Testing for PPP:
Should We Use Panel Methods?[¤]

Anindya Banerjee[y]   Massimiliano Marcellino[z]   Chiara Osbat[x]

Abstract

A common ...nding in the empirical literature on the validity of purchasing power parity (PPP) is that it holds when tested for in panel data, but not in univariate (i.e. country speci...c) analysis. The usual explanation for this mis-match is that panel tests for unit roots and cointegration are more powerful than their univariate counterparts. In this paper we suggest an alternative ex-planation for the mismatch. More generally, we warn against the use of panel methods for testing for unit roots in macroeconomic time series. Existing panel methods assume that cross-unit cointegrating or long-run relationships, that tie the units of the panel together, are not present. However, using empirical examples on PPP for a panel of OECD countries, we show that this assumption is very likely to be violated. Simulations of the properties of panel unit root tests in the presence of long-run cross-unit relationships are then presented to demonstrate the serious cost of assuming away such relationships. The empirical size of the tests is substantially higher than the nominal level, so that the null hypothesis of a unit root is rejected very often, even if correct.

J.E.L. Classi...cation: C23, C33, F31

Keywords: PPP, unit root, panel, cointegration, cross-unit dependence

1 Introduction

The last few years have seen a veritable explosion of papers on testing for purchasing power parity (PPP) using panels of macroeconomic data. A selection would include inter alia Mark (1995), Frankel and Rose (1995), Jorion and Sweeney (1996), Oh (1996), Rogo¤ (1996), O'Connell (1998), Papell (1997), and more recently Bayoumi and MacDonald (1999), Pedroni (1999a) and the papers by Papell and co-authors given below.

The real dollar (or Deutsche mark) exchange rate for country i is constructed as:

$$q_{it} = e_{it} + p_t^{\pi} \; i \; p_{it};$$

where q_{it} is the logarithm of the real exchange rate, p_t^{π} is the logarithm of the US (or German) CPI and p_t is the logarithm of the CPI for country i. The strong form of the test for PPP consists of testing the null of a unit root in the q_{it} series, either individually (i.e. country by country) or using panel methods, discussed brie‡y in the next section.

More generally, a weak form of the hypothesis may be tested by constructing the series $\mathbf{q}_{it}$, where

$$\mathbf{q}_{it} = e_{it} + {}^{\circledR}p_t^{\pi} \; i \; {}^{-}p_{it}:$$

An acceptance of a unit root in the series $\mathbf{q}_{it}$ is taken to be a rejection of the weak form of the PPP hypothesis. The $^{\circledR}$ and $^{-}$ can be either taken to be known a priori or derived from single-equation- or system-cointegration methods, since all three component series of $\mathbf{q}_{it}$ are assumed to be I(1). The test of the weak form in panels may thus be seen to be tests for cointegration in panel data.

Among the large number of papers available in the literature, special attention should be drawn to the work of Papell and his co-authors who have investigated and drawn attention to a large number of issues that are relevant to a proper consideration of the evidence derived from panel tests.

The starting-o¤ point for the use of panel unit root and panel cointegration tests is the presumption that univariate (i.e country-speci...c) tests have low power. Within this framework, Papell (2000) for example has noted that the use of panel tests does not always rescue the various forms of the PPP hypothesis. In particular, Papell and Theodoridis (2000) ...nd that the choice of the numeraire currency is important in determining rejections or acceptance of a unit root in the real exchange rate series. Using a recursive estimation exercise, they show that the choice of the Deutsche mark as the numeraire currency leads to rejections of the unit root hypothesis for

1

sample sizes of steadily increasing lengths (with a ...xed starting point) much more strongly than when the dollar is taken to be the numeraire.

Papell (2000) argues in favour of the use of methods analogous to those employed by Perron (1989) in his much-cited paper that introduced the possibility of deterministic breaks in tests for unit roots. Papell showed that by modelling the appreciation and then depreciation of the dollar in the 1980's as shifts in the deterministic components of the series, stronger rejections of the unit root hypothesis in real exchange rate series are obtained. While the experience with such models in a purely time-series context does not make this a particularly surprising ...nding, it nevertheless alerts us to some of the subtleties that need to be tackled in determining whether or not weak- or strong-form PPP holds in the data.

The origin of our research in this paper is the evidence presented in Banerjee, Marcellino, and Osbat (2000) (BMO). There we investigated the properties of tests of cointegration in panels of data, particularly those proposed by Groen and Kleibergen (1999) and Larsson and Lyhagen (1999), and showed that not taking account of the cointegrating relationships across the units of the panel would lead to serious di¢culties in inference about cointegration within the units of a panel. These tests are extensions to panels of the Johansen approach to estimating cointegrating rank among time series, and the main restriction imposed is that there are no cointegrating relationships among the variables across the units (typically countries) of the panel. We showed that when this restriction is valid, the tests have the correct size and high power to detect cointegration. If the restriction is invalid however, the tests for cointegration tend to be grossly over-sized especially as T increases, so that the null of no cointegration is rejected too often in relation to the nominal con...dence level (or size) of the test.[1]

In the context of the equations above, the panel approach (both to testing for unit roots and cointegration) as currently advocated, rules out the existence, for example, of cointegrating relationships between e_{it} and e_{jt} or p_{it} and p_{jt}, for all $i \neq j$. In this paper we demonstrate the consequences of the violation of this assumption by looking at the properties of the panel unit root tests commonly used to test for stationarity of the real exchange rate. We show that over-sizing is a major problem also here. While the analysis in BMO is therefore more directly relevant for the consideration of the weak form of PPP, our main aim here is to explore further the arguments within the context of the strong form of the PPP hypothesis:

---

[1] This over-sizing property of panel unit root tests has also been discussed by Engel (2000) and O'Connell (1998) in slightly di¤erent contexts. Engel's paper does not deal with panels while O'Connell shows the e¤ect of short-run linkages among the units (for example through the non-zero covariances of the errors across the units) on unit root tests.

In the next section, we provide a brief overview of three of the most commonly used tests for unit roots in panels. These are the Levin and Lin (1992) and Levin and Lin (1993) (LL), Im, Pesaran, and Shin (1997) (IPS) and Maddala and Wu (1999 Special Issue) (MW) tests. More comprehensive summaries are provided inter alia by Baltagi and Kao (2000) and Banerjee (1999), and readers not interested in technical details may skip this section. Section 3 provides a set of motivating empirical examples demonstrating that while unit-by-unit ADF tests typically tend to accept the null of a unit root for the real exchange rate, panel unit root tests often reject this null. Such rejections must however be treated with a great deal of caution since there is strong evidence, detailed here, of cross-unit or cross-country cointegrating relationships. Therefore, in the light of our discussion above, these rejections instead of being attributed to the higher power of panel unit root tests may be attributed simply to the over-sizing that is present when cointegrating relationships link the units of the panel together. This claim is demonstrated in more detail in section 4 with a set of Monte Carlo experiments that analyze the size and power properties of the LL, IPS and MW tests in a series of cases. The cases studied include notably the presence or absence of cross-unit cointegration and the presence or absence of weak exogeneity of some of the variables. Section 5 o¤ers conclusions and closing remarks.

2 Testing for Unit Roots in Dynamic Panels

2.1 Levin and Lin (1992, 1993)

The model adopted by Levin and Lin allows for ...xed e¤ects and unit-speci...c time trends in addition to common time e¤ects (which may in practice be concentrated out of the equation) and is given by

$$\Phi y_{it} = \circledR_i + \pm_i t + \mu_t + \frac{1}{2}_i y_{it_i \ 1} + {}^3{}_{it}; \quad i = 1, 2, :N; t = 1, 2, ::::, T: \tag{1}$$

The null hypothesis of interest is H_0: $\frac{1}{2}_i = 0$ for all i against the alternative H_A: $\frac{1}{2}_i = \frac{1}{2} < 0$ for all i.

