International Journal of Finance and Economics
*Int. J. Fin. Econ.* 5: 253–263 (2000)

# Is There a Base Currency Effect in Long-Run PPP?

Jerry Coakley* and Ana-Maria Fuertes
London Guildhall University, London, UK

The base currency effect in the purchasing power parity (PPP) literature refers to the stylized fact that tests on German mark real exchange rates are more likely to support mean reversion than analogous tests on US dollar rates. Using a panel of 19 OECD currencies, 1973–1997, we employ different panel unit root approaches to investigate the view that this effect can be attributed to neglected cross-sectional dependence. While the results from panel methods which permit cross-sectional dependence and heterogeneous serial correlation generally support long-run PPP, they provide no evidence of a base currency effect. Copyright © 2000 John Wiley & Sons, Ltd.

**KEY WORDS:** cross-sectional dependence; panel unit root tests; real exchange rates

## INTRODUCTION

The assumption of long-run purchasing power parity (PPP) has become a cornerstone of many open economy models.[1] PPP has recently been the subject of panel unit root tests which have the merit of increased power over standard time series tests for sample periods such as the post-Bretton Woods era.[2] One apparent stylized fact in the literature is that tests on German mark real exchange rates (RERs) support mean reversion more often than analogous tests on US dollar RERs. For instance, employing CPI-based RERs, Jorion and Sweney (1996), Papell (1997, 1998), Koedijk *et al.* (1998) and Papell and Theodoridis (1997, 1998) claim that evidence for PPP is stronger from the mark series than from the corresponding dollar series. Canzoneri *et al.* (1999) using the Balassa–Samuelson model and Wei and Parsley (1995) reach a similar conclusion for tradable prices.

Several attempts have been made in the literature to rationalize this effect, called the numeraire or base currency effect. On one hand it can be related to the Froot and Rogoff (1995) observation that it is easier to reject the no-cointegration null for fixed than for floating exchange rates. In this respect we note that most OECD groups are dominated by European economies which are characterized by the relative fixity of mark denominated exchange rates within the exchange rate mechanism of the EMS 1979–1992. This may be contrasted with the volatile behaviour of dollar RERs in the 1980s. On the other hand, the numeraire effect has been formally addressed by O'Connell (1998) from a statistical point of view. He argues that it is an artifact of empirical studies failing to deal adequately with the cross-sectional dependence induced by a common base currency and foreign price level. O'Connell shows analytically that panel unit root tests for PPP base on generalized least squares (GLS) estimation, which controls for cross-sectional dependence, should be invariant to the base currency when serial correlation is absent or is homogeneous across RERs.

* Correspondence to: Department of Economics, London Guildhall University, 31 Jewry St., London EC3N 2EY, UK. E-mail: jcoakley@econ.bbk.ac.uk

JEL Codes: C32, F31.

One problem in testing for a base currency effect is that it is difficult to disentangle it from the problems associated with tests of the PPP hypothesis itself. Thus testing for a unit root in RERs provides the overall framework for this paper. Within this, situations where identical panel unit root tests produce conflicting evidence on PPP for different numeraires are identified as implying a base currency effect. Attention is restricted to the mark and dollar numeraires for two reasons. On one hand, the choice of the dollar is vindicated by the fact that most international trade is invoiced in dollars while the predominance of European countries in the OECD panel under consideration renders the mark the other obvious candidate. On the other hand, use of these numeraires enables one directly to confront the most popular version of the base currency effect in existing studies.

This paper contributes to the PPP literature in two respects. First, in addressing the numeraire currency issue for 19 OECD countries, 1973–1997, it employs three panel unit root test procedures each of which deals with cross-sectional dependence and heterogeneous serial correlation in different manners.[3] The procedures comprise the Im, Pesaran and Shin (1995) (IPS) approach which averages the augmented Dickey–Fuller (ADF) unit root test statistic across groups, a test based on a seemingly unrelated regressions (SUR) system estimated by feasible GLS (FGLS), and the Taylor and Sarno (1998) likelihood ratio test which builds on the Johansen cointegrating VAR framework.[4] Second, once cross-sectional dependence is accounted for, our empirical findings provide qualitatively identical results for both numeraire currencies and thus no evidence of a base currency effect. Moreover, this finding is robust irrespective of whether the tests are applied both to the full panel of 18 OECD RERs or to various subsamples of this panel for both RER series. In this respect, they support the empirical results of Engel *et al.* (1997), Pedroni (1997) and Higgins and Zakrajšek (1999) but conflict with most other extant findings.

The paper is organized as follows. In the next section the three panel unit root procedures used are outlined. In the following section we describe the data and analyse our findings. A final section concludes.

## PANEL TESTS FOR A BASE CURRENCY EFFECT IN PPP

### Background

Despite the huge number of PPP studies, explicit analysis of the base currency effect generally has not figured prominently with a few exceptions. Firstly, O'Connell (1998) formalizes the base currency effect and attributes it to neglected cross-sectional dependence. More specifically, he showed analytically that, when the serial correlation pattern (both autoregressive parameters and truncation lag length) is homogeneous across groups, the GLS panel unit root test estimator is invariant to the numeraire choice. He illustrated empirically the impact of the latter by testing for PPP using a panel method which ignores cross-sectional dependence—by assuming a diagonal covariance matrix—and an FGLS estimation method which controls for it. In both cases the serial correlation structure is restricted to be homogeneous across countries.[5]

His main argument is that neglected contemporaneous interdependence increases the size and diminishes the power of panel tests. For his panel of 64 CPI RERs 1973:I–1995:IV, the former method indicates that the unit root null is rejected for the whole panel and three subsamples, Europe, Asia, South America but not for Africa. However, using FGLS estimation he finds absolutely no support for PPP. He concludes that, since simulation methods show that the FGLS tests have adequate size and power, the existing panel evidence favouring long-run PPP is an artifact of ignoring cross-sectional dependence. This he refers to as the 'overvaluation of PPP'.

