{it} = \frac{1}{E_{it}} \frac{P_{it}}{P_t^{base}}$$

or, taking logs:

$$q_{it} = p_{it} - p_t^{base} - e_{it}. \tag{8}$$

Since prices and exchange rates are recognized to be nonstationary time series, a natural way of looking at the problem is to ask if there is a linear combination of the series which renders a stationary real exchange rate, *i.e.* if the prices and the exchange rate are cointegrated.

A distinction is made between the strong and weak form of PPP. The weak form allows for coefficients different from (1, -1) on the price indices. The weak form of PPP has its economic justification in the presence of measurement errors, which would persist in the long-run, or varying effects of productivity shocks which may cause the cointegrating coefficients to differ from unity. The weak form of PPP has been tested in an error correction approach, *e.g.* by Cheung and Lai (1993) or Corbae and Ouliaris (1991). Edison et al. (1997) and Kouretas (1997) apply a Johansen (1995) procedure, the latter to investigate PPP of the Canadian dollar and five other currencies.

The strong form of PPP, restricts the coefficients to (1, -1) *a priori* and tests the resulting real exchange rate for a unit root. Only this test is of interest in the panel unit root framework. The PPP hypothesis translates into the stationarity of the real exchange rate q_{it}: Only if this series is mean reverting and does not accumulate shocks permanently, can PPP hold. Interestingly, the majority of the studies apply tests that have a unit root as a null hypothesis and literally *accept* stationarity if nonstationarity is rejected, which clearly is a loose interpretation of the unit root rejection. Kouretas (1997) and Kuo and Mikkola (1999) are two studies which test both stationarity and nonstationarity in a panel framework. In a univariate framework, Engel (2000) points out that even if one rejects the unit root and fails to reject stationarity there is a possibility of a unit root in the series. This might be caused by a size distortion in the unit roots tests and the low power of the stationarity tests.

Several issues make PPP an interesting application from the perspective of panel unit root tests. The increased power when taking into account a set of time series allows for a more precise statement on the stationarity of the series. While single country analyses often reject PPP because a unit root is found in the real exchange rate, this might be due to the low power of single equation unit root tests with an autoregressive coefficient close to unity. Therefore, the panel approach might give more insights. However, there are drawbacks on the use of panel methods. Interestingly, some authors find differing results according to the base currency chosen. Papell (1997) rejects the unit root when the Deutschmark is chosen as a base currency, but has mixed results when the panel is US$ based. Note that the series to be tested for a unit root formed following equation (8) exhibit cross-sectional correlation by construction, as they are expressed with respect to one base currency. Hence, shocks that affect this exchange rate are directly reflected in the entire panel. This means that the degree of cross-sectional correlation depends on the base currency chosen. However, Tables 3 and 4 in Section 3.1 show that the actual value of the cross-sectional correlation does not influence the performance of the test. It is more plausible that the choice of the base currency affects the degree to which the data is contaminated with cointegration.

There is, of course, a debate on what long-run means in this context. While some authors argue that PPP should hold regardless to the exchange rate regime, and consequently apply the tests to long series from, say 1949-1996 (Kuo and Mikkola, 1999), or even over 100 years (Engel, 2000), most of the studies rely on the time period of the current float, *i.e.* from 1973 onwards.

All studies mentioned above, including Pedroni (1999), do not consider the possibility of cross-sectional cointegration. Banerjee et al. (2001) confirm the result of the previous cointegration analysis that if cross-sectional cointegration is not taken into account when the real exchange rate is computed, severe distortions may arise. Although one should be aware of the possibility of cross-sectional cointegrating relations and the serious distortions this causes, one has to recognize that large dimensional systems cannot be estimated without an *a priori* restriction. To illustrate this argument, a full Johansen estimation of the weak form of PPP would yield a system of N countries, each with 3 variables, so that an unrestricted estimation of the cointegration matrix Π would not be feasible with some 100 observations.

As mentioned earlier, the study by O'Connell (1998) examines PPP in the presence of short run dependencies in the form of cross-sectional correlations. Moreover, the size and power of the LLC test are explicitly analyzed. O'Connell comes to the conclusion that the performance of the LLC test in the presence of cross-sectional correlation is very poor and suggests a new GLS type estimator. The impact of the O'Connell critique was considerable and has to some extent discredited the LLC test. There are some things worthwhile noticing.

Apparently O'Connell does not use the adjusted t–value when he evaluates his simulations results. Not adjusting the t–values means that the finite sample adjustments are not made. Also, common time effects are not eliminated in his simulation exercise. This becomes clear as the distortion in size he reports can only be reproduced if one does not perform this elimination. The poor power properties that are attested to the LLC test are not related to the cross-sectional correlation (see Table 3). It should be pointed out that with a specification of $\rho = 0.96$ even univariate unit root tests have poor power results (Schwert, 1989). Thus the poor power properties are not panel specific. The proposed GLS estimate may seem more appealing than the removal of common time effects. However, this procedure involves the estimation of a covariance matrix and relies on the consistency and accuracy of this point estimate.

4.2. Shortrun dependence

The main finding of the simulation exercise above is that it is essential to know more about the covariance structure of the data before applying unit roots tests. This poses some methodological problems because estimators have to deal with possible heteroscedasticity and serial correlation in the data. Robust estimators are needed. In addition, once a point estimate of a covariance matrix is obtained, it is necessary to conduct some inference on the parameters in order to asses the significance of the correlations. Parametric (Den Haan and Levin, 1996) and nonparametric (Newey and West, 1987) methods for robust estimations of covariance matrices are discussed in Appendix A.2. In addition, a bootstrap algorithm (Politis and Romano, 1994) is suggested to test for significance of the estimated correlations. The data used are a panel of real exchange rates for 18 OECD countries, using the US$ as the base currency.[7] To be consistent with the covariance estimators that operate under stationarity, the first difference of the real exchange rates form the basis for the following analysis. This is consistent while working under the null hypothesis of a unit root. The parametric (Table 9) and the nonparametric (Table A.6 in the Appendix) estimation yield similar results for the covariance matrix and show clear signs of significant positive correlation.[8]

One can argue that the parametric estimate is superior because it explicitly considers prewhitening which is the main drawback of the nonparametric estimator used. On the other hand the differenced exchange rates do not, in general, have very high order autoregressive

[7]The data used is from the IMF data sets, namely the International Financial Statistics and covers quarterly nominal spot exchange rates and CPI, for the period 1973:1 to 1997:3 for 18 OECD countries. A plot of the data can be found in the Appendix.

[8]The estimates presented are not sensitive to the choice of parameters (information criterion, lag lengths, truncation $\bar{K}$). Although the two estimates are not identical, their results are very similar, and the deviations from each other are in a plausible range (see, *e.g.*[Section 6]Den Haan and Levin (1997)).

Table 8: Testing different covariance structures for 16 currencies

	$p = 0.25$	
ω	Q	p-val
0.5	295.82	1.00
0.6	85.75	0.01
0.7	2.06	0.00
0.8	44.76	0.00
0.9	213.84	1.00

Note: Q is χ^2_{120} distributed.

components,[9] so the impact of serial correlation on the nonparametric estimator might be limited.

More than the nonparametric estimate, the parametric estimates detect negative correlation of the Canadian Dollar with most other currencies in the point estimates. However, the standard errors indicate that it is not significant. Recalling that all variables are constructed the following way: $\mathbf{x}_{1t} = \frac{GBP}{US\$}, \mathbf{x}_{2t} = \frac{ATS}{US\$}$ and that from the 18 countries chosen most are European, it is not surprising that Canada seems to react in a different way to shocks – if affected at all. The same is true for Korea. The Japanese Yen, on the other hand, does exhibit similar reactions to the European currencies. In the parametric case, the standard errors are in a plausible range of 0.02 to 0.2, whereas in the nonparametric case, the standard errors become very small, especially if the estimated correlation is close to unity. Overall, the parametric estimation seems more plausible.