Levin and Lin (1992) derive the asymptotic distributions of the panel estimator of $\frac{1}{2}$ under di¤erent assumptions on the existence of ...xed e¤ects or heterogeneous time trends. For example, if ${}^3{}_{it} \gg IID(0; \frac{3}{4}^2)$ for ...xed i; the errors are also assumed to be independent across the units of the sample, $\circledR_i = \pm_i = 0$ for all i and there are no common time e¤ects, then the asymptotic distributions of the ordinary least

squares (OLS) pooled panel estimator and associated t-statistic are given by

$$T\sqrt{N}\hat{\delta} \Rightarrow N(0;2); \quad T;N \to \infty \tag{2}$$
$$t_{\delta=0} \Rightarrow N(0;1):$$

Note that, compared to unit-by-unit Dickey-Fuller tests, the LL statistics have (a) a faster rate of convergence to their asymptotic distributions and (b) the limiting distributions are normal.

In a slightly more general model, where $\alpha_i \neq 0; \pm_i = 0$ in (1),

$$T\sqrt{N}\hat{\delta} + 3\sqrt{N} \Rightarrow N(0;10.2); \quad \sqrt{N}=T \to 0; T;N \to \infty \tag{3}$$
$$\sqrt{1.25}\,t_{\delta=0} + \sqrt{1.875N} \Rightarrow N(0;1); \quad \sqrt{N}=T \to 0; T;N \to \infty:$$

The null hypothesis for this second model is given by $H_0 : \delta = 0; \alpha_i = 0$ for all i against the alternative $H_A : \delta < 0; \alpha_i$ unrestricted.

As a further generalisation, let us now consider the second of the two models discussed above with the error process following a stationary invertible ARMA process for each unit but still being distributed independently across units. Thus,

$$\epsilon_{it} = \sum_{j=0}^{\infty} \theta_{ij}\,\epsilon_{it-j} + \varepsilon_{it}; \quad t=1;2;...T: \tag{4}$$

A multi-step procedure is prescribed by Levin and Lin (1993) for this case. The starting point is the application of augmented Dickey-Fuller (ADF) test to each individual series. Thus, the regression

$$\Delta y_{it} = \alpha_i + \rho_i y_{it-1} + \sum_{j=1}^{p_i} \theta_{ij}\,\Delta y_{it-j} + \varepsilon_{it}; \quad i=1;2;:N; t=1;2;...T \tag{5}$$

is estimated (for each i) by regressing ...rst Δy_{it} and then y_{it-1} on the remaining regressors in the ADF regression above, providing the residuals $\hat{e}_{it}$ and $\hat{v}_{it-1}$ respectively. Then the regression of $\hat{e}_{it}$ on $\hat{v}_{it-1}$

$$\hat{e}_{it} = \rho_i \hat{v}_{it-1} + \varepsilon_{it} \tag{6}$$

is estimated to derive $\hat{\rho}_i$ from the i-th cross-section. The following expressions are next required:

4

$$\hat{\sigma}_{e_i}^2 = (T_i - p_i - 1)^{-1} \sum_{t=p_i+2} (\Delta y_{it} - \hat{\rho}_i \Delta y_{it-1})^2 \tag{7}$$

$$e_{it} = \Delta y_{it} / \hat{\sigma}_{e_i}$$

$$v_{it-1} = \Delta y_{it-1} / \hat{\sigma}_{e_i}$$

$$\hat{\sigma}_{y_i}^2 = (T_i - 1)^{-1} \sum_{t=2} \Delta y_{it}^2 + 2 \sum_{L=1}^{\bar{K}} w_{\bar{K}L} \left[(T_i - 1)^{-1} \sum_{t=L+2} \Delta y_{it} \Delta y_{it-L} \right]^2$$

$$\hat{S}_{NT} = N^{-1} \sum_{i=1}^{N} (\hat{\sigma}_{y_i} / \hat{\sigma}_{e_i});$$

where $\bar{K}$ is the lag truncation parameter and $w_{\bar{K}L} = (L / (\bar{K} + 1))$.

The ...nal step is to estimate the panel regression (making use of all i and t)

$$e_{it} = \rho v_{it-1} + \tilde{\varepsilon}_{it} \tag{8}$$

and compute the t-statistic

$$t_{\rho=0} = \frac{\hat{\rho}}{RSE(\hat{\rho})}; \tag{9}$$

where

$$RSE(\hat{\rho}) = \hat{\sigma}_{\tilde{\varepsilon}} \left[\sum_{i=1}^{N} \sum_{t=p_i+2} v_{it-1}^2 \right]^{-1/2} \tag{10}$$

$$\hat{\sigma}_{\tilde{\varepsilon}}^2 = (N\bar{T})^{-1} \sum_{i=1}^{N} \sum_{t=p_i+2} (e_{it} - \hat{\rho} v_{it-1})^2$$

$$\bar{p} = N^{-1} \sum_{i=1}^{N} p_i; \quad \bar{T} = (T - \bar{p} - 1):$$

The Levin and Lin statistic is an adjusted version of the t-statistic above and is given by

$$t_\rho^* = \frac{t_{\rho=0} - N\bar{T}\hat{S}_{NT}\hat{\sigma}_{\tilde{\varepsilon}}^{-2} RSE(\hat{\rho})\hat{\mu}_{\bar{T}}}{\frac{3}{4}_{\bar{T}}}; \tag{11}$$

where $\hat{\mu}_{\bar{T}}$ and $\frac{3}{4}_{\bar{T}}$ are mean and standard deviation adjustment terms which are computed by Monte Carlo simulation and tabulated in their paper for three separate speci...cations of the deterministic terms above.

The main result of the Levin and Lin analysis is that under the null hypothesis that $\rho = 0$, the panel test statistic t_ρ^* has the property that

$$t_\rho^* \Rightarrow N(0, 1); \quad T, N \to 1; \tag{12}$$

5

provided the ADF lag order p_{max} increases at some rate T^p where $0 < p \cdot 1=4$ and the lag truncation parameter $\overline{K}$ increases at rate T^q where $0 < q < 1$. Under the alternative hypothesis, $t_{\frac{1}{2}}^{\pi}$ diverges to negative in...nity at rate $p\overline{NT}$, a property that can thereby be exploited to construct a consistent one-tailed test of the null hypothesis against the alternative.

2.2   Im, Pesaran and Shin (1997)

Im, Pesaran, and Shin (1997) extend the Levin and Lin framework to allow for heterogeneity in the value of $\frac{1}{2}_i$ under the alternative hypothesis. Let

$$\Phi y_{it} = \circledR_i + \tfrac{1}{2}_i y_{it_i 1} + {}^3_{it}; \quad i = 1; 2; ::::; N; t = 1; 2; :T; \tag{13}$$

where the errors ${}^3_{it}$ are serially autocorrelated with di¤erent serial correlation (and variance) properties across the units. The null and alternative hypotheses are given by $H_0 : \tfrac{1}{2}_i = 0$ for all i; against the alternative $H_A : \tfrac{1}{2}_i < 0; i = 1; 2; ::::; N_1; \tfrac{1}{2}_i = 0; i = N_1 + 1; N_1 + 2; :::N:$

Following the critique of Pesaran and Smith (1995) on pooled panel estimators, such as those used by Levin and Lin (1992) and Levin and Lin (1993), Im, Pesaran, and Shin (1997) propose a group-mean Lagrange multiplier (LM) statistic. The ADF regressions

$$\Phi y_{it} = \circledR_i + \tfrac{1}{2}_i y_{it_i 1} + \sum_{j=1}^{X_i} \mu_{ij} \Phi y_{it_i j} + {}^{"}_{it}; \quad t = 1; 2; :::T \tag{14}$$

are estimated for each i and the LM-statistic for testing $\tfrac{1}{2}_i = 0$ is computed. De...ning

$$\overline{LM}_{N;T} = N^{i\,1} \sum_{i=1}^{X} LM_{iT}(p_i; \mu_i); \tag{15}$$

where $\mu_i = (\mu_{11}; \mu_{12}; :::\mu_{1p_i})^0$ and $LM_{iT}(p_i; \mu_i)$ is the individual LM-statistic for testing $\tfrac{1}{2}_i = 0$, the standardized LM-bar statistic is given by

$$a_{\overline{LM}} = \frac{p\overline{N}\,\overline{LM}_{N;T\;i}\; N^{i\,1} \sum_{i=1}^{N} E\,[LM_{iT}(p_i; 0 j\tfrac{1}{2}_i = 0]}{q\frac{}{N^{i\,1} \sum_{i=1}^{N} Var\,[LM_{iT}(p_i; 0 j\tfrac{1}{2}_i = 0]}}: \tag{16}$$

The values of $E\,[LM_{iT}(p_i; 0 j\tfrac{1}{2}_i = 0]$ and $Var\,[LM_{iT}(p_i; 0 j\tfrac{1}{2}_i = 0]$ are obtainable by stochastic simulation and are tabulated in their paper using 50,000 replications for di¤erent values of T and p_i's.