Second, Papell and Theodoridis (1997) provide a wide-ranging empirical study of the base currency issue by using panels of 20 different CPI-based numeraire currencies over the 1973:I–1996:IV period. They establish that the degree of serial correlation varies considerably across countries which motivates allowing for heterogenous serial dependence in their analysis in contrast with that of O'Connell. By estimating via FGLS a system of ADF equations with heterogenous intercepts and serial correlation, they not only support the established wisdom on German mark based PPP but they generalize this to the proposition that tests on

 

European RERs are more likely to support PPP than analogous tests on non-European RERs.[6] They explain the numeraire effect by distance and exchange rate volatility. The latter is linked to the sharp appreciation and depreciation of the US$ in the early to mid-1980s whose effect is confirmed in a study of the same OECD panel by Papell and Theodoridis (1998) and, separately, by Lothian (1998).

Finally, Pedroni (1997) shows for the weak PPP formulation that a panel procedure which corrects for cross-sectional dependence with time dummies renders his tests asymptotically invariant to the numeraire choice even with heterogenous serial correlation. However, his approach differs from those of Papell and Theodoridis (1997) and O'Connell (1998) in that he compares the results of two separate procedures for controlling for cross-sectional dependence: a group demeaning or time dummies procedure on its own and a combination of a GLS correction and time dummies. Pedroni draws two main conclusions. First, accounting for cross-sectional dependence tends to strengthen the evidence in favour of weak PPP. Second, in no case does the GLS correction reverse the conclusions of his prior analysis which includes time dummies only.

From the above it is clear that the issue of cross-sectional dependence is critical for any investigation of the base currency effect. Higgins and Zakrajšek (1999) also underline its relevance for testing the stationarity of RERs, although their bottom line is diametrically opposed to that of O'Connell (1998). Using a series of panel unit root tests which take account of contemporaneous cross-sectional dependence, they find very strong evidence of RER stationarity for an OECD panel, a group of European economies and for a larger sample of open economies 1973:I–1997:IV.[7] Thus, for our empirical investigation, cross-sectional dependence played a central role in our choice of panel unit root test procedures which are described in the next section.

#### Panel Unit Root Tests of PPP

The real exchange rate $q_t$ is defined as the nominal exchange rate $s_t$ minus the difference between the domestic price index $p_t$ and the foreign price index $p_t^*$ as follows:

$$q_t = s_t - p_t + p_t^*, \tag{1}$$

where $s_t$ is defined as units of domestic per foreign currency and all variables are in logs. Assuming the variables $s_t$, $p_t$ and $p_t^*$ contain a unit root or are nonstationary, the long-run PPP hypothesis implies that $q_t$ is a mean reverting process or, equivalently that (1) represents a cointegrating relationship. Both univariate methods, such as the ADF unit root test, and multivariate methods, such as the Johansen reduced-rank procedure, have been extensively used to the test the stationarity of $q_t$. The panel test employed in this paper build on these methods.

One problem with the standard ADF test is its low power against near unit root alternatives for short data spans. Recent panel unit root tests have sought to overcome this problem by adding a cross-sectional dimension to the basic ADF equation. This initially involved building on the basic least squares dummy variable (LSDV) representation given in Levin and Lin (1992):

$$\Delta q_{it} = \alpha_{0i} + \beta q_{i,t-1} + \sum_{j=1}^{L} \alpha_j \Delta q_{i,t-j} + v_{it},$$
$$t = 1, \ldots, T, \quad i = 1, \ldots, N, \tag{2}$$

where $v_{it} \sim$ i.i.d. $(0, \sigma^2)$ for all $i$ and $E(v_t v_t') = \sigma^2 I_N$, with $v_t = (v_{1t}, v_{2t}, \ldots, v_{Nt})'$ and $I_N$ the identity matrix. Many of the early panel unit root PPP applications such as Frankel and Rose (1996), Oh (1996) and Wu (1996) employed variations of the above LSDV or fixed effects model which typically imposes a common speed of adjustment and homogeneous (if any) serial correlation patterns and no contemporaneous cross correlation between the residuals.

The more recently proposed tests have relaxed some or all of these assumptions. They vary in many respects such as the formulation of the null and alternative hypotheses, heterogeneity (or otherwise) in intercepts, $\alpha_{0i}$, speed of adjustment coeffcient, $\beta_i$, higher order dynamics including both AR parameters, $\alpha_{ji}$, and truncation lag, $L_i$, and the degree of cross-sectional dependence induced by the covariance matrix $E(v_t v_t')$.[8] Below we outline three tests which permit varying degress of both cross-sectional dependence and heterogeneous serial correlation.