Having in mind the results from the simulation and the asymptotic considerations of Section 3, it is desirable to have a homogeneous dataset in terms of error correlation. Therefore, the two countries with a different error correlation (Canada and Korea) were dropped from the sample yielding a panel with almost equally correlated errors. The tests on different structures of the the covariance matrix of the remaining 16 countries reported in Table 8 indicate that a common correlation coefficient in the order of .6 to .8 cannot be rejected, with 0.7 yielding an exceptionally low test statistic.

[9]The average lag length is 2.7, with a range from zero to 7 in one case.

Table 9: Estimated sectional correlation in the differenced PPP Data, US$ as base currency.

	UK	AT	BE	DK	FR	DE	NL	CA	JP	FN	GR	ES	AU	IT	CH	KO	NW	SW
UK	1.00																	
AT	0.55 0.09	1.00																
BE	0.59 0.08	0.96 0.02	1.00															
DK	0.62 0.07	0.92 0.03	0.94 0.02	1.00														
FR	0.62 0.07	0.89 0.04	0.89 0.04	0.88 0.04	1.00													
DE	0.59 0.08	0.99 0.01	0.97 0.02	0.94 0.02	0.90 0.03	1.00												
NL	0.63 0.07	0.97 0.02	0.96 0.02	0.93 0.03	0.88 0.05	0.98 0.01	1.00											
CA	0.06 0.10	−0.03 0.11	−0.04 0.11	0.02 0.11	−0.01 0.10	−0.03 0.11	−0.04 0.11	1.00										
JP	0.50 0.08	0.62 0.07	0.63 0.07	0.64 0.07	0.61 0.08	0.63 0.07	0.64 0.07	−0.01 0.11	1.00									
FN	0.64 0.08	0.64 0.12	0.65 0.11	0.63 0.12	0.60 0.13	0.65 0.10	0.66 0.10	0.11 0.11	0.44 0.09	1.00								
GR	0.65 0.07	0.75 0.12	0.77 0.13	0.79 0.11	0.79 0.11	0.79 0.13	0.79 0.12	−0.01 0.14	0.55 0.07	0.58 0.08	1.00							
ES	0.63 0.08	0.74 0.07	0.73 0.07	0.73 0.07	0.77 0.07	0.74 0.09	0.75 0.07	−0.04 0.12	0.51 0.11	0.58 0.11	0.65 0.07	1.00						
AU	0.21 0.12	0.22 0.10	0.24 0.10	0.28 0.10	0.28 0.10	0.22 0.10	0.25 0.10	0.27 0.10	0.24 0.12	0.21 0.11	0.33 0.10	0.22 0.09	1.00					
IT	0.62 0.08	0.69 0.09	0.72 0.08	0.72 0.09	0.81 0.07	0.71 0.09	0.71 0.09	−0.04 0.13	0.50 0.09	0.59 0.12	0.69 0.09	0.74 0.08	0.23 0.08	1.00				
CH	0.53 0.09	0.84 0.04	0.82 0.05	0.82 0.05	0.87 0.05	0.85 0.04	0.82 0.06	0.05 0.12	0.64 0.06	0.57 0.13	0.70 0.09	0.66 0.10	0.23 0.10	0.70 0.08	1.00			
KO	−0.07 0.09	0.05 0.12	0.03 0.12	0.03 0.12	0.09 0.11	0.05 0.11	0.03 0.11	0.22 0.09	0.13 0.09	0.03 0.11	0.01 0.09	0.10 0.09	0.29 0.14	0.08 0.09	0.09 0.15	1.00		
NW	0.60 0.10	0.85 0.03	0.82 0.04	0.79 0.06	0.85 0.05	0.85 0.03	0.81 0.04	−0.01 0.12	0.52 0.08	0.64 0.13	0.73 0.10	0.74 0.08	0.30 0.09	0.68 0.09	0.76 0.05	0.09 0.09	1.00	
SW	0.61 0.08	0.71 0.07	0.69 0.08	0.68 0.09	0.68 0.08	0.70 0.08	0.69 0.07	−0.02 0.11	0.42 0.11	0.70 0.09	0.60 0.12	0.76 0.09	0.21 0.09	0.68 0.09	0.62 0.09	0.05 0.12	0.77 0.06	1.00
	UK	AT	BE	DK	FR	DE	NL	CA	JP	FN	GR	ES	AU	IT	CH	KO	NW	SW

Note: The values reported are correlation coefficients obtained after the correlation transformation of the estimator in (A.3). The standard errors reported are based on 1000 replications of the bootstrap procedure described in Section A.3 and computed as in (A.5).

Table 10: Trace test for cointegration in a subsample of 9 exchange rates

H_0: rank=p	$-T\sum\log(1-\hat{\lambda}_i)$	95%
$p=0$	234.5**	192.9
$p\leq 1$	183.3**	156.0
$p\leq 2$	136.7**	124.2
$p\leq 3$	97.49*	94.2
$p\leq 4$	65.67	68.5

Note: ** indicates that the hypothesis is rejected at least at the 95%-level.

4.3. Long-run dependence

A full assessment of the long-run dependency in the real exchange rate data is not possible using a maximum likelihood approach due to the few numbers of observations in relation to the entire system. A *VAR*(p) specification of the process that satisfies minimal residual properties would require a lag order higher than $p = 2$, which is the highest feasible in the system of 16 exchange rates. Estimation might be achievable by imposing further *a priori* restrictions on the parameter matrices, but theory does not give any further guidance. However, the interesting question whether there is cointegration or not can positively be answered in subsystems of the 16 exchange rates. For the sake of presentation, here the result of a subsample of 9 exchange rates is presented.[10] A cointegration analysis following Johansen (1995) suggests that the data are cointegrated. The trace test detects at least three cointegrating relations in this subsample of the data (see Table 10). This exercise could be repeated with varying subsamples yielding similar results.

4.4. Results

The individual lags that were included in the different sections were determined after the removal of common time effects. This lag structure differs from the optimal lag structure if each of the series would be tested individually before demeaning. However, because the absence of serial correlation is essential for the LLC and IPS test this should be carried out after the demeaning. A series of tests was used to analyze the residuals of each series for their white noise properties. Table 11 reports the results and the main white noise indicators. Further, the t–values of the included lags were considered which had to be significant at least for the highest lag considered. Normality is not rejected for all residual series.

[10] A *VAR*(3) was fitted allowing for seasonal dummies and a constant. All Box Pierce statistics testing for the absence of serial autocorrelation up to 11 lags cannot be rejected, the same is true for ARCH(4) effects. Absence of vector autocorrelation is rejected at the 5%–level, vector normality is not rejected.

Table 11: ADF - lags selection

	US$ based						
	lags	AR(4)	BP		lags	AR(4)	BP
DE	3	.59	.99	FN	4	.25	.58
UK	1	.90	.34	GR	5	.87	.71
AT	1	.57	.87	ES	1	.78	.60
BE	3	.77	.24	AU	3	.96	.84
DK	2	.44	.45	IT	2	.80	.85
FR	1	.74	.83	CH	1	.54	.85
NL	3	.47	.90	NW	3	.87	.12
JP	5	.41	.60	SW	2	.84	.85

Note: In all cases the null hypothesis is absence of the respective disturbance. The reported values are the p values at which this hypothesis can be rejected. AR(4) stands for a test on autocorrelation to the 4th order, BP is the Breush Pagan test for heteroscedasticity.

Table 12: Test results

test	coefficient	p–values N(0,1)	simulated $b = 15$	simulated $b = 8$	adj. power (5 %) $\rho = .9$	adj. power (5 %) $\rho = .95$
LLC t^*	-1.911	0.028	0.014	0.020	0.94	0.38
IPS $\Psi_{\bar{t}}$	-2.856	0.002	0.001	0.002	0.99	0.73

Note: Simulated values based on 4,000 replications.