Im, Pesaran, and Shin (1997) show that under H_0: $\tfrac{1}{2}_i = 0$ for all i,

$$a_{\overline{LM}} =) \; N(0; 1) \tag{17}$$

6

as $T, N \to 1$ and $N=T \to k$ where k is a ...nite positive constant. For the test to be consistent under the alternative, it is also required that $\lim_{N_i \to 1} (N_i/N) = {}_{,1},$ $0 < {}_{,b}. 1$. Under this further assumption, $a_{\overline{LM}}$ diverges to positive in...nity at rate $T\sqrt{N}$ under the alternative.

Im, Pesaran, and Shin (1997) also propose the use of a group-mean t-bar statistic given by

$$
a_{\bar{t}} = \frac{\sqrt{N}\left\{ \bar{t}_{N,T} \ _i \ N^{\ i \ 1} \sum_{i=1}^{N} E\left[t_{iT}(p_i, 0 j\frac{1}{2}_i = 0\right]\right\}}{\sqrt{N^{\ i \ 1} \sum_{i=1}^{N} Var\left[t_{iT}(p_i, 0 j\frac{1}{2}_i = 0\right]}} \tag{18}
$$

where,

$$
\bar{t}_{N,T} = N^{\ i \ 1} \sum_{i=1}^{N} t_{iT}(p_i, \mu_i) \tag{19}
$$

and $t_{iT}(p_i, \mu_i)$ is the individual t-statistic for testing $\frac{1}{2}_i = 0$ for all i: $E\left[t_{iT}(p_i, 0 j\frac{1}{2}_i = 0\right]$ and $Var\left[t_{iT}(p_i, 0 j\frac{1}{2}_i = 0\right]$ are tabulated in the paper. The convergence result stated for $a_{\overline{LM}}$ holds for $a_{\bar{t}}$ also, and consistency is guaranteed under the controlled rate of divergence of N and T to in...nity.

In a Monte Carlo study, Im, Pesaran, and Shin (1997) demonstrate better ...nite sample performances of the $a_{\overline{LM}}$ and $a_{\bar{t}}$ tests in relation to the Levin and Lin test given by $t_{\frac{1}{2}}^{\alpha}$. We undertake a comparison of these tests in section 4 below for our various cases of interest. In the comparison we also include the Maddala and Wu test, which is described next.

2.3   Maddala and Wu (1999)

Maddala and Wu, relying on Fisher (1932), suggest combining the p-values of a test-statistic for a unit root in each cross-sectional unit. The Fisher test is an exact and non-parametric test, and may be computed for any arbitrary choice of a test for the unit root in a cross-sectional unit. The statistic is given by

$$
MW : \ _i \ 2 \sum_{i=1}^{N} \ln(\frac{1}{4}_i),
$$

and is distributed as a chi-squared variable with $2N$ degrees of freedom under the assumption of cross-sectional independence. $\frac{1}{4}_i$ is the p-value of the test statistic in unit i, the p-value of the ADF test statistic in the cases that we consider.

The obvious simplicity of this test and its robustness to the choice of lag length and sample size make its use attractive. However, our experience with the MW test is somewhat less encouraging, as reported below.

3 Testing for PPP

In this section we evaluate empirically whether the strong version of the purchasing power parity (PPP) holds, namely, whether the real exchange rate (rer) is stationary. This question has attracted considerable interest, as mentioned in the Introduction. Following Pedroni (1999b) and Mark (1995), we use quarterly data on nominal exchange rates and cpi, for the period 1973:1-1997:4, for 19 OECD countries.

The rers when US is used as the numeraire country are graphed in Figure 1. Note that the nominal exchange rates are expressed as US$ per national currency, so that a decline in the rer corresponds to a depreciation of the national currency. It is di¢cult to determine from the graphs whether the rers are stationary or not, but it is evident that most rers tend to move together, in particular for the countries currently in the European Monetary Union. This suggests that if the rers are integrated, they are also cointegrated across countries.

The augmented Dickey-Fuller (ADF) test accepts the null hypothesis of a unit root in the rer for each country and any choice of lag length, see Table 1.

The results of the panel unit root tests are reported in Table 2. Both versions of the IPS tests and the MW statistic reject the presence of a unit root at the 5% signi...cance level, while the LL test in general accepts. Overall, on the basis of the panel tests we would conclude that PPP holds. Yet, before drawing such a conclusion, we should check that the countries under analysis are not linked by cointegrating relationships, an important hypothesis underlying all panel unit root tests (for the tests to have correct size).

As mentioned earlier, it may be seen from Figure 1 that there appear to be substantial comovements of the rers across countries. To evaluate more formally whether the rers are cointegrated, we calculate the Johansen trace test for the rers of all pairs of countries, using in each case a VAR(4) with an unrestricted constant to model the variables (the outcome is not sensitive to the choice of the lag length). The results are reported in Table 3. In many cases the null hypothesis of no bivariate cointegration is rejected at the 10% and also at the 5% signi...cance level. Given that this test can be also expected to be biased toward acceptance of the null hypothesis in small samples, even more cross-country cointegration can be presumed to exist.[2]

Another interesting tool to analyze jointly the properties of large systems of possibly non-stationary variables was developed by Gonzalo and Granger (1995). It allows the determination of the total number of nonstationary trends driving all the variables under analysis. To implement this procedure, ...rst we extract the

---

[2]If the rers were stationary, the hypothesis of cointegrating rank equal to one versus rank equal to two should be also rejected, while it is not in most cases.

common trend from each of 7 bivariate systems that are cointegrated according to Table 3 (UK-Austria, Belgium-Netherlands, France-Germany, Japan-Australia, Finland-Norway, Spain-Italy, Greece-Switzerland). Then, we extract the trends from a 4-variable system for Denmark, Canada, Korea and Sweden. The rers for these countries turn out not to be cointegrated, hence we have 4 additional stochastic trends. Finally, we test for cointegration among the 11 trends (7 obtained in the ...rst step plus the 4 trends obtained in the second step) using the Johansen trace statistic in a VAR(3). The hypothesis of cointegrating rank equal to 5 is accepted at the 10% level, while that of rank equal to 4 is accepted at the 5% level. Hence, on the basis of the Gonzalo and Granger (1995) procedure, only 6 or 7 stochastic trends drive the joint evolution of the 18 rers. As a consequence, there exist 12 or 11 cointegrating relationships among the rers.

Both the bivariate cointegration analysis and the Gonzalo and Granger procedure indicate the presence of many cointegrating relationships across the countries, so that a basic assumption underlying the panel unit root tests is violated. The simulations in the next section show that in this case the panel tests can be substantially biased. The actual size is much larger than the nominal level, i.e. the null hypothesis of a unit root is rejected too often, and this can help to explain the con‡icting outcomes of the univariate and panel unit root tests. From the results of the simulation experiments in Tables 10 and 13 for T = 100, whose parameters are very close to those of our empirical analysis, it may be seen that LL is the least distorted panel test. Indeed this is the test which accepts the null hypothesis of a unit root at the 5% level for our empirical example, in agreement with the unit by unit ADF tests, and provides evidence against the validity of the strong version of PPP.[3]

According to Papell and Theodoridis (2000), the choice of the numeraire country can a¤ect the procedures to evaluate the presence of PPP in the long run. Hence, we repeat our analysis using Germany as the numeraire. The resulting rers are graphed in ...gure 2. Though there are some changes with respect to the graphs in Figure 1, the overall pattern is quite similar.