*Im et al. (1995) t-Bar Tests*

The first approach considered is the IPS procedure which employs OLS estimation to compute standard unit root statistics for the individual groups and averages them over the panel. The IPS test accounts for heterogeneity across groups by allowing for different intercepts, speed of adjustment coefficients, higher order AR coefficients and truncation lags in the individual ADF regression as follows:

$$\Delta q_{it} = \alpha_{0i} + \beta_i q_{i,t-1} + \sum_{j=1}^{L_i} \alpha_{ji} \Delta q_{i,t-j} + v_{it},$$
$$t = 1, \ldots, T, \quad i = 1, \ldots, N, \tag{3}$$

where $v_{it} \sim$ i.i.d. $(0, \sigma_i^2)$ and $E(v_t v_t') = \bar{\sigma}' I_N$ with $\bar{\sigma} = (\sigma_1^2, \sigma_2^2, \ldots, \sigma_N^2)'$. That is, $v_t$ is a column vector of random errors which are neither autocorrelated nor contemporaneously correlated but can have different variances. The hypotheses being tested are:

$$\mathrm{H}_0: \quad \beta_i = 0, \forall i$$
$$\mathrm{H}_A: \quad \beta_i < 0, i = 1, \ldots, N_1;$$
$$\beta_i = 0, \; i = N_1 + 1, N_1 + 2, \ldots, N. \tag{4}$$

The test is consistent under the alternative that at least one of the series is stationary on condition that $0 < N_1/N \leq 1$ as $N(T) \to \infty$.

The IPS test is based on the following standardized mean-group or $t$-bar unit root statistic which averages the standard ADF $t$-ratio across group:[9]

$$\Psi_t = \frac{\sqrt{N(T)}\left\{\bar{t}_{NT}(L_i, \alpha_{ji}) - \frac{1}{N}\sum_{i=1}^{N} E[t_{iT}(L_i, 0)|\beta_i = 0]\right\}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N} \mathrm{Var}[t_{iT}(L_i, 0)|\beta_i = 0]}}, \tag{5}$$

where

$$\bar{t}_{NT}(L_i, \alpha_{ji}) = \frac{1}{N}\sum_{i=1}^{N} t_{iT}(L_i, \alpha_{ji})$$

and $t_{iT}(L_i, \alpha_{ji})$ is the individual $t$-statistic for testing $\beta_i = 0$ in (3). The values of $E[t_{iT}(L_i, 0)|\beta_i = 0]$ and $\mathrm{Var}[t_{iT}(L_i, 0)|\beta_i = 0]$, which represent the asymptotic mean and variance, respectively, of the individual $t$-statistics under the null have been evaluated via stochastic simulations for different $L_i$ and $T$ and tabulated in Im *et al.* (1995) The $\Psi_t$ statistic converges in probability to a standard normal variate under the null as $T \to \infty$, $N(T) \to \infty$ and $N = O(T)$,[10] while $\Psi_t \to -\infty$ under the alternative.

The standard IPS (hereafter IPS) panel test based on (3) can be used as a baseline in investigating the numeraire effect in the absence of cross-sectional dependence. Another version (hereafter demeaned IPS) extends the above test to allow for cross-sectional dependence or a nondiagonal covariance matrix, $E(v_t v_t') = \Omega$. To do this Im *et al.* (1995) assume that the error term in (3) comprises two orthogonal random components, $v_{it} = \theta_t + e_{it}$, where $\theta_t$ is a stationary time-specific common effect across groups and $e_{it}$ is a group-specific or idiosyncratic effect which is independent across groups. Accordingly, they propose a demeaning procedure which substracts cross-sectional means $\bar{q}_t$ from the observed series $q_{it}$ and they construct a demeaned version $\tilde{\Psi}_t$ of the estimator (5) on the basis of the individual $t$-statistics for the demeaned series $\tilde{q}_{it} = q_{it} - \bar{q}_t$. This demeaning procedure is equivalent to including time dummies in (3) to capture the common interdependence between countries. Heterogeneous serial correlation in $e_{it}$ can be dealt with by adding higher order dynamics as for $v_{it}$ in regression (3).

The unit root $t$-statistics obtained by estimating the demeaned equations separately for each group are asymptotically valid. Put differently, for the assumed $v_{it}$, the demeaning procedure is asymptotically equivalent—as $N(T) \to \infty$—to estimating the associated $\Omega$ directly. However the demeaned IPS approach cannot capture any dynamic interdependencies resulting in feedbacks over time.[11] Pedroni (1997) provides an interesting and original interpretation of the demeaning procedure: it amounts to a small open economy assumption since such economies are affected by shocks to the rest of the world but not vice versa.

*SUR-FGLS Test*

The second panel approach considered is based on a SUR system as in Abuaf and Jorion (1990) and is estimated by iterative FGLS.[12] Accordingly, the ADF($L_i$) equation for each group $i$ in the panel is:

$$\Delta q_{it} = \alpha_{0i} + \beta q_{i,t-1} + \sum_{j=1}^{L_i} \alpha_{ji} \Delta q_{i,t-j} + v_{it},$$
$$t = 1, \ldots, T, \quad i = 1, \ldots, N, \tag{6}$$

where $v_{it} \sim$ i.i.d. $(0, \sigma_i^2)$ and $E(v_t v_t') = \Omega$, where $v_t = (v_{1t}, v_{2t}, \ldots, v_{Nt})'$ and $\Omega$ is an $N \times N$ (possibly) nondiagonal covariance matrix with diagonal elements $\Omega_{ii} = \sigma_i^2$, $i = 1, \ldots, N$.[13] The proposed model allows for group-specific intercepts ($\alpha_{0i}$) and different higher-order dynamics ($\alpha_{ji}$, $L_i$) but imposes a common speed of adjustment coefficient ($\beta$) across groups.[14] The test statistic is the $t$-statistic on $\beta$ and the hypotheses tested:

$$\mathrm{H}_0: \quad \beta = 0$$
$$\mathrm{H}_A: \quad \beta < 0$$

Collecting the individual ADF regressions (6) in a system yields a SUR model with a cross-equation restriction (common $\beta$) which can be efficiently and consistently estimated by iterated FGLS.