Both dependencies are present in the data. The simulation exercise has shown that in this case it is not possible to make predictions about the direction of a potential size bias. New critical values can be computed simulating panels of exchange rates, and thereby following as close as possible the presumed DGP. Therefore, a panel of 16 variables with 15 (and 8) cointegrating relations and an error correlation structure using the point estimate of Table 9 was simulated and the test statistic was computed. Under the alternative, variables with an autoregressive coefficient of $\rho = \{.9, .95\}$ and the same error structure were simulated. Table 12 reports the results of the LLC and the IPS test, the percentiles of the normal and the simulated distribution. In addition, Figure 1 shows kernel densities of the estimated t–values, the standard normal distribution, the actual test value and the $\rho = 0.9$ alternative.

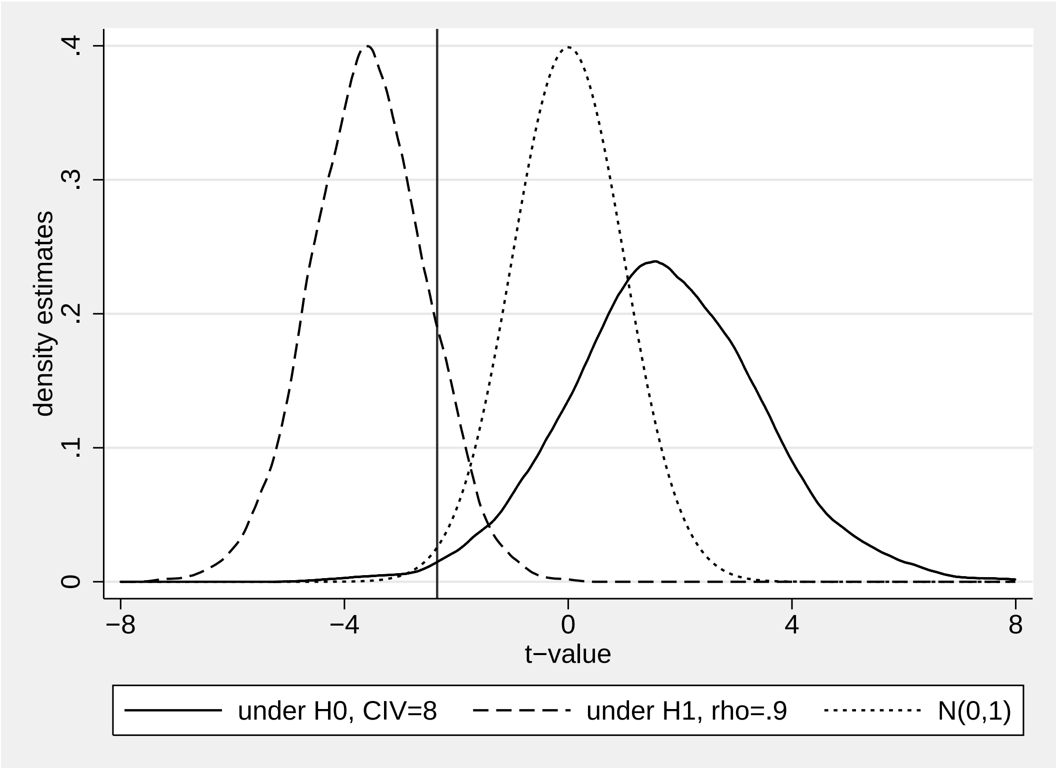

Figure 1: LLC test results.

One has to keep in mind, however, that this test result is sensitive to the assumed structure of the data, in particular the presence of both long and short run dependence. While the short run correlation with all errors sharing the same correlation coefficient does not appear to influence the test result, the presence of cointegration is much more worrisome. If one is willing to assume that the values obtained via the simulation reflect the true properties of the DGP, the null of nonstationarity of the real exchange rates can be rejected at a very low p–value, hence providing some argument for the validity of PPP.

5. Conclusion

The simulation exercise has shown that two of the most popular panel unit root tests are sensitive to dependencies among sections of the panel. Analyzed separately, both short run dependence in the form of correlated errors and long-run dependence in the form of cointegration lead to a significant oversizing of the test. However, if put together, the effect goes into the opposite direction. The determination of the actual presence of dependencies is therefore necessary in order to interpret the test results. The estimation of and the inference on contemporaneous correlation is crucial, although not easy to perform.

The application to a set of real exchange rates has shown that both dependencies are present in the data. Hence, the test results are likely to be biased. In order account for these dependencies, simulated critical values were used which origin from a data generating process that resembles the actual data. The null hypothesis of a unit root can be rejected providing some empirical evidence for the validity of he purchasing power parity. However, in the light of the simulation results obtained earlier, the reliability of the test results are questionable. This exemplifies the the problems with the use of panel unit root tests on sectional dependent data in general.

References

ANDREWS, Donald W. K. and MONAHAN, J. Christopher (1992): «An Improved Heteroscedasticity and Autocorrelation Consistent Covariance Matrix Estimator», *Econometrica*, 60 (4): 953–966.

BANERJEE, Anindya; MARCELLINO, Massimiliano and OSBAT, Chiara (2000): *Some Cautions on the Use of Panel Methods for Integrated Series of Macro-Economic Data*, EUI Working Papers (2000/20), Italy: European University Intitute.

BANERJEE, Anindya; MARCELLINO, Massimiliano and OSBAT, Chiara (2001): «Testing for PPP: Should We Use Panel Methods», *mimeo*.

BILLINGSLEY, Patrick (1986): *Probability and Measure*, Wiley Series in Probability and Mathematical Statistics, New York: Wiley, 2nd ed.

BOSE, A. (1988): «Edgeworth Correction by Bootstrap in Autoregressions», *Annals of Statistics*, 16: 1709–1722.

BREITUNG, Joerg and MEYER, Wolfgang (1991): *Testing for Unit Roots in Panel Data: Are Wages on Different Bargaining Levesl Constrained?*, Institut fuer Quantitative Wirtschaftsforschung Working Paper, Humboldt Universitaet zu Berlin.

CHEUNG, Y. W. and LAI, K. S. (1993): «Long Run Purchasing Power Parity During the Recent Float», *Journal of International Economics*, 34: 181–192.

CORBAE, D. and OULIARIS (1991): «A Test of Long Run PPP Allowing for Structural Breaks», *Economic Record*, 67: 26–33.

CROWDER, William J. (1997): *Purchasing Power Parity During the Modern Fload: A Look Behind The Panels*, Working Paper, University of Texas at Arlington.

DAVIDSON, James (1994): *Stochastic Limit Theory: An Introduction for Econometricians*, Advanced Texts in Econometrics, Oxford: Oxford University Press.

DAVISON, A.C. and HINKLEY, D.V. (1997): *Bootstrap Methods and their Applications*, Cambridge: Cambridge University Press.

DEN HAAN, Wouter J. and LEVIN, Andrew (1996): *Inferences from Parametric and Non-Parametric Covariance Matrix Estimation Procedures*, NBER Technical Working Paper Series, Cambridge, MA: NBER.

DEN HAAN, Wouter J. and LEVIN, Andrew T. (1997): «A Practitioner's Guide to Robust Covariance Matrix Estimation», vol. 15 of *Handbook of Statistics, Chapter 12*, 299–342.

DICKEY, D. and FULLER, W. (1981): «Likelihood Ratio Tests for Autoregressive Time Series with a Unit Root», *Econometrica*, 49: 1057–1072.

ECONOMIST, The (2001): «Economics Focus: The Big Mac Index», *The Economist*, 82.

EDISON, H. J.; GAGNON, J.E. and MELICK, W.R. (1997): «Understanding the Empirical Literature on Purchasing Power Parity: The Post Bretton Woods Era», *Journal of International Money and Finance*, 16: 1–17.

ENGEL, Charles (2000): «Long-Run PPP May Not Hold After All», *Journal of International Economics*, 57: 243–273.

EVANS, Paul and KARRAS, Georgios (1996): «Do Economies Converge? Evidence from a Panel of U.S. States», *Review of Economics and Statistics*, 78 (3): 384–388.