The ADF test still accepts the null hypothesis of a unit root in the rer for each country, see Table 1. The IPS and the MW panel tests reject this hypothesis, while the LL statistic accepts it, see Table 2. As for the case with US as the numeraire, both the graphs in ...gure 2 and the formal bivariate cointegration analysis indicate the presence of several cross-country cointegrating relationships among the rers,

---

[3]Alternative explanations for the failure of Levin and Lin tests to reject the null hypothesis of a unit root have been proposed by Papell (2000), as discussed in the introduction to this paper. In some sense these arguments must be weighed o¤ against those proposed here since these provide grounds for over-rejection of the null.

see Table 4, even though the pattern of bivariate cointegration is di¤erent when Germany is the numeraire. The presence of cross-country cointegrating relationships among the rers is also con…rmed by the Gonzalo and Granger method. Applying the step-wise procedure we described before, it turns out that in this case only 3 or 4 stochastic trends drive the evolution of the 18 rers.[4] Hence, the panel tests can be even more biased towards over-rejection in this case.

In summary, the unit by unit ADF tests favor the hypothesis of a unit root in the real exchange rates, while the IPS and MW panel unit root tests reject this hypothesis. This mismatch can be due to a bias in the panel tests, that reject too often the presence of a unit root when there are cross-unit cointegrating relationships. The panel LL test appears to be more robust in this case, and indeed it accepts the presence of a unit root, in agreement with the ADF tests. These …ndings are robust to the choice of the numeraire country, which instead in‡uences the number and composition of the cross-country cointegrating relationships.

4   Simulation results for panel unit root tests

In this section we describe the Monte Carlo design, and then evaluate the performance of the panel unit root tests, with and without cross-unit cointegrating relationships. This provides evidence on the reliability of the panel methods when applied to test for PPP and, more generally, to evaluate the stationarity of macroeconomic time series for several countries.

4.1   The Monte Carlo Design

Let us consider the variables y_{it}, x_{it}, where $i = 1, \ldots N$ indexes the units and $t = 1, \ldots, T$ is the temporal index. We group the y and x variables for each unit into the $N \pounds 1$ vectors $Y_t = (y_{1t}, \ldots, y_{Nt})'$ and $X_t = (x_{1t}, \ldots, x_{Nt})'$. Then we consider the following Error Correction Model (ECM) as the data generating process (DGP):

$$
\begin{pmatrix} \Phi Y_t \\ \Phi X_t \end{pmatrix} = \circledR^{-0} \begin{pmatrix} Y_{t_i 1} \\ X_{t_i 1} \end{pmatrix} + "_t, \tag{20}
$$

---

[4]We consider in the …rst step 7 bivariate VARs(4), for the pairs US-UK, Austria-Switzerland, France-Spain, Australia-Japan, Finland-Norway, Netherland-Sweden, and Greece-Denmark. Then we analyze a 4-variable VAR(4) for Belgium, Canada, Italy and Korea, but no cointegrating relationships are found. Finally, we run a VAR(3) for the 7 stochastic trends we obtain in the …rst step, plus the 4 trends from the second step. The cointegrating rank is 7 or 8 when using 5% or 10% critical values, so that 4 or 3 stochastic trends drive the 18 rers.

where $"_t$ is i:i:d: $N(0; \S)$, with

$$\S = \begin{pmatrix} \frac{3}{4}_1 & 0 & & 0 \\ 0 & \frac{3}{4}_2 & ::::: & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & & \frac{3}{4}_{2N} \end{pmatrix};$$

and $\frac{3}{4}_j$ is extracted from a uniform distribution on $[0:5; 1:5]$, for $j = 1; :::; 2N$.

We also considered DGPs with lagged di¤erences as regressors and serially cor-related errors. There were no qualitative di¤erences in the results, and hence we focus on the speci...cation in (20). The di¤erent experiments for size and power, with or without cross unit cointegration, are obtained by imposing proper restrictions on the ® and ¯ matrices, as detailed below. In all cases the results are based on 5000 replications, and are reported for di¤erent values of N (5, 10, 25, 50, 100) and of T (25, 50, 100), to mimic situations often encountered in empirical applications.

4.2   No cross-unit cointegration

The ...rst case we consider is testing for unit roots when there are no cross-unit coin-tegrating relationships, i.e., we are in a situation where the assumptions underlying panel tests hold. We are interested in evaluating the small sample size and power of the tests, to have a benchmark for the other cases.

For the size experiments we set ® = 0 in the DGP, so that X_t and Y_t are independent random walks. For the power experiments we use

$$® = \begin{pmatrix} I_{N\pounds N} & I_{N\pounds N} \\ 0_{N\pounds N} & 0_{N\pounds N} \end{pmatrix}; \quad ¯^0 = \begin{pmatrix} I_{N\pounds N} & 0_{N\pounds N} \\ c\, I_{N\pounds N} & 0_{N\pounds N} \end{pmatrix};$$

where c = 0:9 or c = 0:8. Hence, X_t is an N variate random walk, while Y_t is made up of N stationary AR(1) processes, with roots equal to c.

The estimated model for each unit is

$$\Phi y_{it} = °_0 y_{it_i\,1} + °_1 \Phi y_{it_i\,1} + °_2 \Phi y_{it_i\,2} + e_{it}: \tag{21}$$

This choice re‡ects the fact that when the lag length is underspeci...ed relative to the DGP the tests become very undersized, while overspeci...cation improves the size, even though the power can deteriorate slightly. The choice of the lag truncation in computing $\hat{\mathstrut}\sigma^2_{y_i}$ in (7) is also important in determining size, and we follow Levin and Lin's recommendations for comparability with existing results in the literature.

The empirical sizes of the unit root tests are reported in Table 5. We can read this table along three directions: for varying N, for varying T, and for both N and T varying. For ...xed T, in general the performance of the IPS-LM test improves when N increases, while it deteriorates for the IPS-t and, in particular, for the MW tests; the size is always decreasing for LL. When N is ...xed, in general the size distortions of all tests decrease with T, and this is particularly evident for MW. When both N and T vary, we consider the cases where N=T is about 0.25 (($N = 5$; $T = 20$), ($N = 10$; $T = 50$), ($N = 25$; $T = 100$)) or 0.5 (($N = 10$; $T = 20$), ($N = 25$; $T = 50$), ($N = 50$; $T = 100$)). The size improvements for larger dimensions are clear, in particular when N=T $= 0:5$. Overall, the IPS-LM has very low size distortions when $N > 25$ and $T > 50$, while the other three tests experience deviations from the nominal level also for these rather large values of units and temporal observations. It is worth noting that the LL test performs better when N is small, less than 25.

The power results are reported in Tables 6 and 7 for, respectively, $c = 0:8$ and $c = 0:9$. For $T = 20$ and N increasing, the IPS-t and the MW have the highest power, which is not surprising because of their large size distortions. The IPS-LM and the LL tests, with better size properties, have good power for N large when $c = 0:8$, but not for $c = 0:9$ (the power is less than 40% also for $N = 100$). When $T = 50$, the situation improves substantially, in particular for the IPS-LM and already for $N = 25$, also with $c = 0:9$. When $T = 100$ the power is close or equal to one for all tests, with the exception of LL when $N = 5$; 10 and $c = 0:9$. For ...xed N, the power increases substantially with T already for $N = 5$, and a similar result holds when N and T vary jointly.

Overall, these ...gures indicate that the temporal dimension is quite important for high power also in a panel context, and that the IPS-LM test has not only good empirical size but also rather high power, particularly when $N > 25$ and $T > 50$.[5]

4.3   Cross-unit cointegration, weak exogeneity

We now consider the following speci...cation of the DGP, to allow for cross-unit cointegrating relationships to be present:

$$
® = 0:1 @ \begin{matrix} O & 1 \\ i\,I_{N£N} & 0_{N£N} \\ 0_{N£N} & i\,I_{N£N} \end{matrix} A ; \quad -^0 = @ \begin{matrix} O & 1 \\ i\,I_{N£N} & I_{N£N} \\ 0_{N£N} & B_{N£N} \end{matrix} A ;
$$

---

[5]We also considered di¤erent values for c across the units. This violates the condition of the LL test on the alternative hypothesis, and the LL test turns out to have systematically lower power than the IPS tests and MW in this case.

$$
B = \begin{pmatrix}
1 & -1 & 0 & 0 & 0 & 0 & \cdots & 0 \\
0 & 1 & -1 & 0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 1 & -1 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\
\cdots & & & & & & & \\
0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0
\end{pmatrix} :
\tag{22}
$$

Hence, all the variables are I (1), but the system is driven by a number of independent random walks equal to the number of zero rows of B, say b. For example, in the formulation in (22), $b = N - 3$. The remaining $2N - b$ stationary roots of the system are equal to 0:9. There exist N within unit cointegrating relationships, $(y_{it} - x_{it})$, $i = 1, \ldots, N$, plus $N - b$ cross-unit cointegrating relationships of the type $(x_{it} - x_{i+1t})$, $i = 1, \ldots N - b$. Moreover, the x_i variables are weakly exogenous in the subsystem for the y_i variables.[6] In BMO, weak exogeneity was an important factor in reducing the size distortions, which is our reason for considering it here. The next subsection deals with the case of cross unit-relationships but without weak exogoneity.