This is the estimation approach adopted by Papell (1997), Papell and Theodoridis (1997) and O'Connell (1998) to control for cross-sectional dependence in panel unit root tests. Note that while O'Connell assumes common serial correlation properties for all RERs ($\alpha_{ji} = \alpha_{jk}$ and $L_i = L_k$ for all $i$, $k$) the others allow for heterogeneous serial correlation. We follow the latter less restrictive approach because it makes the SUR-FGLS and IPS test results more comparable. Finally, since the distribution of the unit root $t$-statistic on $\beta$ under the null is not standard, for this test we employ the empirical critical values generated by Monte Carlo simulations.

### *Taylor and Sarno (1998) Johansen LR Test*

Finally, Taylor and Sarno (1998) propose an innovative multivariate unit root test which builds on the Johansen (1988) maximum likelihood (ML) cointegration approach. In the latter the hypothesis of cointegration is formulated as a reduced-rank condition in the following cointegrating VAR representation:

$$\Delta y_t = \Gamma + \Pi y_{t-1} + \sum_{j=1}^{L} \Phi_j \Delta_{y_{t-j}} + v_t, \quad t = 1, \ldots, T, \tag{7}$$

where $y_t = (y_{1t}, y_{2t}, \ldots, y_{Nt})'$, $\Gamma$ is an $N \times 1$ vector of constants, $\Pi$ is an $(N \times N)$ long-run multiplier matrix and $v_t$ is an $N \times 1$ vector of disturbances such that $v_t \sim$ i.i.d. $N(0, \Omega)$.[15] The rank of $\Pi$ or, equivalently, the number of its nonzero latent roots is equal to the number of cointegrating vectors $r$. If $\Pi$ has full rank, $r = N$, this implies that all the series $y_t$ are realizations of stationary processes.

The Taylor and Sarno approach fits the above VAR to a panel of $N$ time series and tests the null that at least one series has a unit root against the alternative that they all are stationary. This amounts to testing the null that $\Pi$ has less than full rank or that its smallest characteristic root ($\lambda_{\min}$) is zero. On this basis, Taylor and Sarno propose the following Johansen likelihood ratio (JLR) statistic:

$$\mathrm{JLR} = -T \ln(1 - \lambda_{\min}), \tag{8}$$

which has a $\chi^2(1)$ limiting distribution under the null. Taylor and Sarno (1998) show via simulation that the empirical distribution of the JLR statistic is quite close to the asymptotic distribution for $T > 100$. While the JLR test imposes a common truncation lag (given by the order of the VAR) on all RERs, the AR parameters (coefficients on $\Delta y_{t-j}$) can vary across groups much as in the IPS and FGLS tests.

The small sample properties of the IPS, SUR-FGLS and JLR panel unit root methods are different, as expected, given that they all formally test distinct hypotheses and that they account for cross-sectional dependence and the serial correlation pattern in various ways. In this respect, the joint application of all three tests to our panel of RERs seeks to provide a robust analysis of both the base currency effect and of long-run PPP.

## EMPIRICAL ANALYSIS

### Data and Test Results

Our monthly data were taken from *Datastream* and span the 24-year period, January 1973 to January 1997, giving a total of 289 observations per variable. The RER is defined using the wholesale price index (hereafter WPI) with both the German mark (DM) and the US dollar (US$) as numeraires. The panel comprises 19 OECD countries and contains a high proportion of European countries.[16] The ADF and semiparametric Phillips–Perron unit root tests fail to reject the unit root null at the 5% level in virtually all nominal exchange rates (excluding Austria and Switzerland for the DM series) and price levels (excluding Denmark, Ireland and the UK). The combined ADF and Kwiatkowski *et al.* (1992) unit root test indicates $I(0)$ behaviour in just two RERs for the DM but for no US$ RER series. The Johansen procedure provides support for weak PPP in all series, but rejects the $[1, -1, 1]$ cointegrating restriction or, equivalently, fails to reject $I(1)$ behaviour in all but three DM and US$ RER series. Overall, these results are interpreted as providing very little evidence in favour of stationary RERs which accords with the existing time series literature on PPP.[17]

Prior to applying the panel approaches discussed in the section 'Panel Unit Root Tests of PPP' to the full OECD panel and subpanels, we stacked the individual ADF($L_i$)regressions in a SUR system and tested for the significance of the off-diagonal elements of the error covariance matrix $E(v_t v_t') = \Omega$. We employ the log-likelihood ratio statistic $\mathrm{LR}_\Omega = 2(\mathrm{LL}_U - \mathrm{LL}_R)$ where $\mathrm{LL}_U$ is the maximized value of the log-likelihood function under $\mathrm{H}_A$ (the unrestricted model) and $\mathrm{LL}_R$ is the maximized value of the log-likelihood function under $\mathrm{H}_0$ (the restricted model). The large sample distribution of $\mathrm{LR}_\Omega$ is a $\chi^2_{N(N-1)/2}$. The test results indicate decisive rejection of the diagonality null for the different panels and both numeraires. For instance, $\mathrm{LR}_\Omega$ yields values of 208.9 and 956.8 for the DM-based G5 and G10 panels, and 667.9 and 1889.4 for the US$-based G5 and G10 counterparts. These results vindicate our choice of panel tests which correct for cross-sectional dependence.

The results for the standard and demeaned IPS test and the SUR-FGLS test are presented in Table 1.[18] For the baseline IPS procedure which ignores cross-sectional dependence, the overall results are more supportive of long-run PPP for the DM than for the US$ series both in terms of number of rejections of the unit root null and in terms of the absolute value of the statistic. In this sense they seem to provide evidence of a base currency effect. All the subsequent panel results take account of cross-sectional dependence.

