Department of Economics

Getting PPP Right: Identifying Mean-Reverting Real Exchange Rates in Panels

Georgios Chortareas and George Kapetanios

Working Paper No. 629 July 2008 ISSN 1473-0278

Queen Mary
University of London

Getting PPP Right: Identifying Mean-Reverting Real Exchange Rates in Panels*

Georgios Chortareas†, George Kapetanios‡

June 2008

Abstract

Recent advances in testing for the validity of Purchasing Power Parity (PPP) focus on the time series properties of real exchange rates in panel frameworks. One weakness of such tests, however, is that they fail to inform the researcher as to which cross-section units are stationary. As a consequence, a reservation for PPP analyses based on such tests is that a small number of real exchange rates in a given panel may drive the results. In this paper we examine the PPP hypothesis focusing on the stationarity of the real exchange rates in up to 25 OECD countries. We introduce a methodology that when applied to a set of established panel-unit-root tests, allows the identification of the real exchange rates that are stationary. We apply procedures that account for cross-sectional dependence. Our results reveal evidence of mean-reversion that is significantly stronger as compared to that obtained by the existing literature, strengthening the case for PPP. Moreover, our approach can be used to provide half-lives estimates for the mean-reverting real exchange rates. We find that the half-lives are shorter than the literature consensus and therefore that the PPP puzzle is less pronounced than initially thought.

Key Words: PPP, panel unit root tests, real exchange rates, half-lives, PPP puzzle.

JEL Classification: C12, C15, C23, F31

Previous JBF Manuscript numbers: JBF no. 9089, 9684

*We are grateful to two anonymous referees for their very insightful comments. This paper was originally submitted to Professor Giorgio Szego in December 2006 and was revised three times prior to submission through EES.

†Department of Economics, University of Athens, 8 Pesmazoglou Street, Athens, 10559, Tel. +(30)210 3689805, Fax+(30)210 3689810, email: GChortar@econ.uoa.gr

‡Department of Economics, Queen Mary University of London, Mile End Road, London E1 4NS, UK. email: G.Kapetanios@qmul.ac.uk. Tel: +44(0)20 78825097.

1. Introduction

Given the central role of the Purchasing Power Parity (PPP) concept in theoretical open economy models and the inconclusive results of the existing empirical literature on its validity, PPP has emerged as the most popular topic of empirical research in international macroeconomics. Testing for unit roots in real exchange rates using panels is popular partly because the results of such studies tend to uncover more evidence for PPP. Other advantages of panel unit root tests include the ability to mitigate problems such as the "survivorship bias" and the presence of structural shifts in exchange rate behavior.

Panel frameworks are not free of drawbacks, however, and most recent developments emphasized those relating to cross-sectional dependence. Nevertheless, from an economist's point of view, a major weakness of the existing unit root panel methodologies is that the null of non-stationarity is a joint hypothesis for all the real exchange rates in the panel. As a consequence the null hypothesis of a unit root may be rejected even if only one of the real exchange rates is stationary.[1] Thus, the possibility emerges that small groups of cross-sectional units in the panel, that share particular features, may drive the results. Therefore, panel unit root tests are sensitive to the selection of series included in the panel.

In this paper we consider the PPP hypothesis in panels of up to 25 OECD countries using an approach that overcomes the limitations mentioned above. In particular, we introduce a methodology that when applied to a battery of panel-unit-root tests, allows the identification of the real exchange rates that are stationary within the panel. We apply those procedures to a set of tests that accounts for a number of other potential pitfalls in panels, such as cross-sectional dependence. Our results reveal evidence of mean-reversion that is significantly stronger compared to that obtained by standard stationarity tests, strengthening the case for PPP. Our methodology has some straightforward advantages as compared to the typical panel unit root approaches. In particular, while we exploit all the advantages of the panel structure (such as the potential enhanced power of panel unit root tests), we are able to identify the stationary real exchange rates within the panel. This allows a direct comparison of the panel test results with the univariate tests results, i.e., focusing on individual real exchange rates - something that the existing literature on real exchange rates and PPP was not able to do so far.