The estimated model for each unit remains

$$
\Delta y_{it} = \delta_0 y_{it-1} + \delta_1 \Delta y_{it-1} + \delta_2 \Delta y_{it-2} + e_{it},
\tag{23}
$$

and we are interested in evaluating if and how the presence of cross-unit cointegration affects the performance of the panel unit root tests. Note that each y_i is I (1), so that the empirical rejection frequencies of the tests should be around 5%.

The results are reported in Tables 8 to 12, for different values of N and varying percentages of cross-unit cointegration, q, where $q = ((N - b)/N) \times 100$. The parameter q typically ranges from 20% to 80%, i.e., from 20% to 80% of the units are related by bivariate cointegrating relationships. Five points are worth making. First, the distortions for IPS-t and MW are substantial in all cases, with values often in the range 20% - 50% when $N > 25$ and $T = 20$. Second, the IPS-LM performs reasonably well for $T = 20$, while the LL test appears to be quite robust to the presence of cross-unit cointegration, with only minor distortions. Third, focusing on IPS-LM and LL, when T increases the distortions increase. Fourth, the distortions decrease for larger N in the case of LL, and do not increase with N for IPS-LM. Fifth, for $N > 25$ the distortions increase with q, while for fewer units the outcome is less clear cut.

---

[6] Essentially this requires that the cointegrating relationships between y_i and x_i do not affect the x_i variables. See Johansen (1995) for a more precise definition of weak exogeneity in cointegrated systems.

4.4 Cross-unit cointegration, no weak exogeneity

We still want all the variables to be $I(1)$, with the system driven by fewer than 2N stochastic trends and the x_i variables no longer weakly exogenous in the subsystem for the y_i variables. Hence, we consider the following speci…cation of the DGP:

$$\Phi = 0.1 \begin{pmatrix} I_{N \times N} & 0_{N \times N} \\ A_{N \times N} & I_{N \times N} \end{pmatrix}; \quad \Theta = \begin{pmatrix} I_{N \times N} & I_{N \times N} \\ B_{N \times N} & 0_{N \times N} \end{pmatrix};$$

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ \cdots & & & & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}; \quad B = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ \cdots & & & & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix};$$

$$\tag{24}$$

There are $2N - b$ stationary roots equal to 0.9, where we recall that b is the number of zero rows of B (i.e., the system is driven by b stochastic trends). The stationary roots are associated with N within unit cointegrating relationships, $(y_{it} - x_{it})$, $i = 1, \ldots, N$, plus $N - b$ cross-unit cointegrating relationships of the type $(y_{it} - y_{i+1t})$, $i = 1, \ldots N - b$. The cointegrating relationship between y_i and y_{i+1} a¤ects x_i, which is therefore not weakly exogenous.

The estimated model for each unit remains that in equation (21).

The results are summarized in Table 13. Ranking the tests in terms of size distortions, MW is the worst, followed by IPS-t, which though performs much better than in the weakly exogenous case. IPS-LM is reasonably well behaved for $T = 20$ and $T = 50$, but still presents substantial size distortions for $T = 100$, that increase with the number of units, N. The LL statistic appears again to be the most robust, with empirical size in the range 3%-8% also for $T = 100$ and $N = 100$. It may also be noted by comparing Tables 8 to 12 with Table 13 that the lack of weak exogeneity does not pose any special problem in the context of our analysis here, while the consequences were much more serious for the results in BMO.

5 Conclusions

The empirical analysis in our paper demonstrates clearly the importance of taking proper account of the presence of cross-unit cointegrating relationships in inter-

preting the results of unit root tests in panel. These cross-unit relationships are detectable by means of bivariate or small-system cointegration analysis and by the Gonzalo and Granger (1995) procedure on the estimated common trends. Using the latter, the amount of cross-unit cointegration in the PPP panel in section 3 is found to be around 50% with US as the numeraire and around 80% with Germany as the numeraire country. Looking at the relevant entries of Tables 8-13, for T=100, N=25 and q=50% or 80%, the size-distortions are in the range of 0.07 to 0.32 when x_t is weakly exogenous. When x_t is not weakly exogenous, the corresponding numbers are 0.03 and 0.24 respectively. It should also be noted that of all the tests considered, the LL test appears to su¤er the least from size distortion in the presence of cross-unit cointegrating relationships, but the DGPs considered here are on the whole those most favourable to the LL testing framework because of the homogeneity across the units of the autoregressive parameter. In some results not reported here, we also ...nd that when heterogeneity is allowed for in the DGP, the power of the LL test is lower than for the other panel tests considered.

Overall, our results provide yet a further serious warning against the use of standard panel unit root tests but also a stimulus for further research. In particular, the aim should be to understand the theoretical properties of these tests under various relaxations of the underlying assumption of cross-unit long run independence and to devise procedures that are robust to such generalisations.

References

Baltagi, B., and C. Kao (2000): "Nonstationary Panels, Cointegration in Panels and Dynamic Panels: A Survey," mimeo.

Banerjee, A. (1999): "Panel Data Unit Roots and Cointegration: An Overview," Oxford Bulletin of Economic and Statistics, 61, 607–630.

Banerjee, A., M. Marcellino, and C. Osbat (2000): "Some Cautions on the Use of Panel Methods for Integrated Series of Macro-Economic Data," Discussion Paper working paper n.170, IGIER.

Bayoumi, T., and R. MacDonald (1999): "Deviations of Exchange Rates from Purchasing Power Parity: A Story Featuring Two Monetary Unions," Discussion Paper 46, 1, IMF Sta¤ Papers.

Engel, C. (2000): "Long-Run PPP May Not Hold After All," Journal of International Economics, 57, 243–273.

Fisher, R. (1932): Statistical Methods for Research Workers. Oliver and Boyd, Edimburgh.

Frankel, J. A., and A. K. Rose (1995): "Empirical Research on Nominal Exchange Rates," in Handbook of International Economics Vol. 3, ed. by G. M. Grossman, and K. Rogo¤, pp. 1689–1729. Elsevier Science, Amsterdam.

Gonzalo, J., and C. W. J. Granger (1995): "Estimation of Common Long-Memory Components in Cointegrated Systems," Journal of Business and Economic Statistics, 13, 27–36.

Groen, J. J., and F. R. Kleibergen (1999): "Likelihood-Based Cointegration Analysis in Panels of Vector Error Correction Models," Discussion Paper Discussion Paper 55, Tinbergen Institute.

Im, K. S., M. H. Pesaran, and Y. Shin (1997): "Testing for Unit Roots in Heterogeneous Panels," Discussion paper, Department of Applied Economics, University of Cambridge.

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

Jorion, P., and R. Sweeney (1996): "Mean Reversion in Real Exchange Rates: Evidence and Implications for Forecasting," Journal of International Money and Finance, August, 535–550.

Larsson, R., and J. Lyhagen (1999): "Likelihood-Based Inference in Multivariate Panel Cointegration Models," Discussion Paper Working Paper Series in Economics and Finance 331, Stockholm School of Economics and Finance.

Levin, A., and C. Lin (1992): "Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties," Discussion Paper 92-93, Department of Economics, University of California at San Diego.

Levin, A., and C. F. Lin (1993): "Unit Root Tests in Panel Data: New Results," Discussion Paper Discussion Paper 56, University of California at San Diego.