Table 1. IPS and SUR-FGLS unit root tests

| Panel | $\Psi_{\bar{t}}$ Statistic, Equation (3) | |
|---|---|---|
| | DM series | US$ series |
| A. IPS results | | |
| OECD | −4.14* | −2.61* |
| G10 | −2.71* | −1.49** |
| G7 | −2.13* | −1.24 |
| G5 | −1.29** | −0.90 |
| B. Demeaned IPS results | | |
| OECD | −4.86* | −2.40* |
| G10 | −2.77* | −2.84* |
| G7 | −1.65* | −1.77* |
| G5 | −0.67 | −0.71 |
| C. SUR-FGLS results | $t(\beta)$ Statistic, Equation (6) | |
| OECD | −8.37* | −8.38* |
| G10 | −6.13* | −5.65* |
| G7 | −5.19* | −4.82* |
| G5 | −3.93 | −3.92 |

* Indicates rejection of the unit root null at the 5% level.
** Indicates rejection of the unit root null at the 10% level.

The results from the demeaned IPS and SUR-FGLS panel tests are supportive of long-run PPP for the full OECD panel and the G10 and G7 subsamples at the 5% level for both numeraires. The lack of rejection for the smaller G5 panel ($N = 4$) from both tests may well reflect some power loss as the cross-sectional dimension $N$ diminishes for fixed $T$.[19] What is interesting is that both methods produce qualitatively identical results across the two numeraires for all panels. These can be interpreted as vindicating O'Connell's (1998) argument on the importance of allowing for cross-sectional dependence. The fact that the demeaned IPS procedure and the SUR-FGLS test produced identical results for both numeraires and the same verdict on PPP supports Pedroni's (1997) empirical findings and suggests that a cross-section demeaning or time dummies procedure may be adequate for capturing most of the contemporaneous dependence in RER panels.

The results from the Taylor and Sarno (1998) multivariate JLR test procedure are reported in Table 2.[20] The JLR statistic clearly supports RER mean reversion for the G7 and G5 panels using both the DM and US$ numeraire. These findings supporting PPP are in line with those of Sarno and

 

Table 2. Multivariate JLR unit root test

| Panel | DM series | | US$ series | |
|---|---|---|---|---|
| | $\lambda_{\min}$ ($\times 10^3$) | Test statistic | $\lambda_{\min}$ ($\times 10^3$) | Test statistic |
| OECD | 1.0074 | 2.92 | 1.0105 | 2.94 |
| G10 | 1.2047 | 3.50 | 1.2273 | 3.57 |
| G7 | 1.3475 | 3.92* | 1.3141 | 3.83* |
| G5 | 1.6965 | 4.94* | 1.6741 | 4.87* |

* Indicates rejection of the null that at least one of the series has a unit root at the 5% level using the asymptotic $\chi^2(1)$ critical values.
A lag order $L = 1$ in (7) was selected by the Akaike and Schwarz criteria in all cases.

Taylor (1998) for the G5 using US$-CPI monthly RERs for the 1973:1–1996:12 period and also with the Taylor and Sarno (1998) results for a similar G5 panel with quarterly frequency. However, the JLR statistic fails to reject the null that at least one of the series has a unit root for the two larger panels.[21] Again, no evidence of a base currency effect is found in line with the results of the previous two panel tests.

Overall the results from the three panel procedures which permit cross-sectional dependence and hetererogeneous serial correlation are invariably consistent for the two numeraire currencies employed. They contrast with the results from the standard IPS procedure, where cross-sectional dependence is ignored, which seem to provide some evidence of a base currency effect. They can be interpreted as supporting O'Connell's argument that the numeraire effect is a statistical artifact of neglected cross-sectional dependence and are also in accord with the findings of Engel et al. (1997) and Pedroni's (1997) results for weak PPP. Finally, for our particular sample at least, allowing for heterogeneous serial correlation does not appear to lead to a base currency effect.[22] We conjecture that other factors such as the volatility of dollar exchange rates in the 1980s combined with the relative stability of German mark rates over this period may have less of an impact in our longer sample period as compared with earlier studies employing shorter spans.

### Implications for PPP

An implication of our empirical investigation is that the verdict on long-run PPP from the three panel tests employed is generally favourable for the different (sub)panels. Three separate tests reject the nonstationarity null for the G7, two tests reject the null for both the full panel and the G10 while the JLR test rejects the null for the G5. Note however that our verdict on PPP contrasts sharply with O'Connell's (1998) blanket rejection. O'Connell's testing approach and results have recently been called into question by Higgins and Zakrajšek (1999).They query his truncation lag selection procedure which leads to a choice of relatively high order AR models. They use Monte Carlo simulations to show that this leads to a substantial loss of power which can explain O'Connell's blanket rejection of long-run PPP. On the other hand, in keeping with the literature our panel test results contrast with non-panel time series test results. Though the failure of the latter tests to reject the unit root null has usually been explained on the basis of low power for near unit root alternatives and short data spans, a number of recent studies rationalize it as an artifact of employing linear time series specifications.

Nonlinear time series models have been applied to RERs to capture the effects induced by transaction costs and other market frictions that may impede profitable arbitrage opportunities for small deviations from equilibrium or (PPP) parity.[23] Thus, RER behaviour might be better characterized by models such as the threshold autoregression (TAR) or smooth transition autoregression (STAR) specifications that allow for the speed of adjustment to vary with the size of the deviation from equilibrium.[24] In the light of the generally more favourable verdict on long-run PPP from panel procedures which also rely on linear specifications, it is interesting to examine the effect of these nonlinearities on the small sample properties of panel tests.