Our ability to identify mean reverting series within the panel, allows us to focus on the half-lives of the mean-reverting real exchange rates only. We find that the half-lives are shorter than the prevailing literature consensus. Thus, we revisit the so-called "PPP puzzle" in the light of our new results providing half-life estimates that pertain only to the stationary real exchange rates of the panel and comparing them with those based on the full panels.

Finally, we discuss the implications for a number of issues including the validity of PPP across different exchange rate regimes and the role of the numeraire currency. The implications of the methodological innovations of the paper go beyond the issue of PPP. Clearly, the proposed methodology can be used to consider a number of other topics which require focusing on the stationarity properties of a series.

The next section provides a brief discussion of the evidence and the issues that emerge from

[1] Taylor and Sarno (1998) emphasize this point.

recent studies on PPP that use panel unit root tests. Section 3 describes the methodology for separating stationary from nonstationary and poolable from nonpoolable series. Section 4 discusses the data and section 5 presents and discusses the results of our analysis. Section 6 revisits the "PPP puzzle" using the results of section 5. Finally, Section 7 concludes.

2. A Review of Some Issues Related to PPP

The relevance of PPP for policy purposes is important in both traditional and new approaches in open economy macroeconomics. In the traditional framework for example, whether PPP holds is a valuable piece of information for policymakers who want to assess the effects of a devaluation, since under PPP the effects of the devaluation on competitiveness will disappear in the long-run. In the recent new open economy macroeconomics literature PPP is a required condition for market completeness and the equalization of the marginal utility of home and foreign currency that in turn allows for perfect risk sharing. A stylized fact of the post-Bretton Woods float, however, is the difficulty of distinguishing real exchange rate behavior from random walks and therefore the relatively weak evidence for PPP. Empirical research has successively relied on various methodological approaches to consider the validity of PPP, including cointegration tests for nominal exchange rates and prices, variance ratios tests, long horizon regressions (Serletis and Goras (2004)), quantile regressions (Nikolaou (2008)), and unit root tests on real exchange rate series[2] but despite the voluminous literature the profession's conventional wisdom concerning PPP remains, in general, inconclusive.

Hakkio (1984), Abuaf and Jorion (1991) and Wu (1996) represent early attempts to utilize panel data sets as a means of increasing the power of unit root tests in PPP studies. Tests for unit roots within heterogeneous panels, however, are currently well established, and most of them utilize the frameworks of Levin and Lin (1992), and Im, Pesaran, and Shin (2003) (IPS).[3] Until the emergence of nonstationary panel techniques the evidence supporting the existence of PPP had not only been weak (see Macdonald (1995)) but also lacked robustness. In particular, the results tended to depend on the length of the sample period, the frequency of the series, the choice of countries in the sample, and the choice of numeraire currency. Evidence in favour of PPP was more likely to be found if the tests were based on long samples (of around 100 years) of annual data and if the US dollar was not used as a numeraire (see, e.g., Papell and Theodoridis (2001)). Studies of PPP using panel unit-root tests reversed the relatively gloomy PPP picture. Research focusing on industrial countries provided increased evidence of real exchange stationarity using panel frameworks (see Frankel and Rose (1996), MacDonald (1996), Oh (1996), Papell (1997), Taylor and Sarno (1998) and so on). Despite the increased ability to uncover evidence that validates PPP when panel data are used the existing evidence of panel data studies are still inconclusive. A set of evidence based on panel data methodologies exists that is less favorable to PPP (O'Connell (1998), Papell and Theodoridis (1998), Papell and Theodoridis (2001)). In summary, while the results on balance are supportive of PPP, the fact that a number of studies

[2]For surveys on the stationarity properties of the real exchange rates, see Boucher Breuer (1994), Froot and Rogoff (1995), Mark (2001). See Murphy and Zhu (2008) for a general discussion of empirical irregularities in exchange rates.

[3]Other approaches exist in testing for the presence of unit roots in heterogeneous panels, such as, e.g, Harris and Tzavalis (1999).

employing panel tests fail to always rescue the PPP hypothesis makes the issue more contentious.