Maddala, G., and S. Wu (1999 Special Issue): "A Comparative Study of Unit Root Tests and a New Simple Test," Oxford Bulletin of Economic and Statistics, 61, 631–652.

Mark, N. (1995): "Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability," American Economic Review, 85, 201–218.

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

Oh, K. (1996): "Purchasing Power Parity and Unit Root Tests Using Panel Data," Journal of International Money and Finance, 15, 405–418.

Papell, D. (1997): "Searching for Stationarity: Purchasing Power Parity Under the Current Float," Journal of International Economics, 43, 313–332.

Papell, D. H. (2000): "The Great Appreciation, the Great Depreciation, and the Purchasing Power Parity Hypothesis," Discussion paper, University of Huston.

Papell, D. H., and H. Theodoridis (2000): "The Choice of Numeraire Currency in Panel Tests of Purchasing Power Parity," Forthcoming: the Journal of Money, Credit and Banking.

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

Pedroni, P. (1999b): "Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors," Oxford Bulletin of Economics and Statistics, 61 Special Issue, 653–70.

Perron, P. (1989): "The Great Crash, the Oil Price Shock and the Unit Root Hypothesis," Econometrica, 58, 1361–1401.

Pesaran, H., and R. Smith (1995): "Estimating Long-Run Relationships from Dynamic Heterogeneous Panels," Journal of Econometrics, 68, 79–113.

Rogoff, K. (1996): "The Purchasing Power Parity Puzzle," Journal of Economic Literature, 34(2), 647–668.

Tables

Table 1: Univariate ADF unit root tests for real exchange rates

lags	US numeraire				Germany numeraire			
	1	2	3	4	1	2	3	4
US	i	i	i	i	¡ 2:96¤	¡ 3:04¤	¡ 2:41	¡ 2:26
UK	¡ 2:35	¡ 2:01	¡ 2:41	¡ 2:48	¡ 1:81	¡ 1:71	¡ 1:90	¡ 1:94
Austria	¡ 1:90	¡ 1:67	¡ 2:04	¡ 2:43	¡ 1:76	¡ 1:63	¡ 1:81	¡ 1:97
Belgium	¡ 1:76	¡ 1:54	¡ 2:04	¡ 2:36	¡ 1:98	¡ 1:81	¡ 2:22	¡ 2:45
Denmark	¡ 1:95	¡ 1:63	¡ 2:12	¡ 2:32	¡ 1:83	¡ 1:64	¡ 1:92	¡ 2:02
France	¡ 2:01	¡ 1:84	¡ 2:18	¡ 2:53	¡ 2:10	¡ 1:99	¡ 2:23	¡ 2:46
Germany	¡ 1:93	¡ 1:72	¡ 2:11	¡ 2:47	i	i	i	i
Netherland	¡ 2:02	¡ 1:71	¡ 2:15	¡ 2:58	¡ 2:09	¡ 1:86	¡ 2:14	¡ 2:38
Canada	¡ 0:95	¡ 0:95	¡ 1:47	¡ 1:34	¡ 1:50	¡ 1:52	¡ 1:71	¡ 1:72
Japan	¡ 1:69	¡ 1:61	¡ 1:79	¡ 2:06	¡ 1:62	¡ 1:59	¡ 1:62	¡ 1:70
Finland	¡ 2:06	¡ 1:88	¡ 2:53	¡ 2:92¤	¡ 2:14	¡ 2:08	¡ 2:32	¡ 2:47
Greece	¡ 1:55	¡ 1:48	¡ 1:52	¡ 2:28	¡ 1:13	¡ 1:13	¡ 1:16	¡ 1:62
Spain	¡ 1:83	¡ 1:88	¡ 2:28	¡ 2:36	¡ 1:90	¡ 1:92	¡ 2:13	¡ 2:14
Australia	¡ 1:81	¡ 1:74	¡ 1:96	¡ 1:90	¡ 1:98	¡ 1:98	¡ 2:32	¡ 2:29
Italy	¡ 2:14	¡ 2:00	¡ 2:15	¡ 2:51	¡ 1:72	¡ 1:65	¡ 1:73	¡ 1:95
Switzerland	¡ 2:19	¡ 2:01	¡ 2:26	¡ 2:67	¡ 1:98	¡ 1:90	¡ 1:98	¡ 2:13
Korea	¡ 0:84	¡ 0:77	¡ 0:99	¡ 1:24	¡ 1:45	¡ 1:45	¡ 1:44	¡ 1:45
Norway	¡ 2:11	¡ 1:76	¡ 2:11	¡ 2:18	¡ 1:96	¡ 1:84	¡ 2:01	¡ 2:01
Sweden	¡ 1:83	¡ 1:63	¡ 2:24	¡ 2:04	¡ 1:96	¡ 1:83	¡ 2:41	¡ 2:21

¤ and ¤¤ indicate rejection at the 5% and 1% of the null hypothesis of a unit root

Table 2: Panel unit root tests for real exchange rates

lags	US numeraire			
	1	2	3	4
IPS-LM	3.47	2.42	2.43	2.83
	(0.0003)	(0.008)	(0.008)	(0.002)
IPS-t	−3.56	−2.91	−3.00	−3.39
	(0.0002)	(0.002)	(0.001)	(0.0003)
LL	−2.23	−1.38	−1.36	−1.58
	(0.013)	(0.083)	(0.087)	(0.057)
MW	66.57	55.74	56.04	59.51
	(0.001)	(0.019)	(0.018)	(0.008)
	Germany numeraire			
IPS-LM	3.16	2.03	2.37	2.90
	(0.0008)	(0.021)	(0.009)	(0.002)
IPS-t	−3.34	−2.63	−2.97	−3.46
	(0.0004)	(0.004)	(0.001)	(0.0003)
LL	−1.73	−0.86	−1.17	−1.50
	(0.042)	(0.195)	(0.120)	(0.066)
MW	63.40	51.90	55.19	60.06
	(0.003)	(0.042)	(0.021)	(0.007)

p-values in parentheses

Table 3: Rejection of the null hypothesis of no cointegration (rank=0) using Johansen trace statistic, US is numeraire

	UK	Au	Be	Fr	Ge	Ne	Ca	Ja	Fi	Gr	Sp	Aul	It	Sw	Ko	No	Swe	De
UK		**	**	*	*	*	–	**	**	–	**	–	**	**	–	–	*	**
Au	**		–	–	**	–	–	–	–	**	–	–	–	**	–	–	–	–
Be	**	–		*	*	*	–	–	–	*	–	–	–	–	–	–	–	–
Fr	*	–	*		**	**	–	–	–	*	–	–	–	–	–	–	–	**
Ge	*	**	*	**		–	–	–	–	**	–	–	–	*	–	–	–	*
Ne	*	–	*	**	–		–	–	–	–	–	*	–	–	–	–	–	–
Ca	–	–	–	–	–	–		–	–	–	–	–	–	–	–	–	–	–
Ja	**	–	–	–	–	–	–		–	*	–	*	–	*	–	–	–	–
Fi	**	–	–	–	–	–	–	–		**	–	**	**	–	*	**	–	–
Gr	–	**	*	*	**	–	–	*	**		–	–	*	**	–	–	–	**
Sp	**	–	–	–	–	–	–	–	–	–		–	**	*	–	–	–	–
Aul	–	–	–	–	–	*	–	*	**	–	–		–	–	–	–	–	–
It	**	–	–	–	–	–	–	–	**	*	**	–		–	–	–	–	–
Sw	**	**	–	–	*	–	–	*	–	**	*	–	–		–	–	–	–
Ko	–	–	–	–	–	–	–	–	*	–	–	–	–	–		–	–	–
No	–	–	–	–	–	–	–	–	**	–	–	–	–	–	–		–	–
Swe	*	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–		–
De	**	–	–	**	*	–	–	–	–	**	–	–	–	–	–	–	–	

** and * indicate rejection at 5% and 10%. Results are based on a VAR(4) with unrestricted constant.

Table 4: Rejection of the null hypothesis of no cointegration (rank=0) using Johansen trace statistic, Germany is numeraire