FRANKE, J. and HAERDLE, W. (1992): «On Bootstrapping Kernel Spectral Estimates», *Annals of Statistics*, 20 (1): 121–145.

FROOT, K. and ROGOFF, K. (1995): «Perspectives of PPP and Long-Run Real Exchange Rates», in: «Handbook of International Economics», (eds. GROSSMAN, G. and ROGOFF, K.), Amsterdam: North-Holland.

GREENE, William H. (2000): *Econometric Analysis*, New Jersey: Prentice Hall, 4th ed.

HALL, Peter; HOROWITY, Joel L. and JING, Bing Yi (1995): «On Blocking Rules for the Bootstrap with Dependent Data», *Biometrika*, 82 (3): 561–574.

IM, K.; PESARAN, M.H. and Y., Shin (1997): *Testing for Unit Roots in Heterogeneous Panels*, Discussion Paper, Department of Applied Economics, University of Cambridge.

JOHANSEN, Søren (1995): *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford and New York: Oxford University Press.

KOURETAS, Georgios P. (1997): «The Canadian Dollar and Purchasing Power Parity During the Recent Float», *Review of International Economics*, 5 (4): 467–477.

KUO, Biing Shen and MIKKOLA, Anne (1999): *How Sure are We About PPP? Panel Evidence With the Null of Stationary Real Exchange Rates*, CEPR Discussion Paper: 2120, CEPR.

LEVIN, Andrew and LIN, Chien-Fu (1992): *Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Porperties*, University of California Discussion Paper (92-23), University of California, San Diego.

LEVIN, Andrew and LIN, Chien-Fu (1993): *Unit Root Tests in Panel Data: New Results*, University of California Discussion Paper (93-56), University of California, San Diego.

LEVIN, Andrew; LIN, Chien-Fu and CHU, Chia-Shang James (2002): «Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties», *Journal of Econometrics*, 108 (1): 1–24.

LYHAGEN, Johan (2000): *Why Not Use Standard Panel Unit Root Test for Testing PPP*, vol. 413 of *Stockholm School of Economics Working Paper Series in Economics and Finance*.

NEWEY, W. K. and WEST, K.D. (1987): «A Simple Positive Semi-Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix», *Econometrica*, 55: 703–708.

NEWEY, Whitney K. and WEST, Kenneth D. (1994): «Automatic Lag Selection in Covariance Matrix Estimation», *Review of Economic Studies*, 61: 631–653.

NICKELL, Stephen (1981): «Biases in Dynamic Models with Fixed Effects», *Econometrica*, 49 (6): 1417–1426.

O'CONNELL, Paul G.J. (1998): «The Overvaluation of Purchasing Power Parity», *Journal of International Economics*, 44: 1–19.

PAPELL, D. (1997): «Searching for Stationarity: Purchasing Power Parity under the Current Float», *Journal of International Economics*, 43: 313–323.

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

PESARAN, M.H. and SMITH, R. (1995): «Estimating Long-Run Relationships from Dynamic Heterogeneous Panels», *Journal of Econometrics*, 68: 79–113.

PHILLIPS, P. C. B. and MOON, H.R. (1999): «Linear Regression Limit Theory for Nonstationary Panel Data», *Econometrica*, 67: 1057–1111.

PHILLIPS, P.C.B. and DURLAUF, S.N. (1986): «Multiple Time Series Regression with Integrated Processes», *Review of Economic Studies*, 53: 473–495.

POLITIS, Dimitris N. and ROMANO, Joseph P. (1994): «The Stationary Bootstrap», *Journal of the American Statistical Association*, 89 (428): 1303–1313.

QUAH, Danny (1990): *International Patterns of Growth: I. Persistence in Cross-Country Disparities*, MIT Working Paper.

SCHWERT, G.W. (1989): «Tests for Unit Roots: A Monte Carlo Investigation», *Journal of Business and Economics Statistics*, 7: 147–159.

STRAUSS, Jack (2000): «Is There a Permanent Component in US Real GDP», *Economics Letters*, 66 (2): 137–142.

WHITE, Halbert (2001): *Asymptotic Theory for Econometricians*, London: Academic Press, 2nd ed.

WU, Jyh Lin (2000): «Mean Reversion of the Current Account: Evidence from the Panel Data Unit-Root Test», *Economics Letters*, 66 (2): 215–222.

A. Appendix

A.1. Asymptotic normality of LLC with common shocks

Absence of shocks

Consider a model without constant and no additional lagged differences, along the lines of Levin et al. (2002). Following the notation from Section 2.1, in this case, for large N and T, no adjustments are necessary and $t^*_\delta = t_\delta$. The least squares estimator of δ proposed by LLC under the null hypothesis is:[11]

$$\hat{\delta} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}\epsilon_{it}x_{it-1}}{\sum_{i=1}^{N}\sum_{t=1}^{T}x_{it-1}^2}$$

Define

$$\xi_{1iT} = \frac{1}{\sigma^2 T}\sum_{t=1}^{T}\epsilon_{it}x_{it-1} \text{ and } \xi_{2iT} = \frac{1}{\sigma^2 T^2}\sum_{t=1}^{T}x_{it-1}^2$$

and, using an estimator for the standard deviation[12] σ, the corresponding t–value is:

$$t_\delta = \frac{\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\xi_{1iT}}{\left(\frac{\hat{\sigma}}{\sigma}\right)\left[\frac{1}{N}\sum_{i=1}^{N}\xi_{2iT}\right]^{1/2}}.$$

Sectional correlation is a violation concerning the N. Taking the easiest form of multi index asymptotics, namely sequential limits (Phillips and Moon, 1999) first the limiting distributions when T goes to infinity is, for N fixed:[13]

$$\begin{aligned} \lim_{T\to\infty}\xi_{1iT} &= \xi_{1i} \text{ with } E[\xi_{1i}] = 0 \text{ and } Var[\xi_{1i}] = \frac{1}{2} \\ \lim_{T\to\infty}\xi_{2iT} &= \xi_{2i} \text{ with } E[\xi_{2i}] = \frac{1}{2} \text{ and } Var[\xi_{2i}] = \frac{1}{3} \end{aligned} \tag{A.1}$$

one obtains

$$t_\delta = \frac{\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\xi_{1i}}{\left(\frac{\hat{\sigma}}{\sigma}\right)\left[\frac{1}{N}\sum_{i=1}^{N}\xi_{2i}\right]^{1/2}}. \tag{A.2}$$

If the errors were uncorrelated and $\hat{\sigma}$ a consistent estimator of σ, averaging over the sections of the panel would give the known result that $t_\delta \Rightarrow N(0,1)$. The convergence in probability of the denominator of (A.2) is established by the following application of a law of large numbers (Billingsley, 1986, p. 290):

[11] See Section 2.1.

[12] For example:

$$\hat{\sigma} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}x_{it}x_{it-1}}{\sum_{i=1}^{N}\sum_{t=1}^{T}x_{it-1}^2}.$$

[13] These results are due to Phillips and Durlauf (1986), cf. Levin and Lin (1992, p. 14).

Theorem A.1 Suppose that for each time-series dimension T, the variables Z_{iT} are independent and identically distributed across individuals i, with mean μ_T and variance $0 < \sigma_T^2 < \infty$, and that $\mu = \lim_{T\to\infty} \mu_T$. If $\lim_{T\to\infty} \frac{\sigma_T^2}{N_T} = 0$. Then $\frac{1}{N_T}\sum_{i=1}^{N_T} Z_{iT} \xrightarrow{p} \mu$.

The inner expression of the denominator has expectation 1/2 and for all i the expectations of the variance are finite. Hence, the denominator converges to $\sqrt{0.5}$. The convergence in distribution of the numerator is established by applying the following central limit theorem (Billingsley, 1986, p. 368):

Theorem A.2 Suppose that for each time-series dimension T, the variables Z_{iT} are independent and identically distributed across individuals i, with mean μ_T and variance $0 < \sigma_T^2 < \infty$, and that $\mu = \lim_{T\to\infty} \mu_T$, and $\sigma^2 = \lim_{T\to\infty} \sigma_T^2$. Then $\frac{1}{\sqrt{N_T}}\sum_{i=1}^{N_T}(Z_{iT} - \mu_T) \Rightarrow N(0, \sigma^2)$.