A critical issue that emerges when panel unit roots are employed is the problem of cross-sectional dependence. As O'Connell (1998) suggests, the non-zero covariances of the errors across the units in panel tests for unit roots (and cointegration) imply short-run linkages among the units.[4] Using a generalized least squares (GLS) approach to control for intercountry dependence O'Connell produces results that are not supportive to PPP. Subsequent studies that employed GLS, however, -including Papell and Theodoridis (1998) and Taylor and Sarno (1998)- came to the rescue of PPP. Papell (1997), using the Levin and Lin (1992) tests, shows that the rejection of the unit root hypothesis depends critically on the cross-sectional size, and whether or not the critical values have been adjusted to account for serial correlation. Recent advances have provided sophisticated methods which are clearly advantageous to the conventional practice of simply de-meaning the series. Being aware that one cannot completely eliminate cross-sectional dependence, we use some tests that account for this possibility. Our two chosen tests are put forward by Chang (2002) and Pesaran (2003). In Section 5, we provide details on why we choose these two tests.

Many authors, however, have pointed out some fundamental problems in using panel unit-root tests (e.g., Mark (2001), Taylor and Sarno (1998)). In particular, attention has been drawn to the fact that the null hypothesis in such tests is specified as a joint nonstationarity hypothesis. Thus, cases may exist where the panel appears as stationary but a (possibly large) number of individual series display unit roots. In fact, even one stationary series may suffice to reject the unit root null for the whole panel. In this case one may incorrectly conclude that the panel is on balance stationary or -in the best case- they will not be able to distinguish which are the cross-sectional units that display stationarity. While some attempts have been made to circumvent this problem (Taylor and Sarno (1998)), to our knowledge there is no formal procedure available so far that directly considers stationarity of the individual cross-sectional units in a panel framework.

Another closely related dimension of analyzing PPP issues in panels that has received scant -if any- attention refers to the validity of pooling specific sets of real exchange rate series. Applying panel methods on a set of real exchange rates that are not poolable may lead to wrong conclusions. Inappropriate pooling across cross-sectional units, in the case where different real exchange rate series exhibit different rates of convergence, is likely to lead to upwardly biased panel estimators (see Choi, Mark, and Sul (2004)). We avoid such potential pitfalls using a new methodology that tests for the poolability of the series. Our results show that almost all series we find stationary are also poolable.

The ability to separate stationary from nonstationary and poolable from nonpoolable series becomes particularly important when a relatively large number of countries is considered. In such cases the size of the panel and the choice of the countries included can be a contentious issue when standard panels are employed. When discretion is exercised in removing or adding cross-section units in the panel the (summary) result can be affected. Rogoff (1996), for example, expresses reservations along these lines for the 150-country study of Frankel and Rose (1996). Our approach, however, is robust to such problems not only because we provide evidence for

[4]More recently, Banerjee, Marcellino, and Osbat (2003) suggest that since the panel unit root tests assume away the presence of cross-section cointegrating relationships, if this assumption is violated the tests become oversized.

each individual real exchange rate but also because we conduct tests that validate the poolability of the series. The methodological innovations of our analysis render it robust to a number of other weaknesses that plague many PPP studies. Rogoff (1996), for example, questions some favorable-to-PPP results obtained with panel tests on the basis that they include a large number of high-inflation countries. This is a special case where a subsection of cross-section units sharing some specific features drive the results. Providing an analysis of the time series properties of individual series removes any scepticism about the results based on such concerns.

3. Methodology

An attractive feature of panel unit root tests is the ability to exploit coefficient homogeneity under the null hypothesis of a unit root for all series involved in order to obtain a more powerful test of the unit root hypothesis. However, under the alternative hypothesis of heterogeneous panel unit root tests such as, e.g., IPS, of at least one series being stationary, the results are not illuminating enough. In particular if one rejects the unit root hypothesis, one cannot know which series caused the rejection.

We introduce a new procedure to the PPP literature that enables us to distinguish the set of series into a group of stationary and a group of nonstationary series. This method uses a sequence of panel unit root tests to distinguish between stationary and nonstationary series. If more than one series are actually nonstationary then the use of panel methods to investigate the unit root properties of the set of series may indeed be more efficient compared to univariate methods. The proposed method starts by testing the null of all series being unit root processes along the lines considered in many heterogeneous panel unit root tests such as, e.g., the IPS panel unit root test. We use this test as a vehicle for illustrating our method below - which is nevertheless compatible with any other panel unit root test. We first implement this test to all real exchange rates in the panel and if the null is not rejected we accept the nonstationarity hypothesis and the procedure stops. If the null is rejected then we remove from the set of series the one with the minimum individual DF t-test and redo the panel unit root test on the remaining set of series. The procedure is continued until either the test does not reject the null hypothesis or all the series are removed from the set. The end result is a separation of the set of variables into a set of stationary variables and a set of nonstationary variables. The method is presented in detail in subsection .