	UK	Au	Be	Fr	US	Ne	Ca	Ja	Fi	Gr	Sp	Aul	It	Sw	Ko	No	Swe	De
UK		–	–	–	**	–	–	**	**	–	–	*	–	*	–	–	–	–
Au	–		–	**	–	–	–	*	–	**	–	**	–	**	–	–	–	–
Be	–	–		–	–	–	–	–	–	**	–	**	–	–	–	–	–	–
Fr	–	**	–		–	**	–	*	–	**	*	**	–	*	–	–	–	**
US	**	–	–	–		–	–	*	–	–	–	**	–	–	–	–	*	–
Ne	–	–	–	**	–		–	–	–	**	–	**	–	–	–	–	*	–
Ca	–	–	–	–	–	–		–	*	–	–	–	–	–	–	–	–	–
Ja	**	*	–	*	*	–	–		*	**	*	**	–	**	–	*	–	–
Fi	**	–	–	–	–	–	*	*		*	*	**	**	–	–	*	–	–
Gr	–	**	**	**	–	**	–	**	*		–	–	–	**	–	–	*	**
Sp	–	–	–	*	–	–	–	*	*	–		**	–	–	–	**	–	–
Aul	*	**	**	**	**	**	–	**	**	–	**		–	**	**	**	**	**
It	–	–	–	–	–	–	–	–	**	–	–	–		–	–	*	–	–
Sw	*	**	–	*	–	–	–	**	–	**	–	**	–		–	–	–	–
Ko	–	–	–	–	–	–	–	–	–	–	–	**	–	–		–	*	–
No	–	–	–	–	–	–	–	*	*	–	**	**	*	–	–		–	–
Swe	–	–	–	–	*	*	–	–	–	*	–	**	–	–	*	–		–
De	–	–	–	**	–	–	–	–	–	**	–	**	–	–	–	–	–	

** and * indicate rejection at 5% and 10%. Results are based on a VAR(4) with unrestricted constant.

Table 5: Unit root tests, no cross-unit cointegration - Size

N/T		20	50	100
5	IPS LM	0:067	0:073	0:070
	IPS t	0:137	0:083	0:071
	LL	0:091	0:056	0:060
	MW	0:185	0:092	0:067
10	IPS LM	0:066	0:073	0:073
	IPS t	0:132	0:081	0:072
	LL	0:090	0:055	0:061
	MW	0:181	0:097	0:070
25	IPS LM	0:041	0:056	0:055
	IPS t	0:189	0:098	0:069
	LL	0:049	0:039	0:044
	MW	0:325	0:118	0:075
50	IPS LM	0:041	0:052	0:050
	IPS t	0:271	0:109	0:071
	LL	0:042	0:031	0:037
	MW	0:480	0:143	0:081
100	IPS LM	0:031	0:047	0:049
	IPS t	0:382	0:129	0:077
	LL	0:023	0:019	0:029
	MW	0:674	0:176	0:097

Table 6: Unit root tests, no cross-unit cointegration - Power, roots=0.8

N/T		20	50	100
5	IPS LM	0:165	0:780	1:000
	IPS t	0:324	0:844	1:000
	LL	0:176	0:465	0:944
	MW	0:329	0:767	1:000
10	IPS LM	0:227	0:970	1:000
	IPS t	0:532	0:990	1:000
	LL	0:232	0:748	0:999
	MW	0:511	0:967	1:000
25	IPS LM	0:412	1:000	1:000
	IPS t	0:859	1:000	1:000
	LL	0:380	0:987	1:000
	MW	0:797	1:000	1:000
50	IPS LM	0:661	1:000	1:000
	IPS t	0:988	1:000	1:000
	LL	0:586	1:000	1:000
	MW	0:966	1:000	1:000
100	IPS LM	0:903	1:000	1:000
	IPS t	1:000	1:000	1:000
	LL	0:822	1:000	1:000
	MW	0:992	1:000	1:000

Table 7: Unit root tests, no cross-unit cointegration - Power, roots=0.9

N/T		20	50	100
5	IPS LM	0:100	0:338	0:870
	IPS t	0:218	0:413	0:905
	LL	0:128	0:161	0:402
	MW	0:252	0:346	0:826
10	IPS LM	0:114	0:523	0:992
	IPS t	0:305	0:690	0:998
	LL	0:133	0:241	0:717
	MW	0:348	0:542	0:986
25	IPS LM	0:156	0:870	1:000
	IPS t	0:553	0:978	1:000
	LL	0:181	0:502	0:987
	MW	0:569	0:890	1:000
50	IPS LM	0:228	0:991	1:000
	IPS t	0:780	1:000	1:000
	LL	0:231	0:800	1:000
	MW	0:793	0:993	1:000
100	IPS LM	0:374	1:000	1:000
	IPS t	0:971	1:000	1:000
	LL	0:333	0:978	1:000
	MW	0:965	1:000	1:000

Table 8: Unit root tests, cross-unit cointegration, weak exogeneity - N=5

q/T		20	50	100
20%	IPS LM	0.091	0.118	0.132
	IPS t	0.176	0.140	0.136
	LL	0.101	0.080	0.087
	MW	0.205	0.132	0.122
40%	IPS LM	0.176	0.328	0.363
	IPS t	0.292	0.361	0.369
	LL	0.117	0.126	0.144
	MW	0.303	0.335	0.340
60%	IPS LM	0.081	0.165	0.338
	IPS t	0.165	0.198	0.358
	LL	0.103	0.095	0.121
	MW	0.208	0.178	0.313
80%	IPS LM	0.080	0.193	0.432
	IPS t	0.171	0.240	0.473
	LL	0.103	0.105	0.158
	MW	0.207	0.194	0.393

q: percentage of cross-unit cointegration

$(q = ((N-b)/N)*100)$

Table 9: Unit root tests, cross-unit cointegration, weak exogeneity - N=10

q/T		20	50	100
20%	IPS LM	0:055	0.076	0.110
	IPS t	0:157	0.105	0.123
	LL	0:071	0.048	0.070
	MW	0:237	0.113	0.120
50%	IPS LM	0.050	0.091	0.175
	IPS t	0.158	0.128	0.199
	LL	0.070	0.055	0.084
	MW	0.232	0.126	0.174
70%	IPS LM	0.066	0.148	0.347
	IPS t	0.183	0.211	0.398
	LL	0.083	0.071	0.119
	MW	0.255	0.183	0.324
90%	IPS LM	0.062	0.132	0.271
	IPS t	0.177	0.196	0.337
	LL	0.082	0.065	0.116
	MW	0.240	0.162	0.251

q: percentage of cross-unit cointegration

Table 10: Unit root tests, cross-unit cointegration, weak exogeneity - N=25

q/T		20	50	100
25%	IPS LM	0:044	0:081	0:133
	IPS t	0:202	0:137	0:160
	LL	0:051	0:046	0:060
	MW	0:340	0:145	0:161
50%	IPS LM	0:049	0:118	0:246
	IPS t	0:227	0:206	0:310
	LL	0:059	0:052	0:073
	MW	0:359	0:192	0:255
70%	IPS LM	0:053	0:129	0:238
	IPS t	0:227	0:210	0:316
	LL	0:064	0:055	0:078
	MW	0:362	0:188	0:245
90%	IPS LM	0:054	0:134	0:265
	IPS t	0:246	0:235	0:356
	LL	0:069	0:064	0:087
	MW	0:376	0:201	0:266

q: percentage of cross-unit cointegration

Table 11: Unit root tests, cross-unit cointegration, weak exogeneity - N=50

q/T		20	50	100
25%	IPS LM	0:043	0:081	0:108
	IPS t	0:287	0:156	0:159
	LL	0:044	0:039	0:04
	MW	0:500	0:177	0:144
50%	IPS LM	0:051	0:114	0:190
	IPS t	0:317	0:232	0:273
	LL	0:049	0:041	0:051
	MW	0:513	0:219	0:221
70%	IPS LM	0:053	0:147	0:301
	IPS t	0:343	0:298	0:425
	LL	0:051	0:053	0:065
	MW	0:523	0:260	0:326
90%	IPS LM	0:060	0:183	0:365
	IPS t	0:364	0:354	0:504
	LL	0:060	0:060	0:084
	MW	0:535	0:296	0:381