For each i, the numerator has expectation 0 and finite variance 1/2. Hence, it converges in distribution to $N(0, 0.5)$. Using the results obtained in (A.1) and applying both theorems, (A.2) converges to $N(0, 1)$.[14]

Common shocks

In the case of sectional correlation, however, the crucial assumption used in both theorems about the independence of the random variables is violated and their application fails. The numerator of (A.2) no longer converges to $N(0, 0.5)$. To be more precise, assume the easiest case in which the correlation among sections takes the following form:[15]

$$E[\xi_{1i}\xi_{1j}] \neq 0 \text{ and } Cov[\xi_{1i}\xi_{1j}] = \tilde{\omega}_1 \text{ for } i \neq j$$

To see which central limit theorem can be applied, it is necessary to check the properties of $\xi_1 \equiv \lim_{N\to\infty} \frac{1}{\sqrt{N}}\sum_{i=1}^{N} \xi_{1i}$:

$$\begin{aligned} E[\xi_1] &= E[\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\xi_{1i}] = 0 \text{ and} \\ Var[\xi_1] &= Var[\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\xi_{1i}] = \frac{1}{N}\Big(\sum_{i=1}^{N} Var(\xi_{1i}) + 2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} Cov(\xi_{1i}, \xi_{1j})\Big) \\ &= \frac{1}{N}\Big(\frac{N}{2} + N(N-1)\tilde{\omega}_1\Big) = \frac{1}{2} + \tilde{\omega}_1(N-1) \end{aligned}$$

The variance of the numerator increases with N. Central limit theorems for dependent random variables require a finite variance to establish convergence (see, *e.g.* Billingsley (1986, p. 376) and White (2001, p. 122)). Hence, no convergence result can be stated for this general form of dependence.

However, the analysis of the elimination of common time effects above (see page 8) has shown that the *effective* disturbance to the correlation matrix *after* removing common time effects is itself a function of N. More specifically, using the result from equation (6) that $\tilde{\omega}_1 = (1 - \tilde{\omega}_1)(\frac{N-1}{N}\frac{-1}{N-1}) = \frac{\tilde{\omega}_1 - 1}{N}$ one can rewrite the above after removing common time effects as:

$$Var[\xi_1] = \frac{1}{2} + \frac{\tilde{\omega}_1 - 1}{N}.$$

[14] For the variance, notice that $Var(\frac{N(0,0.5)}{\sqrt{0.5}}) = \frac{0.5}{0.5} = 1$.

[15] Note that if $Var(\epsilon_{it}) = 1$ the covariances equal the correlation coefficients.

As N goes to infinity the variance converges to the same value as in the case without sectional correlation. Using a central limit theorem which does not require independence (White, 2001, p. 125):

Theorem A.3 Suppose that for each time series dimension T, Z_{iT} is a stationary process with mean μ_T and variance $0 < \sigma_T^2 < \infty$, and that $Var(\frac{1}{\sqrt{N_T}} \sum_{i=1}^{N_T} Z_{iT}) \xrightarrow{p} \sigma_N^2$ where $0 < \sigma_N^2 < \infty$. Then $\frac{1}{\sqrt{N_T}} \sum_{i=1}^{N_T} (Z_{iT}) \Rightarrow N(0, \sigma_N^2)$.

Since each section has a finite variance and the variance of the average over all sections converges to 1/2, the numerator converges to $N(0, 0.5)$. For the denominator write:

$$E[\xi_{2i}\xi_{2j}] \neq 0 \text{ and } Cov[\xi_{2i}\xi_{2j}] = \tilde{\omega}_2 \text{ for } i \neq j.$$

Again, checking the properties of $\xi_2 \equiv \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \xi_{2i}$ it is easily seen that:

$$\begin{aligned}
E[\xi_2] &= \frac{1}{2} \text{ as before and} \\
Var[\xi_2] &= Var[\frac{1}{N}\sum_{i=1}^{N} \xi_{2i}] = \frac{1}{N^2}\Big(\sum_{i=1}^{N} Var(\xi_{2i}) + 2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} Cov(\xi_{1i}, \xi_{1j})\Big) \\
&= \frac{1}{N^2}\Big(\frac{N}{3} + N(N-1)\tilde{\omega}_2\Big) = \frac{1}{3N} + \tilde{\omega}_2 \frac{N-1}{N}.
\end{aligned}$$

The variance of the denominator decreases with N. Following White (2001, p. 44), the following law of large numbers is applicable to weakly dependent data:

Theorem A.4 Suppose Z_i is a stationary ergodic scalar sequence with $E[Z_i] = \mu < \infty$. Then $Z_i \xrightarrow{a.s.} \mu$.

Almost sure convergence (a.s.) implies convergence in probability (Davidson (1994)). Hence, the denominator converges in probability to $\sqrt{(0.5)}$ regardless to the dependence in the data. Summarizing, with common sectional correlation and after removing common time effects, (A.2) will converge in distribution to $N(0, 1)$ as it is the case without sectional correlation. This is in line with the simulation results obtained earlier, that the problem of oversizing diminishes with N and T.

Other shocks

In the case where the covariance matrix takes the form of a band matrix (see page 9) if common time effects were not eliminated, the variance of the numerator, in terms of the expressions above, would again not converge for large N as it becomes $Var(\xi_1) = 1/2 + \sum_{a=1}^{N-1} a\omega^{N-a}$. If common time effects are eliminated, the structure of the correlation matrix becomes even less homogeneous and no convergence is achieved. The resulting matrix is a straightforward but rather unpleasant combination of N's and ω's. Here is a numerical example:

$$\text{if } \mathbf{\Sigma} = \begin{pmatrix} 1 & & & \\ .7 & 1 & & \\ .49 & .7 & 1 & \\ .34 & .49 & .7 & 1 \end{pmatrix} \text{ then } \mathbf{Q\Sigma Q'} = \begin{pmatrix} 1 & & & \\ .46 & 1 & & \\ -.81 & -.42 & 1 & \\ -.79 & -.81 & .46 & 1 \end{pmatrix}.$$

Hence, the elimination of common time effects in this case does not provide any remedy for the test. The test statistic will not converge to a N(0,1)

A.2. Estimation of and Inference on (Co-)Variances

Much attention has been devoted to so called heteroscedasticity and autocorrelation consistent (HAC) or robust estimators of covariances from to stationary series. Under nonstationarity, covariances are not constant over time and methods designed for stationary series can no longer be used. In this case one can either use the differences of the series to compute the covariances or the residuals from regression on the lagged variable. Schwert (1989) finds that, in the univariate case, the difference based approach has a smaller bias in finite samples. Therefore, and in order to proceed consistently under the null of non-stationarity, the following lines apply to the first differences of an I(1) process without drift.

Analogous to the univariate problem of variance estimation, the aim is to get a consistent estimate of the covariance matrix at zero frequency.[16] To estimate the spectrum of an unknown DGP correctly, all T autocovariances have to be estimated, which is not feasible with T observations. The class of parametric estimators imposes a certain structure on the data and constructs estimators that would be implied by the model, while nonparametric procedures use a weighted average of autocovariances.

Parametric estimators

The parametric estimator VARHAC (vector autoregressive heteroscedasticity and autocorrelation consistent) was developed by Den Haan and Levin (1996) and fits a vector autoregressive (*VAR*) model to the series under consideration using an information criterion to determine the optimal lag length. To the residuals of that *VAR* a standard covariance estimator is applied.