An additional and highly related issue that emerges when panel data sets are employed, however, is the assumption of poolability, i.e. the validity of the assumption that panel units described by a given model have a common parameter subvector for that model. This assumption is typically being overlooked in the literature. Relevant econometric work, however, has concentrated on whether a given dataset is poolable as a whole, i.e., whether the null hypothesis $H_0 : \beta_j = \beta$, $j = 1, \ldots, N$ holds, where β is the assumed common parameter subvector of the N cross-sectional units of the dataset. In that vein a common approach, discussed, in some detail, in Baltagi (2001), is to use an extension of the Chow (1960) parameter stability test on the pooled dataset. Other tests for this null hypothesis have been developed by Ziemer and Wetzstein (1983) and Baltagi, Hidalgo, and Li (1996).

If such tests reject the null hypothesis, however, the researcher is left with a problem mirroring that of the distinguishing the stationary from nonstationary series in a panel. In other words, although one knows that the null hypothesis of poolability in the panel can be rejected, he cannot identify the series that caused the rejection. Thus, the need for a method that allows the distinction of the set of series into a group of poolable and a group of nonpoolable series emerges. If more than one series are actually poolable then the use of panel methods to investigate the properties of this set of series is indeed more efficient compared to univariate methods. Such methods seem indeed possible and one has been suggested by Kapetanios (2003). This methodology for determining the poolability of the series is discussed, in more detail, in subsection .

3.1. Separating stationary from nonstationary series

We will carry out our analysis using the Im, Pesaran, and Shin (2003) heterogeneous panel unit root test. So we give a few details on the version of the test we use as an expository tool for discussing our method. Consider a sample of N cross sections observed over T time periods.

Let the stochastic process $y_{j,t}$ be generated by

$$y_{j,t} = (1-\phi_j)\mu_j + \phi_j y_{j,t-1} + \epsilon_{j,t}, \quad j = 1, \ldots, N, \quad t = 1, \ldots, T \tag{1}$$

where initial values $y_{j,0}$ are given. We are interested in testing the null hypothesis of $\phi_j = 1$ for all j. Rewriting (1) as

$$\Delta y_{j,t} = (1-\phi_j)\mu_j + \beta_j y_{j,t-1} + \epsilon_{j,t} \tag{2}$$

where $\beta_j = \phi_j - 1$, the null hypothesis becomes $H_0 : \beta_j = 0, \ \forall j$. We make an assumption needed in what follows

Assumption 1 *The $\epsilon_{j,t}$ in (1) are i.i.d. random variables for all j and t with zero means and heterogeneous variances σ_j^2.*

The test is based on the average of individual Dickey-Fuller (DF) statistics. The standard DF statistic for the j-th unit is given by the t-ratio of β_j in the regression of $\Delta \mathbf{y}_j = (\Delta y_{j,1}, \ldots, \Delta y_{j,T})'$ on a matrix of deterministic regressors $\boldsymbol{\tau}_T$ and $\mathbf{y}_j = (y_{j,0}, \ldots, y_{j,T-1})'$. $\boldsymbol{\tau}_T$ could include just a constant, i.e. $\boldsymbol{\tau}_T = (1, \ldots, 1)'$ or a constant and a time trend, i.e. $\boldsymbol{\tau}_T = ((1,1)', (1,2)', \ldots, (1,T)')'$. Denoting the t-statistic by $t_{j,T}$ we have

$$t_{j,T} = \frac{\Delta \mathbf{y}_j' \mathbf{M}_\tau \mathbf{y}_j}{\hat{\sigma}_{j,T} (\mathbf{y}_j' \mathbf{M}_\tau \mathbf{y}_j)^{1/2}} \tag{3}$$