q: percentage of cross-unit cointegration

Table 12: Unit root tests, cross-unit cointegration, weak exogeneity - N=100

q/T		20	50	100
25%	IPS LM	0:037	0:069	0:112
	IPS t	0:407	0:187	0:178
	LL	0:028	0:023	0:032
	MW	0:694	0:212	0:165
50%	IPS LM	0:044	0:115	0:242
	IPS t	0:464	0:290	0:360
	LL	0:036	0:034	0:042
	MW	0:719	0:281	0:298
75%	IPS LM	0:059	0:198	0:444
	IPS t	0:521	0:467	0:631
	LL	0:046	0:053	0:073
	MW	0:745	0:396	0:485
90%	IPS LM	0:064	0:228	0:487
	IPS t	0:540	0:504	0:668
	LL	0:051	0:057	0:075
	MW	0:757	0:416	0:522

q: percentage of cross-unit cointegration

Table 13: Unit root tests, cross-unit cointegration, no weak exogeneity

N/T		q=50%			q=80%		
		20	50	100	20	50	100
5 [a]	IPS LM	0.077	0.069	0.121	0.081	0.098	0.235
	IPS t	0.068	0.097	0.124	0.072	0.137	0.260
	LL	0.091	0.054	0.057	0.092	0.060	0.067
	MW	0.179	0.095	0.112	0.197	0.128	0.209
10	IPS LM	0.063	0.054	0.100	0.062	0.056	0.127
	IPS t	0.055	0.095	0.121	0.060	0.103	0.163
	LL	0.067	0.039	0.042	0.070	0.032	0.037
	MW	0.229	0.097	0.099	0.231	0.100	0.122
25	IPS LM	0.050	0.055	0.114	0.058	0.062	0.171
	IPS t	0.046	0.127	0.157	0.055	0.167	0.241
	LL	0.050	0.034	0.038	0.058	0.033	0.035
	MW	0.348	0.133	0.129	0.365	0.153	0.181
50	IPS LM	0.060	0.066	0.176	0.071	0.062	0.376
	IPS t	0.054	0.208	0.268	0.068	0.224	0.546
	LL	0.049	0.031	0.036	0.056	0.023	0.049
	MW	0.515	0.205	0.214	0.538	0.185	0.404
100	IPS LM	0.071	0.060	0.252	0.077	0.116	0.478
	IPS t	0.057	0.307	0.417	0.071	0.485	0.726
	LL	0.033	0.024	0.032	0.041	0.038	0.040
	MW	0.715	0.289	0.322	0.739	0.394	0.528

q: percentage of cross-unit cointegration, (q = ((N-b)/N)*100)
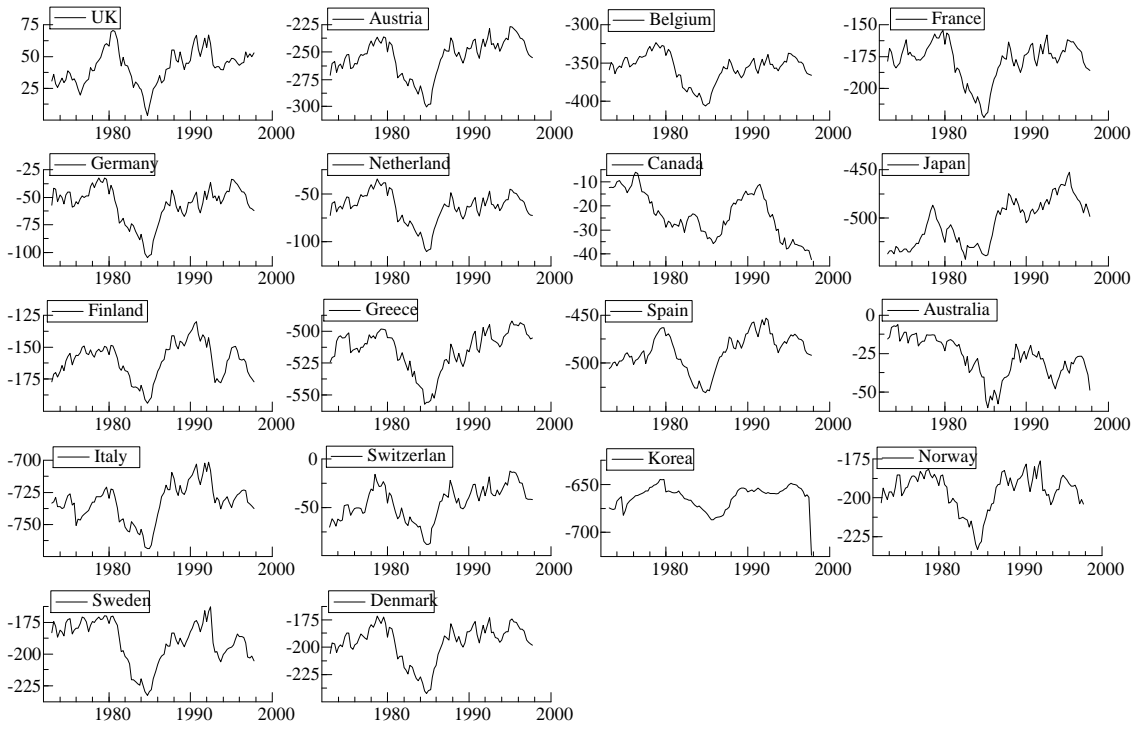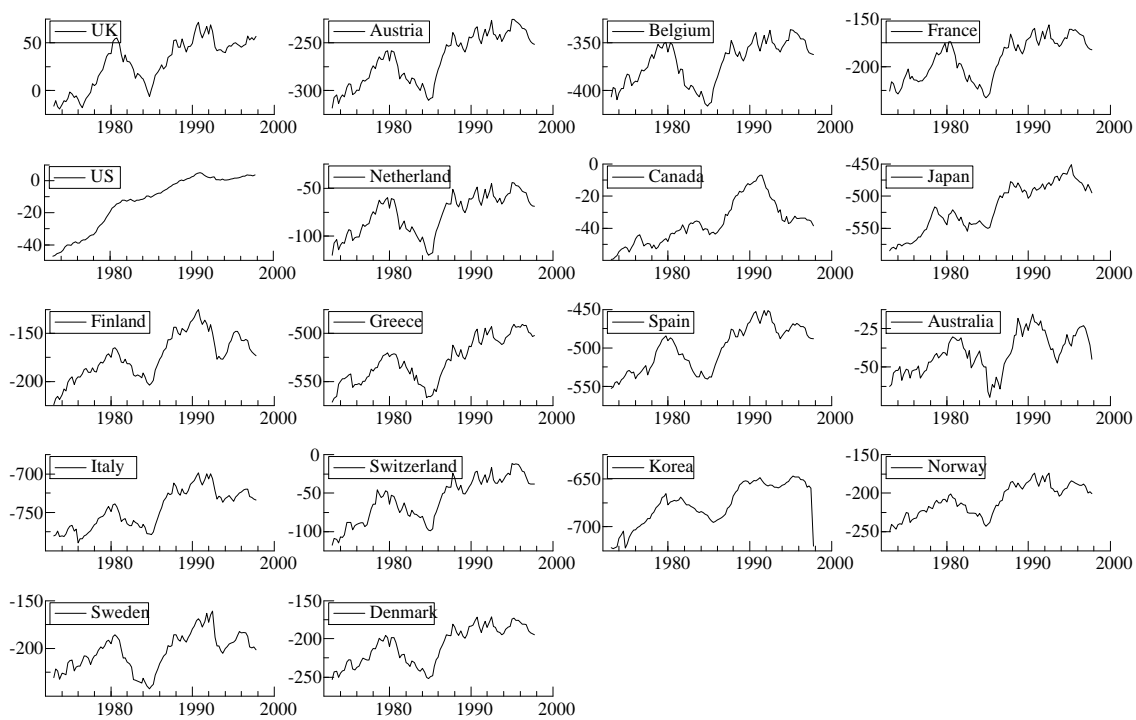
[a] For N=5, q=40% and 80%.

Figure 1: Real exchange rates, US is numeraire

Figure 2: Real exchange rates - Germany is numeraire