More specifically, for each section i of the N-dimensional stationary of the process $\mathbf{x}_t$, an autoregressive process is fitted using a lag order suggested by either the Akaike (AIC) or the Schwarz' Bayesian (BIC) information criterion, and given a maximum lag order. The optimal lag order may differ across sections. The coefficients are collected in a matrix $\hat{\mathbf{A}}_{k(N\times N)}$ for each lag k, taking zero values for section i if k exceeds the maximum lag order of that section. For the highest lag length $\bar{K}$ chosen, a *VAR* is fitted and the residuals $\hat{\mathbf{e}}_t$ are used to compute the innovation matrix:

$$\hat{\mathbf{\Sigma}}_T^{VARHAC} = \frac{1}{T}\sum_{t=\bar{K}}^{T} \hat{\mathbf{e}}_t \hat{\mathbf{e}}_t'.$$

The the spectral density estimator is then given by:

$$\hat{\mathbf{S}}_T^{VARHAC} = [I - \sum_{k=1}^{\bar{K}} \hat{\mathbf{A}}_k]^{-1} \hat{\mathbf{\Sigma}}^{VARHAC} [I - \sum_{k=1}^{\bar{K}} \hat{\mathbf{A}}_k]^{-1}. \tag{A.3}$$

Den Haan and Levin (1996) analyze the performance of this estimator compared to some nonparametric alternatives and find better finite sample properties. According to their results, the individual choice of lag lengths for each section makes this procedure superior to nonparametric estimates, in which one weighting function is applied to all sections.

Nonparametric estimators

In the nonparametric case two concepts are introduced to handle the problem of estimating the covariances: windowing and weighting. The most frequently used kernels in the time series literature

[16] For the following, see Den Haan and Levin (1997).

are the Bartlett kernel and the Parzen kernel. For the Bartlett kernel, the weights assigned to the autocovariances decline from 1 (the sample variance) to 0 (when the truncation is reached). This kernel ensures a positive estimation of the long-run variance – or, in the multivariate case – a positive definite estimate of the covariance matrix (Newey and West, 1987). Since the theoretical guidance in the choice of the truncation is quite unsatisfactory, it might be useful in empirical applications to conduct robustness checks in terms of varying kernels and truncation parameters. Starting point for the estimation of the covariance matrix in the presence of serial correlation is:

$$\mathbf{S}(m) = \hat{\mathbf{\Gamma}}_{\mathbf{0}} + \sum_{\tau=1}^{K} w_{K\tau}(\hat{\mathbf{\Gamma}}_\tau + \hat{\mathbf{\Gamma}}'_\tau), \tag{A.4}$$

where $w_{K\tau}$ is a kernel, K a truncation parameter, and

$$\hat{\mathbf{\Gamma}}_\tau = \frac{1}{T}\sum_{t=\tau+1}^{T}(\mathbf{x}_t - \bar{\mathbf{x}})(\mathbf{x}_{t-\tau} - \bar{\mathbf{x}})'.$$

Refinements to this estimator are possible. Kernel based estimations of the long-run variance matrix in the presence of serial correlation were found to give quite poor results. The major source of bias is that kernels, which ensure a positive definite spectral density matrix place weights less then unity on autocovariances other then at lags zero. Andrews and Monahan (1992) therefore suggest a kernel based prewhitening of the series and observe an improvement using this technique. In an expression similar to equation (A.3) the covariance matrix is placed between the inverse of the prewhitening coefficients. Newey and West (1994) propose an automated bandwidth selection procedure for the estimator in (A.4).

A.3. Bootstrap methods for dependent data

Inference on covariance matrix estimators is rarely done. But the estimation results itself are meaningless if they remain unrelated to some standard errors. In both the parametric and nonparametric case, bootstrap methods may be used to make inference on the estimates.

There is little known about the properties of bootstrap algorithms when the underlying process contains a unit root. But even if the root of the process comes close to unity, Bose (1988) shows that bootstrap approximations deteriorate. However, bootstrapping results will remain valid if the bootstraps are applied to the differenced data. Hence, the following applies to the first differences of a non-stationary process.

For time dependent data, however, the algorithms have to be extended because random resampling would not account for the time dependency of the data, which, as it is the case for the covariance matrix, is a crucial part of the estimator.[17] For time dependent processes, resampling in the frequency domain is suggested *e.g.* by Franke and Haerdle (1992) for the univariate case. This method is designed for making inference about the entire spectrum. In this context the only estimate of interest is the variation of the covariance matrix at zero frequency, and therefore these methods do not seem appropriate. In the time series domain the following methods are suggested. The so called model-based resampling requires reasonable good knowledge of the true model. In short, the assumed DGP is applied to the series, innovations are computed and then used to resample a series again assuming the same DGP. Among the methods that do not require knowledge of the DGP is the so called block

[17] See *e.g.* Davison and Hinkley (1997).

resampling. The basic idea here is to divide the data into b blocks of equal length l. The new resampled series is a randomly order of blocks. Typically, those estimates will be biased, because the resampled series are more independent than the original one, since whenever a block changes, artificial independence is introduced. Furthermore, this break causes the artificial series to exhibit nonstationarity properties, because distribution parameters become time dependent.[18]

The stationary bootstrap

The stationary bootstrap suggested by Politis and Romano (1994) is a sophistication of the aforementioned methods. Moreover, this bootstrap is unbiased and does not produce nonstationarity in the above sense. Another advantage of this method is that its validity for the covariance estimation of a multivariate process was shown, which is precisely what is needed in this context (Politis and Romano, 1994, Theorem 4). The algorithm is as follows:

- Let $\mathbf{x}_t$ be a N-dimensional vector of time series from $t = 1, ..., T$.
- Define $b_{t,l} = \{\mathbf{x}_t, \mathbf{x}_{t+1}, ...\mathbf{x}_{t+l-1}\}$ as a block of l subsequent observations in the sample, starting at some t. If the end of the sample is reached before the end of the block (*i.e.* $t + l > T$), the block is filled up with observations of the beginning of the sample ($\mathbf{x}_N = \mathbf{x}_0$, $\mathbf{x}_{N+1} = \mathbf{x}_1$ ...).
- The length l of the blocks is determined randomly, where the lengths follow a geometrical distribution with some fixed parameter $p \,\epsilon\, [0, 1]$. The probability of a block length m is Pr$\{l = m\} = (1 - p)^{(m-1)} p$ for m = 1, 2, ... Denote those random numbers by L_i.
- Once the lag length is determined, the beginning of the block is determined by a random variable I_i which is discretely uniformly distributed on $[1, T]$.
- The pseudo time series $\mathbf{x}_t^* = \{\mathbf{x}_1^* ... \mathbf{x}_T^*\}$ is generated by the random sequence of blocks $B_{I_1,L_1}, B_{I_2,L_2} ...$, where the end is trimmed at T. The resampling is done B times.
- Let the true vector of parameters of interest be $\boldsymbol{\theta}$. In the same way as the distribution of $\mathbf{x}$ can be approximated by the large number of pseudo series $\mathbf{x}^*$, the distribution of $\hat{\boldsymbol{\theta}}$ conditional on $\mathbf{x}$ can be approximated by the distribution of $\hat{\boldsymbol{\theta}}^*(\mathbf{x}^*)$.
- Applying this procedure to the inference on a covariance matrix, $\boldsymbol{\theta}$ is a vector containing the correlations between the N units of the the vector $\mathbf{x}_t$. If one restricts attention to the triangle below the diagonal, this amounts to $(\frac{1}{2}(N-1)N) = d$ elements. Denote by $\hat{\boldsymbol{\theta}}(\mathbf{x})$ a consistent (parametric or nonparametric) estimator of the covariance matrix. After computing the covariances for each resampled $\mathbf{x}^*$, one can estimate the asymptotic variance of the estimator by:[19]

$$\hat{\mathbf{V}}[\hat{\boldsymbol{\theta}}] = \frac{1}{B} \sum_{b=1}^{B} [\hat{\boldsymbol{\theta}}^*(b) - \hat{\boldsymbol{\theta}}][\hat{\boldsymbol{\theta}}^*(b) - \hat{\boldsymbol{\theta}}]'. \tag{A.5}$$

  The diagonal elements of $\hat{\mathbf{V}}$ contain the variances of each element of the estimator, hence the root of the diagonal contains the standard error to be placed around the point estimates $\hat{\boldsymbol{\theta}}$.