where $\mathbf{M}_\tau = \mathbf{I}_T - \boldsymbol{\tau}_T (\boldsymbol{\tau}_T' \boldsymbol{\tau}_T)^{-1} \boldsymbol{\tau}_T'$ and $\hat{\sigma}_{j,T}^2 = \frac{\Delta \mathbf{y}_j' \mathbf{M}_\tau \Delta \mathbf{y}_j}{T}$. Then, the panel unit root test is based on the following test statistic

$$\bar{t}_T = 1/N \sum_{j=1}^{N} t_{j,T} \tag{4}$$

which we will refer to as the $\bar{t}$-statistic. For one version of the panel unit root test this statistic is normalised to give

$$z_{\bar{t}} = \frac{\sqrt{N}(\bar{t}_T - E(t_{j,T}))}{\sqrt{Var(t_{j,T})}} \tag{5}$$

As Im, Pesaran, and Shin (2003) discuss, this test has a standard normal distribution if $N \to \infty$. $E(t_{j,T})$ and $Var(t_{j,T})$ denote the first and second central moments of the null distribution of $t_{j,T}$. These can be obtained via simulation. Further, for fixed N the distribution of $z_{\bar{t}}$ has no closed form solution. Critical values can be obtained however using simulations as discussed in Im, Pesaran, and Shin (2003). The main asymptotic framework is one where T and N go to infinity but $N/T \to 0$ or where T goes to infinity but N remains fixed. For further use define the following. Let $\mathbf{Y_i} = (\mathbf{y}_{j_1}, \dots, \mathbf{y}_{j_M})$, $\mathbf{i} = \{j_1, \dots, j_M\}$ and $\mathbf{t_i} = (t_{j_1,T}, \dots, t_{j_M,T})'$, for some $M \leq N$, denote subsets of the dataset, the set of indices $\{j_1, ..., j_N\}$ and the vector of unit root test statistics, $(t_{j_1,T}, \dots, t_{j_N,T})'$. Also define $\mathbf{i}^j = \{j\}$, $\mathbf{i}^{1,N} = \{1, \dots, N\}$ and $\mathbf{i}^{-j}$ such that $\mathbf{i}^{-j} \cup \mathbf{i}^j = \mathbf{i}$.

We now define the object we wish to estimate. For every series $y_{j,t}$ define the binary object $\mathcal{I}_j$ which takes the value 0 if $\beta_j = 0$ and 1 if $\beta_j < 0$. We do not consider the case $\beta_j > 0$. Then, $\mathcal{I}_\mathbf{i} = (\mathcal{I}_{j_1}, \dots, \mathcal{I}_{j_M})'$ for some $M \leq N$. We wish to estimate $\mathcal{I}_{\mathbf{i}^{1,N}}$. We denote the estimate by $\hat{\mathcal{I}}_{\mathbf{i}^{1,N}}$. To do so we consider the following algorithm.

1. Set $j = 1$ and $\mathbf{i}_j = \{1, \dots, N\}$.

2. Calculate the $z_{\bar{t}}$-statistic for the set of series $\mathbf{Y}_{\mathbf{i}_j}$. If the test does not reject the null hypothesis $\beta_i = 0$, $i \in \mathbf{i}_j$, stop and set $\hat{\mathcal{I}}_{\mathbf{i}_j} = (0, \dots, 0)'$. If the test rejects go to step (3).

3. Set $\hat{\mathcal{I}}_{\mathbf{i}^l} = 1$ and $\mathbf{i}_{j+1} = \mathbf{i}_j^{-l}$, where l is the index of the series associated with the minimum $t_{s,T}$ over s. Set $j = j + 1$. Go to step (2).

In other words, we estimate a set of binary objects that indicate whether a series is stationary or not. We do this by carrying out a sequence of panel unit root tests on a reducing dataset where the reduction is carried out by dropping series for which there is evidence of stationarity. A low individual t-statistic is used as such evidence. We refer to the new method as Sequential Panel Selection Method (SPSM).