In an important recent contribution Taylor *et al.* (1999) address this question via Monte Carlo experiments.[25] In particular they replicate panels from a nonlinear DGP calibrated on the estimated parameters of a STAR model for four bilateral RERs and the cross-sectional dependence induced by the estimated covariance matrix of residuals. One of their main findings is that for sample sizes typical of many PPP studies ($T = 200$) the ADF test suffers from a dramatic power loss in the presence of nonlinearities, with only a 30–40% probability of rejecting the false unit root null. By contrast, for the JLR and multivariate ADF (MADF) procedures, the rejection frequencies are close to 100%.[26] The overall conclusion is that panel unit root tests appear to be far more robust than univariate tests to the presence of nonlinearities in the (stationary) DGP, which helps to reconcile the nonlinear time-series and (linear) panel empirical literatures on RER mean reversion. In this sense, our findings may also be viewed as consistent with the presence of nonlinearities in RERs.

## CONCLUSIONS

This paper investigates empirically the apparent stylized fact of a base currency effect in long-run PPP. To this end it applies several panel unit root tests, which deal differently with both cross-sectional dependence and heterogeneous serial correlation, to a sample of US dollar and German mark RERs for 19 OECD countries, 1973–1997. These include the IPS mean group ADF estimator in its standard and demeaned versions, a SUR-FGLS panel test procedure, and the Taylor and Sarno (1998) likelihood ratio approach. Our results seem to yield some evidence of a base currency effect if cross-sectional dependence is ignored. However, they are invariant across both numeraires and panels when the contemporaneous cross correlation in the disturbances is accounted for. This provides clear empirical evidence against the base currency effect. While it is at odds with conventional wisdom in the empirical literature, it supports the results of Engel *et al.* (1997), Pedroni (1997) and O'Connell (1998).

Our panel results generally support the stationarity of RERs for the post-1973 era in contrast with standard time series test results. It is plausible that, as a number of recent studies suggest, nonlinearities in the generating process of RERs—possibly fostered by transaction costs—may explain the lack of support from the latter studies. The effect of these nonlinearities on the small sample properties of panel unit tests, such as those employed in the present study, has been investigated by Taylor *et al.* (1999) via Monte Carlo analysis. Their overall conclusion that panel unit root tests retain most of their power in the presence of nonlinearities provides one means of reconciling the nonlinear and panel literatures on PPP.

## ACKNOWLEDGEMENTS

The authors are grateful to Ron Smith and an anonymous referee for detailed and constructive comments which greatly improved the exposition of this paper. The authors also thank participants at the 2000 MMF Annual Conference at South Bank University and at seminars at Birkbeck College, London Guildhall and SOAS, University of London, for helpful comments.

## NOTES

1. See Taylor (1995) for a recent review of PPP. The PPP assumption is also relevant to debates such as the Feldstein–Horioka puzzle. See Coakley *et al.* (1998).
2. A selective sample of panel studies on long-run PPP includes Pedroni (1995, 1997), Frankel and Rose (1996), Wu (1996), Coakley and Fuertes (1997), Sarno and Taylor (1998), Taylor and Sarno (1998), Boucher-Breuer *et al.* (1999) and Boyd and Smith (1999).
3. We restrict out attention to these three methods for space limitations. However, a number of alternative panel unit root approaches are available. These include the multivariate ADF (MADF) test based on SUR estimators proposed by Taylor and Sarno (1998), the long memory panel test of Andersson and Lyhagen (1999) and the Boucher-Breuer *et al.* (1999) SUR-ADF approach which allows one to identify which groups contain a unit root.
4. The IPS test has been extended to a cointegration framework by Pedroni (1995, 1997, 1998), Larsson *et al.* (1998) and Larsson and Lyhagen (1999).
5. An interesting empirical issue is whether imposing homogeneity on the serial correlation structure may be overly restrictive for many of the PPP panels employed in the literature.