[18] For methods on how to overcome this and other problems, see Hall et al. (1995).

[19] See Greene (2000, p. 174).

- Assuming that the bootstrapped values follow a normal distribution, then a simple test for $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}_1$ is:

$$Q = (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_1)' \tilde{\mathbf{V}}^{-1} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_1) \tag{A.6}$$

where $\tilde{\mathbf{V}}$ is an estimate of the covariance matrix of the form:

$$\tilde{\mathbf{V}}[\hat{\boldsymbol{\theta}}] = \frac{1}{B} \sum_{b=1}^{B} [\hat{\boldsymbol{\theta}}^*(b)][\hat{\boldsymbol{\theta}}^*(b)]'$$

Then Q will be approximately χ^2_d distributed, where d is the dimension of $\boldsymbol{\theta}$.[20]

A.4. Additional tables

Table A.1: Size properties when all assumptions are fulfilled

$\omega = .0$	LLC			IPS		
	nominal size 10%					
	$N=5$	$N=10$	$N=20$	$N=5$	$N=10$	$N=20$
$T=25$	.136	.120	.118	.120	.114	.109
$T=50$	.122	.101	.101	.114	.100	.104
$T=100$	.118	.100	.100	.101	.101	.098
	nominal size 5%					
	$N=5$	$N=10$	$N=20$	$N=5$	$N=10$	$N=20$
$T=25$	.073	.058	.058	.069	.059	.059
$T=50$	.062	.053	.047	.064	.052	.054
$T=100$	.058	.056	.049	.058	.057	.048
	nominal size 1%					
	$N=5$	$N=10$	$N=20$	$N=5$	$N=10$	$N=20$
$T=25$	.015	.011	.010	.018	.012	.012
$T=50$	.013	.011	.009	.015	.013	.009
$T=100$	.013	.010	.009	.013	.012	.009

Note: Based on 4,000 replications.

[20] See also Den Haan and Levin (1997, p. 299).

Table A.2: Size properties with cointegration

$CIV = N/2$	LLC			IPS		
	nominal size 10%					
	$N=5$	$N=10$	$N=20$	$N=5$	$N=10$	$N=20$
$T=25$	.155	.146	.147	.150	.142	.142
$T=50$	.169	.151	.145	.179	.185	.194
$T=100$	.178	.160	.149	.267	.291	.301
	nominal size 5%					
	$N=5$	$N=10$	$N=20$	$N=5$	$N=10$	$N=20$
$T=25$	.081	.080	.077	.081	.083	.080
$T=50$	.086	.077	.072	.099	.103	.116
$T=100$	.010	.088	.078	.167	.188	.202
	nominal size 1%					
	$N=5$	$N=10$	$N=20$	$N=5$	$N=10$	$N=20$
$T=25$	.017	.014	.016	.020	.017	.022
$T=50$	.020	.014	.013	.031	.027	.024
$T=100$	.026	.022	.014	.052	.064	.072

Note: based on 4,000 replications.

Table A.3: Size properties with cointegration, $b = N/4$

$b = N/4$	LLC			IPS		
	nominal size 10%					
	$N=5$	$N=10$	$N=20$	$N=5$	$N=10$	$N=20$
$T=25$	.144	.137	.135	.126	.132	.139
$T=50$	.147	.132	.130	.152	.143	.142
$T=100$	.164	.146	.142	.184	.184	.186
	nominal size 5%					
	$N=5$	$N=10$	$N=20$	$N=5$	$N=10$	$N=20$
$T=25$	.074	.072	.073	.069	.077	.074
$T=50$	.074	.070	.072	.085	.078	.077
$T=100$	.070	.082	.074	.108	.109	.112
	nominal size 1%					
	$N=5$	$N=10$	$N=20$	$N=5$	$N=10$	$N=20$
$T=25$	.016	.014	.014	.019	.021	.017
$T=50$	.014	.016	.017	.021	.016	.019
$T=100$	.021	.017	.019	.029	.026	.024

Note: Based on 4,000 replications.

Table A.4: Size properties with cointegration and correlation for $b = N - 1$

	LLC			IPS		
$b = N - 1$	nominal size 10%					
$\omega = 0.7$	$N = 5$	$N = 10$	$N = 20$	$N = 5$	$N = 10$	$N = 20$
$T = 25$	.092	.056	.023	.079	.045	.022
$T = 50$	.084	.045	.020	.086	.032	.013
$T = 100$	.052	.038	.013	.122	.062	.019
	nominal size 5%					
	$N = 5$	$N = 10$	$N = 20$	$N = 5$	$N = 10$	$N = 20$
$T = 25$	.045	.027	.017	.042	.021	.011
$T = 50$	.042	.019	.007	.043	.013	.005
$T = 100$	.025	.018	.007	.064	.031	.008
	nominal size 1%					
	$N = 5$	$N = 10$	$N = 20$	$N = 5$	$N = 10$	$N = 20$
$T = 25$	.010	.007	.003	.011	.006	.002
$T = 50$	.011	.004	.001	.010	.003	.001
$T = 100$	.008	.005	.001	.018	.007	.001

Note: Based on 4,000 replications.

Table A.5: Size properties with cointegration and correlation for $b = N/4$

	LLC			IPS		
$b = N/4$	nominal size 10%					
$\omega = 0.7$	$N = 5$	$N = 10$	$N = 20$	$N = 5$	$N = 10$	$N = 20$
$T = 25$	.080	.054	.026	.059	.033	.013
$T = 50$	.091	.068	.029	.046	.021	.005
$T = 100$	.118	.068	.039	.048	.016	.005
	nominal size 5%					
	$N = 5$	$N = 10$	$N = 20$	$N = 5$	$N = 10$	$N = 20$
$T = 25$	.043	.029	.012	.034	.017	.005
$T = 50$	.054	.041	.014	.023	.011	.003
$T = 100$	.081	.042	.025	.028	.010	.001
	nominal size 1%					
	$N = 5$	$N = 10$	$N = 20$	$N = 5$	$N = 10$	$N = 20$
$T = 25$	.013	.007	.002	.008	.003	.001
$T = 50$	.017	.012	.004	.005	.002	.001
$T = 100$	.030	.018	.009	.008	.003	.000

Note: Based on 4,000 replications.