Before discussing the asymptotic properties of SPSM, it is worth stepping back and considering what the advantages of and alternatives to SPSM are. A simple alternative is to disregard the possible panel structure of the dataset and simply use univariate unit root tests to determine the stationarity properties of each series in the dataset. In order to appreciate the distinction between the two alternatives it is important to consider the alternative hypotheses underlying the panel and univariate unit root tests. The alternative hypothesis of the univariate unit root test is well known and does not require comment. The traditional alternative hypothesis of the panel unit root test, as discussed in, e.g., Quah (1995) and Levin and Lin (1992), is that $H_{10} : \beta_i = \beta < 0$. All series in the panel are stationary under this alternative hypothesis. Of course, this alternative hypothesis is quite restrictive because it imposes dynamic homogeneity. As Pesaran and Smith (1995) note, this restriction is firstly highly unlikely to hold in large panels and secondly, once assumed wrongly, it will lead to estimation inconsistency. This homogeneous alternative seems particularly inappropriate in the case of, for example, the purchasing power parity (PPP) hypothesis, where $y_{j,t}$ can be taken to be the real exchange rate. There are no theoretical grounds for the imposition of the homogeneity hypothesis, $\beta_i = \beta$, under PPP. Thus, a more general alternative hypothesis is suggested by Im, Pesaran, and Shin (2003) which is

$H_{11} : \beta_i < 0,\ i = 1, ..., N_1;\ \beta_i = 0,\ i = N_1 + 1, ..., N$. This hypothesis is more flexible and allows a mix of stationary and nonstationary series.

When H_{10} is entertained, the motivation for a panel unit root test is superficially clear. The panel unit root test will be much more powerful than the univariate unit root tests. But the problem here is that the alternative hypotheses of the panel and univariate unit root tests are not the same. Rejection of the null hypothesis for the panel unit root test does not say anything about individual series because the panel unit root test has power against H_{11} as well and therefore one cannot conclude that all series of a panel, for which the panel unit root has rejected the null hypothesis, are stationary. A further problem is that under H_{11} the power of the panel unit root test maybe quite low depending on N_1. The new method suggested in this paper can help. It does so in the following way. Unlike panel unit root tests, the focus of SPSM is clearly on the properties of individual series like univariate unit root tests. Therefore, the outputs of SPSM and univariate tests are comparable. A major aim of both methods is to uncover stationarity in the data. If the panel unit root tests, underlying SPSM, are more powerful than individual unit root tests then, it is likely than more series will be correctly identified to be stationary, compared to univariate tests.

To see this, we consider a simple experiment. Let all series in a dataset of, say, 100 units, be stationary. Then, using results from Table 4 of Im, Pesaran, and Shin (2003) on the power of the IPS panel unit root test, and assuming, for the moment, that sequential tests are independent we can derive the number of series that will be found on average to be stationary using either SPSM or individual Dickey-Fuller tests. For SPSM the average number of stationary series will be $\sum_{i=1}^{100} \left(\prod_{j=1}^{i} \pi_{100-j+1} \right) (1 - \pi_{100-i})\, i$, where π_i is the power of the panel unit root test for a dataset with i units which are all stationary[5] and $\pi_0 = 0$. For the univariate test, the average number of stationary series is simply the power of the test expressed as a percentage. For SPSM we get that the average number of series found stationary, for T=25,50,100, to be 17.4, 81.8 and 97.3 respectively. The numbers for the univariate test are: 9.1, 15.1 and 35.1. Clearly, SPSM does much better here. There are a number of problems with this naive approach. Firstly, the sequence of tests are obviously not independent. But accounting for the dependence of the tests analytically is very difficult. A Monte Carlo study can be informative in this respect. Secondly, the setup is one where the alternative hypothesis for the panel unit root test is H_{10}. The panel unit root test will be much less powerful for H_{11}. Therefore, the above naive experiment is just illustrative of the underlying idea that gives rise to SPSM. The asymptotic analysis carried out in the next subsection, together with the Monte Carlo study in subsection will provide a much more rigorous evaluation of SPSM compared with the simple alternative of using univariate unit root tests.