6. While their study is extremely thorough, it is difficult to give an economic interpretation to some of their detailed results such as the fact that New Zealand based RERs produced the highest number of ADF test rejections and were the only non-European numeraire series to indicate evidence of stationarity in a panel test.
7. The panel unit root tests they employ include O'Connell's GLS test, a SUR-GLS test similar in spirit to that of Taylor and Sarno (1998) and GLS extensions of the Levin and Lin (1993) and Im *et al.* (1995) tests.
8. Abuaf and Jorion (1990) were the first to address the issue of cross-sectional dependence in PPP tests. See Smith (2000) or an interesting discussion of panel unit root tests.
9. Im *et al.* (1995) also propose a Lagrange multiplier type statistic, the LM-bar. Our choice is based on their Monte Carlo simulation results which show that in small samples the $t$-bar test performs marginally better than the LM-bar counterpart.
10. This aymptotic condition implies that $N(T)/T \to k$ for some finite positive constant $k$, and thus the IPS test is more generally applicable than those panel tests which require the stronger condition that $N(T)/T \to 0$.
11. This implies that in certain panels the covariance matrix $\Omega$ induced by the assumption $v_{it} = \theta_t + e_{it}$ may not be able to capture all the cross-sectional dependence present in the data.
12. Levin and Lin (1993), Frankel and Rose (1996), Oh (1996) and Lothian (1998) all use panel unit root tests based on GLS estimators.
13. The disturbance covariance matrix for the full $NT$ observations is $E(vv') = V = \Omega \otimes I_T$ where $\otimes$ denotes the Kronecker product. Note that (6) is a generalization—with cross-sectional dependence of disturbances and heterogeneous serial correlation—of the LSDV model (2).
14. Engel *et al.* (1997) argue that there are grounds for assuming that intra-continental slopes might differ from inter-continental slopes but their empirical results did not establish any significant difference in this respect.
15. If $\Pi$ has rank $r$, with $0 < r < N$, then $\Pi$ can be factored as $\Pi = \alpha\beta'$ where $\alpha$ and $\beta$ are $N \times r$ matrices. The cointegrating vectors are proportional to the columns of $\beta$. The approach of Johansen is to estimate (7) subject to the cross-equation constraint $\Pi = \alpha\beta'$ for various values of $r$ using ML.
16. Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Netherlands, Norway, Sweden, Spain, Switzerland, UK and USA. Our panel differs from that of Papell and Theodoridis (1997) which is CPI-based, has a quarterly frequency and additionally contains New Zealand and Portugal for which WPI series were unavailable.
17. For details see Coakley and Fuertes (2000b).
18. For the IPS and FLGS tests the number of lags $L_i$ for each ADF Equation (3) was chosen by the Schwarz and Akaike criteria and corroborated by the Breusch–Godfrey serial correlation LM test. For instance, for the demeaned RERs in the G10 panel this selection procedure gives $L_i = 4$ for Belgium, $L_i = 2$ for Italy, Netherlands, Sweden and the UK, $L_i = 1$ for France and $L_i = 0$ for Japan, Canada and the USA.
19. In the SUR-FGLS test a small amount of power may leak away due to the fact that $\Omega$ is not known and has to be estimated. For this test we rely on Papell's (1997) critical values since our $N$ and $T$ dimensions are similar.
20. The lag order for the cointegrating VAR comprising these RERs was selected using the Akaike and Schwarz criteria.
21. The larger panels' results may be related to the potential high-order-VAR bias problem alluded to by Abadir *et al.* (1999). However, this should not affect tests of the numeraire effect since any high order bias would have similar effects for both the DM-and US$-based VAR.
22. Out verdict on the numeraire effect contrasts with that of Papell and Theodoridis (1997) despite using a similar SUR-FGLS panel test. The difference may stem from their more aggregated data, their markedly higher AR lag orders and the fact that their panel includes two additional currencies (New Zealand dollar and the Portuguese escudo) which display puzzling behaviour.
23. See for example, Michael *et al.* (1997), Obstfeld and Taylor (1997), Bleaney and Leybourne (1998), Baum *et al.* (1999), Taylor *et al.* (1999) and Coakley and Fuertes (2000a,d).
24. The latter has been termed 'amplitude asymmetry' in Coakley and Fuertes (2000c) where bootstrap LR tests are proposed to test for this and other types of asymmetry.
25. This addresses the two major PPP puzzles of nonstationary RERs and slow mean reversion. In so doing they find clear evidence of nonlinear (exponential STAR) mean reversion in four major bilateral dollar RERs 1973:1–1996:12. We are grateful to Lucio Sarno for kindly providing a copy of this unpublished paper.
26. Note that both of these procedures allow for cross-sectional dependence. An extension of the interesting issues raised in Taylor *et al.* (1999) is to explore the small sample power of the mean group estimator $\Psi_r$ for a nonlinear DGP, and the sensitivity of this to the cross-sectional dimension ($N$), the degree of cross-sectional dependence ($\Omega$) and the range of nonlinearities. This issue is being pursued by the authors in the context of a Band-TAR DGP using the numerical algorithm for the estimation of this class of models developed in Coakley *et al.* (2000).

## REFERENCES

Abadir KM, Hadri K, Tzavalis E. 1999. The influence of VAR dimensions on estimator biases. *Econometrica* **67**: 163–181.

Abuaf N, Jorion P. 1990. Purchasing power parity in the long run. *Journal of Finance* **45**: 157–174.

Andersson J, Lyhagen J. 1999. A Long Memory Panel Unit Root Test: PPP Revisited. Stockholm School of Economics Working Paper in Economics and Finance, No. 303.

Baum CF, Caglayan M, Barkoulas JT. 1999. Nonlinear Adjustment to Purchasing Power Parity in the Post-Bretton Woods Era. Boston College Discussion Paper No. 3/98, Revised November, 1999.

Bleaney MF, Leybourne SJ. 1998. Real Exchange Rate Dynamics under the Current Float. University of Nottingham Discussion Paper No. 98/9.

Boucher-Breuer J, McNown R, Wallace M. 1999. Series-specific tests for a Unit Root in a Panel Setting with an Application to Real Exchange Rates. Mimeo, University of Colorado.

Boyd D, Smith R. 1999. Testing for purchasing power parity: econometric issues and an application to developing countries. *Manchester School of Economic and Social Studies* **67**: 287–303.

Canzoneri M, Cumby R, Diba B. 1999. Relative labor productivity and the real exchange rate in the long run: evidence from a panel of OECD countries. *Journal of International Economics* **47**: 245–266.

Coakley J, Fuertes AM. 1997. New panel unit root tests of PPP. *Economics Letters* **57**: 17–22.

Coakley J, Fuertes AM. 2000a. Short run real exchange rate dynamics. *The Manchester School of Economic and Social Studies* **68**: 461–475.

Coakley J, Fuertes AM. (2000b) Nonparametric cointegration analysis of real exchange rate dynamics. *Applied Financial Economics* (forthcoming).