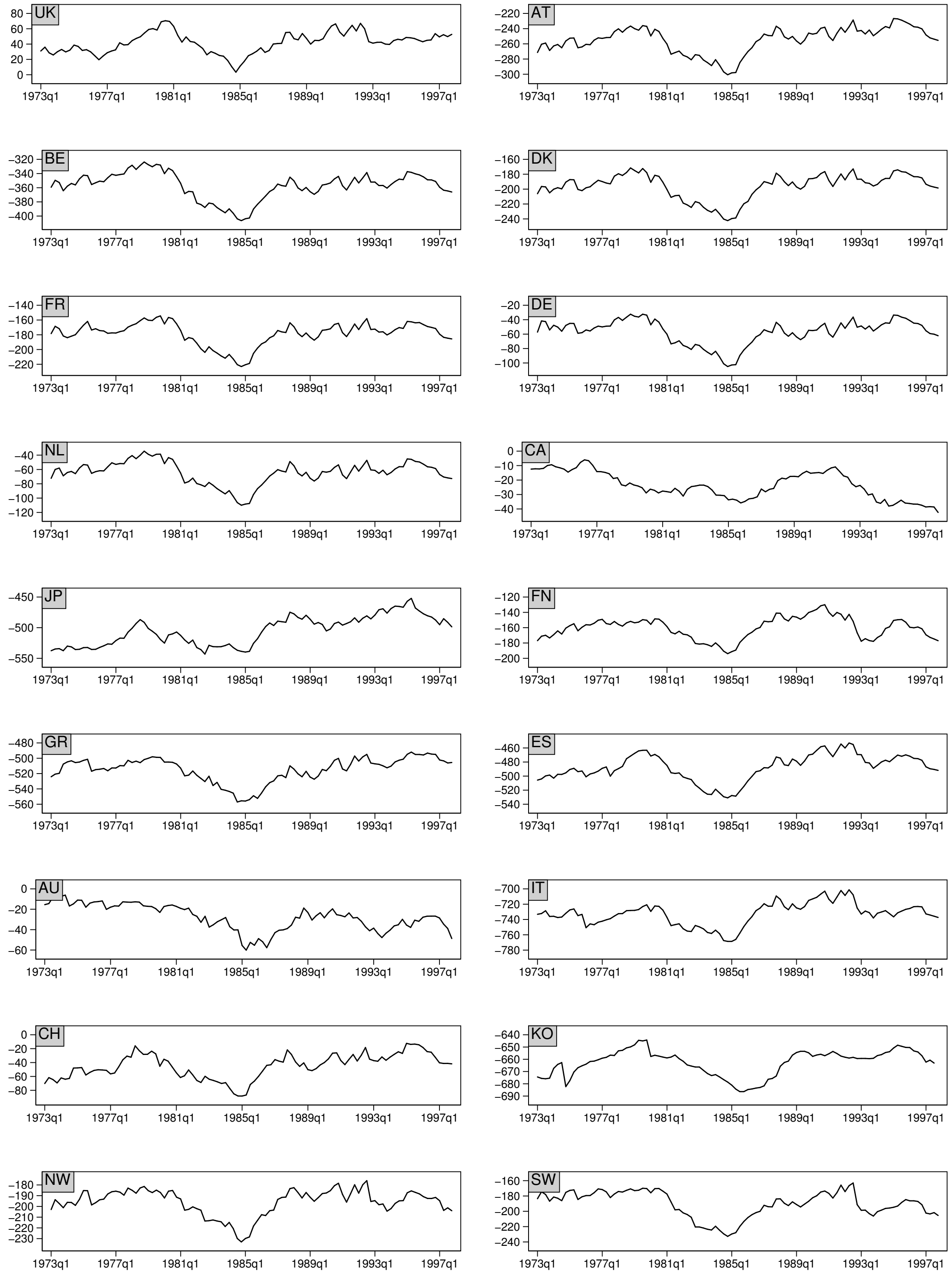

Figure A.1: PPP data with US$ as base currency

Table A.6: Estimated sectional correlation in the differenced PPP Data, US$ as base currency.

	UK	AT	BE	DK	FR	DE	NL	CA	JP	FN	GR	ES	AU	IT	CH	KO	NW	SW
UK	1.00																	
AT	$0.60_{0.10}$	1.00																
BE	$0.64_{0.10}$	$0.97_{0.02}$	1.00															
DK	$0.62_{0.10}$	$0.98_{0.01}$	$0.98_{0.01}$	1.00														
FR	$0.72_{0.09}$	$0.92_{0.04}$	$0.94_{0.03}$	$0.93_{0.04}$	1.00													
DE	$0.60_{0.10}$	$0.99_{0.01}$	$0.95_{0.03}$	$0.96_{0.02}$	$0.92_{0.04}$	1.00												
NL	$0.61_{0.10}$	$0.99_{0.01}$	$0.97_{0.02}$	$0.97_{0.01}$	$0.92_{0.04}$	$0.99_{0.01}$	1.00											
CA	$-0.01_{0.18}$	$0.03_{0.15}$	$0.00_{0.15}$	$0.07_{0.15}$	$-0.06_{0.15}$	$0.02_{0.16}$	$0.01_{0.16}$	1.00										
JP	$0.32_{0.17}$	$0.61_{0.13}$	$0.61_{0.13}$	$0.59_{0.13}$	$0.53_{0.15}$	$0.60_{0.13}$	$0.61_{0.13}$	$-0.05_{0.15}$	1.00									
FN	$0.65_{0.10}$	$0.75_{0.10}$	$0.75_{0.10}$	$0.77_{0.09}$	$0.73_{0.10}$	$0.71_{0.11}$	$0.75_{0.10}$	$0.25_{0.19}$	$0.36_{0.18}$	1.00								
GR	$0.64_{0.11}$	$0.83_{0.09}$	$0.80_{0.10}$	$0.83_{0.09}$	$0.76_{0.13}$	$0.82_{0.10}$	$0.83_{0.10}$	$0.14_{0.16}$	$0.41_{0.13}$	$0.70_{0.09}$	1.00							
ES	$0.68_{0.10}$	$0.81_{0.08}$	$0.82_{0.08}$	$0.84_{0.08}$	$0.83_{0.08}$	$0.79_{0.09}$	$0.81_{0.08}$	$0.09_{0.19}$	$0.35_{0.18}$	$0.78_{0.08}$	$0.79_{0.10}$	1.00						
AU	$0.30_{0.14}$	$0.26_{0.13}$	$0.24_{0.13}$	$0.26_{0.13}$	$0.25_{0.13}$	$0.25_{0.14}$	$0.25_{0.14}$	$0.41_{0.17}$	$0.25_{0.14}$	$0.44_{0.14}$	$0.39_{0.12}$	$0.33_{0.12}$	1.00					
IT	$0.73_{0.11}$	$0.79_{0.09}$	$0.79_{0.09}$	$0.80_{0.09}$	$0.84_{0.07}$	$0.78_{0.10}$	$0.78_{0.09}$	$0.09_{0.19}$	$0.37_{0.20}$	$0.70_{0.13}$	$0.73_{0.11}$	$0.81_{0.08}$	$0.32_{0.13}$	1.00				
CH	$0.60_{0.12}$	$0.91_{0.04}$	$0.90_{0.05}$	$0.91_{0.04}$	$0.88_{0.04}$	$0.90_{0.04}$	$0.91_{0.04}$	$0.02_{0.16}$	$0.65_{0.12}$	$0.69_{0.11}$	$0.77_{0.09}$	$0.73_{0.11}$	$0.24_{0.14}$	$0.71_{0.10}$	1.00			
KO	$0.02_{0.13}$	$0.18_{0.19}$	$0.17_{0.17}$	$0.17_{0.19}$	$0.13_{0.18}$	$0.17_{0.22}$	$0.16_{0.20}$	$0.30_{0.15}$	$0.24_{0.12}$	$0.24_{0.15}$	$0.22_{0.18}$	$0.23_{0.14}$	$0.53_{0.20}$	$0.09_{0.15}$	$0.15_{0.18}$	1.00		
NW	$0.71_{0.09}$	$0.89_{0.03}$	$0.88_{0.04}$	$0.90_{0.04}$	$0.86_{0.05}$	$0.87_{0.04}$	$0.89_{0.04}$	$0.17_{0.19}$	$0.51_{0.13}$	$0.86_{0.06}$	$0.80_{0.09}$	$0.81_{0.09}$	$0.41_{0.14}$	$0.77_{0.10}$	$0.81_{0.06}$	$0.19_{0.19}$	1.00	
SW	$0.71_{0.11}$	$0.76_{0.08}$	$0.80_{0.07}$	$0.80_{0.07}$	$0.80_{0.08}$	$0.73_{0.08}$	$0.76_{0.07}$	$0.15_{0.19}$	$0.36_{0.18}$	$0.83_{0.07}$	$0.70_{0.12}$	$0.86_{0.07}$	$0.39_{0.14}$	$0.81_{0.10}$	$0.67_{0.09}$	$0.15_{0.16}$	$0.86_{0.06}$	1.00
	UK	AT	BE	DK	FR	DE	NL	CA	JP	FN	GR	ES	AU	IT	CH	KO	NW	SW

Note: The values reported are correlation coefficients obtained after the correlation transformation of the estimator in (A.4) using $K = 3$. The standard errors reported are based on 500 replications of the bootstrap procedure described in Section A.3 and computed as in (A.5).