The above discussion makes clear that SPSM while clearly useful, also has limitations, depending on the number of, and extent to which, series are stationary. An ideal situation for SPSM is one where most series considered are stationary and very persistent. Then, it is likely that univariate tests cannot reject the univariate unit root hypothesis, whereas panel unit root tests with their superior power, in this case, can reject the panel unit root hypothesis. The other

[5]Table 4 of Im, Pesaran, and Shin (2003) does not provide estimated power for all values of N. We linearly interpolate from the published numbers to get the missing values.

extreme scenario which does not favour SPSM, is one where most series have a unit root but a minority are stationary and not persistent. Then, whereas univariate tests will likely uncover evidence for stationarity in the stationary series, the panel unit root test will not reject thus leading SPSM to find no evidence of stationarity. Of course, in most cases intermediate scenaria will hold. As a result SPSM is best viewed as an addition to univariate unit root tests, that can provide added value in a number of circumstances, where inference for individuals series is of interest. It is probably the case that real exchange rate panel datasets, where the overwhelming majority of the series are very persistent but, given PPP, few if any are unit root processes, fall more closely to the first extreme scenario discussed above, thereby making SPSM a useful alternative to univariate unit root tests.

3.1.1. Theoretical results

In this section we discuss conditions for the consistency of $\hat{\mathcal{I}}_{\mathbf{i}^{1,N}}$ as an estimator of $\mathcal{I}_{\mathbf{i}^{1,N}}$, both for finite and infinite N, where in the latter case $N/T \rightarrow 0$. Proofs are relegated to the Appendix. Formally, we will show that

Theorem 1 *Under assumption 1 and if (i)* $\lim_{T\rightarrow\infty} \alpha_T = 0$*, (ii)* $\lim_{T\rightarrow\infty} \frac{\ln \alpha_T}{\sqrt{T/N}} = 0$*, where* α_T *is the significance level used for the panel unit root test and (iii)* $N/T \rightarrow 0$ *then*

$$\lim_{T\rightarrow\infty} Pr\left(\sum_{j=1}^{N} |\hat{\mathcal{I}}_{\mathbf{i}^j} - \mathcal{I}_{\mathbf{i}^j}| > 0\right) = 0.$$

Note the similarities between this setup and the variety of tests of rank where a sequence of tests are needed to determine the rank of a matrix (see e.g., Camba-Mendez and Kapetanios (2001)). A weaker result can be established if the significance level, denoted now α, is kept fixed, for infinite N. From before, we note that N_1 denotes the number of stationary series in the panel. We define $N_2 = N - N_1$.

Theorem 2 *Under assumption 1, if* $N, T \rightarrow \infty$*, and* $N/T \rightarrow 0$ *then (i) if* $\mathcal{I}_{\mathbf{i}^j} = 1$*, then* $\lim_{T\rightarrow\infty} Pr(\hat{\mathcal{I}}_{\mathbf{i}^j} = 1) = 1$ *and (ii) if* $\mathcal{I}_{\mathbf{i}^j} = 0$*, then there exists finite* k *such that*

$$\lim_{N_2\rightarrow\infty} \lim_{T\rightarrow\infty} Pr\left(\sum_{\mathcal{I}_{\mathbf{i}^j}=0}^{N_2} |\hat{\mathcal{I}}_{\mathbf{i}^j} - \mathcal{I}_{\mathbf{i}^j}| > k\right) = 0.$$

It is clear that our procedure is very general. It can be applied using any heterogeneous panel unit root test as long as the test satisfies conditions similar to those needed for Theorems 1 and 2. The main ingredients are a panel unit root test and a criterion for choosing which series to classify as stationary at each step in the sequence of tests. Our choices of the Im, Pesaran, and Shin (2003) test for the panel unit root test and the minimum individual t-test, as expository vehicles for the new methodology, seem relatively uncontroversial. Further, a number of other possibilities arise. A reverse search using the panel equivalent of the KPSS test as developed by Shin and Snell (2003) could be envisaged as well. Recent work in the panel unit root test literature has suggested that it is important to take into account cross-sectional dependence when carrying out panel unit root tests. Such tests have been proposed in the literature by Bai and Ng (2005), Chang (2002), Harvey and Bates (2002), Moon and Perron (2004), Phillips and Sul (2002) and Pesaran (2003). As is clear from the above discussion any of these tests can be straightforwardly plugged into our procedure and provide inference on individual series that is likely to be more effective in detecting stationarity than individual unit root tests, if a significant proportion of the series are stationary. In fact, in our empirical work we consider two panel unit root tests of those listed above (those by Chang (2002) and Pesaran (2003)) and apply our methodology to them.