Coakley J, Fuertes AM. 2000c. Bootstrap LR Tests of Sign and Amplitude Asymmetric Adjustment. Birkbeck College Discussion Paper (in *Economics* 4/2000).

Coakley J, Fuertes AM. 2000d. Border costs and real exchange rate dynamics in Europe. *Journal of Policy Modelling* (forthcoming).

Coakley J, Fuertes AM, Pérez T. 2000. An efficient rational interpolation approach to least squares estimation of Band-TARs. *Lecture Notes in Computer Science* (forthcoming).

Coakley J, Kulasi F, Smith R. 1998. The Feldstein-Horioka puzzle and capital mobility: a review. *International Journal of Finance and Economics* **3**: 169–188.

Engel C, Hendrickson MK, Rogers JH. 1997. Intra-National, Intra-Continental, and Intra-Planetary PPP. NBER Working Paper No. 6069.

Frankel J, Rose A. 1996. A panel project on purchasing power parity: mean reversion within and between countries. *Journal of International Economics* **40**: 209–224.

Froot KA, Rogoff K. 1995. Perspectives on PPP and long run real exchange rates. In *Handbook of International Economics*, vol. 3, Grossman GM, Rogoff K (eds). North Holland: Amsterdam; 1647–1688.

Higgins M, Zakrajšek E. 1999. Purchasing Power Parity: Three Stakes in the Heart of the Unit Root Null. Mimeo, Federal Reserve Bank of New York.

Im K-S, Pesaran MH, Shin Y. 1995. Testing for Unit Roots in Heterogeneous Panels. DAE Working Paper 9526, Revised December 1997, University of Cambridge.

Johansen S. 1988. Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control* **12**: 231–254.

Jorion P, Sweney RJ. 1996. Mean reversion in real exchange rates: evidence and implications for forecasting. *Journal of International Money and Finance* **15**: 535–550.

Koedijk KG, Schotman PC, Vasn Dijk MA. 1998. The reemergence of PPP in the 1990s. *Journal of International Money and Finance* **17**: 51–60.

Kwiatkowski D, Phillips P, Schmidt P, Shin Y. 1992. Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics* **54**: 159–178.

Larsson R, Lyhagen J, Löthgren M. 1998. Likelihood-Based Cointegration Tests in Heterogeneous Panels. Stockholm School of Economics Working Paper in Economics and Finance, No. 250.

Larsson R, Lyhagen J. 1999. Likelihood-Based Inference in Multivariate Panel Cointegration Models. Stockholm School of Economics Working Paper in Economics and Finance, No. 331.

Levin A, Lin C-F. 1992. Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties. Discussion Paper No. 92-23, University of California, San Diego.

Levin A, Lin C-F. 1993. Unit Root Tests in Panel Data: New Results. Discussion Paper No. 93-56, University of California, San Diego.

Lothian JR. 1998. Some new stylized facts of floating exchange rates. *Journal of International Money and Finance* **17**: 29–39.

Michael P, Nobay AR, Peel DA. 1997. Transaction costs and nonlinear adjustment in real exchange rates: an empirical investigation. *Journal of Political Economy* **105**: 862–879.

Obstfeld M, Taylor AM. 1997. Non-linear aspects of goods market arbitrage and adjustment: Heckscher's commodity points revisited. *Journal of the Japanese and International Economies* **11**: 441–479.

O'Connell PJG. 1998. The overvaluation of purchasing power parity. *Journal of International Economics* **44**: 1–19.

Oh KY. 1996. Purchasing power parity and unit root test using panel data. *Journal of International Money and Finance* **15**: 405–418.

Papell DH. 1997. Searching for stationarity: purchasing power parity under the current float. *Journal of International Economics* **43**: 313–332.

Papell DH. 1998. The Great Appreciation, the Great Depreciation, and the Purchasing Power Parity Hypothesis. Working Paper 30, Oesterreichische Nationalbank, Vienna, Austria.

Papell DH, Theodoridis H. 1997. The Choice of Numeraire Currency in Panel Tests of Purchasing Power Parity. Mimeo, University of Houston.

Papell DH, Theodoridis H. 1998. Increasing evidence of purchasing power parity over the current float. *Journal of International Money and Finance* **17**: 41–50.

Pedroni P. 1995. Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis. Working Paper No. 95-013, Indiana University. Revised April 1997.

Pedroni P. 1997. Cross Sectional Dependence in Cointegration Tests of Purchasing Power Parity in Panels. Mimeo, Indiana University.

Pedroni P. 1998. Approximate Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors. Mimeo, Indiana University.

Sarno L, Taylor MP. 1998. Real exchange rates under the recent float: unequivocal evidence of mean reversion. *Economics Letters* **60**: 131–137.

Smith RP. 2000. Estimation and Inference with Nonstationary Panel Time Series Data. Mimeo, Birkbeck College.

Taylor MP. 1995. The economics of exchange rates. *Journal of Economic Literature* **33**: 13–47.

Taylor MP, Sarno L. 1998. The behaviour of real exchange rates during the post-Bretton woods period. *Journal of International Economics* **46**: 281–312.

Taylor MP, Peel D, Sarno L. (1999) Nonlinear mean reversion in real exchange rates: toward a solution of the purchasing power parity puzzle. *International Economic Review* (forthcoming).

Wei S-J, Parsley D. 1995. Purchasing Power Dis-Parity During the Floating Rate Period: Exchange Rate Volatility, Trade Barriers and Other Culprits. NBER Working Paper No. 5032.

Wu Y. 1996. Are lone run real exchange rates nonstationary? Evidence from a panel data test. *Journal of Money Credit and Banking* **28**: 54–61.

 