Extending the method to consider models with possibly serially correlated errors is straightforward following, e.g., Im, Pesaran, and Shin (2003). More specifically, assuming that the data are generated by individual ADF(p) regressions

$$\Delta y_{j,t} = a_j + \phi_j y_{j,t-1} + \sum_{s=1}^{p_j} \rho_{j,s} \Delta y_{j,t-s} + \epsilon_{j,t}, \quad j = 1, \ldots, N, \quad t = 1, \ldots, T \tag{6}$$

we can write these regressions as

$$\Delta \mathbf{y}_j = \beta_j \mathbf{y}_j + \mathbf{Q}_j \boldsymbol{\gamma}_j + \boldsymbol{\epsilon}_j \tag{7}$$

where $\mathbf{Q}_j = (\boldsymbol{\tau}_T, \Delta \mathbf{y}_{j,-1}, \ldots, \Delta \mathbf{y}_{j,-p})$ and $\boldsymbol{\gamma}_j = (a_j, \rho_{j,1}, \ldots, \rho_{j,p_j})'$. Then, the $\bar{t}_T$ statistic is given by $\frac{1}{N}\sum_{j=1}^{N} t_{j,T}(p_j, \boldsymbol{\rho}_j)$ where $t_{j,T}(p_j, \boldsymbol{\rho}_j)$ is given by

$$t_{j,T}(p_j, \boldsymbol{\rho}_j) = \frac{\sqrt{T - p_j - 2}(\mathbf{y}_j' \mathbf{M}_{\mathbf{Q}_j} \Delta \mathbf{y}_j)}{(\mathbf{y}_j' \mathbf{M}_{\mathbf{Q}_j} \mathbf{y}_j)^{1/2} (\Delta \mathbf{y}_j' \mathbf{M}_{\mathbf{X}_j} \Delta \mathbf{y}_j)^{1/2}} \tag{8}$$

where $\boldsymbol{\rho}_j = (\rho_{j,1}, \ldots, \rho_{j,p_j})'$, $\mathbf{M}_{\mathbf{Q}_j} = \mathbf{I}_T - \mathbf{Q}_j(\mathbf{Q}_j'\mathbf{Q}_j)^{-1}\mathbf{Q}_j'$, $\mathbf{M}_{\mathbf{X}_j} = \mathbf{I}_T - \mathbf{X}_j(\mathbf{X}_j'\mathbf{X}_j)^{-1}\mathbf{X}_j'$ and $\mathbf{X}_j = (\mathbf{y}_j, \mathbf{Q}_j)$. Obviously for fixed T the distributions of the individual t-statistics involve nuisance parameters whose influence however disappears as T tends to infinity. This occurs even if N remains fixed. Im, Pesaran, and Shin (2003) suggest the use of the following normalised statistic to carry out the panel unit root test.

$$z_{\bar{t}}(\boldsymbol{p}) = \frac{\sqrt{N}\bar{t}_T - E(t_{j,T}(p_j, 0)|\beta_j = 0)}{\sqrt{Var(t_{j,T}(p_j, 0)|\beta_j = 0)}} \tag{9}$$

This converges to $N(0,1)$ if T and then N tend to infinity. However, even if only T tends to infinity the above statistic tends to a nuisance parameter free distribution which only depends on N.

Before presenting our Monte Carlo study we present simulation estimates of $E(t_T)$ and $Var(t_T)$ and the 5% critical values of the $z_{\bar{t}}$ test. For all the results simulations with 10000 replications have been used. We present estimates for $E(t_{j,T}(p_j, 0)|\beta_j = 0)$ and $Var(t_{j,T}(p_j, 0)|\beta_j = 0)$ for $p_j = 0, 1$ for $T \in \{10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 400\}$ in Table 1. Estimates for $E(t_{j,T}(p_j, 0)|\beta_j = 0)$ and $Var(t_{j,T}(p_j, 0)|\beta_j = 0)$ for $p_j = 2, \ldots, 8$ and $T = 100, 1000$ are similar to those presented and are therefore not reported but are available upon request. Critical values for the $z_{\bar{t}}$ for $T \in \{10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 400\}$, $N \in \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30\}$ and $p_j \in \{0, 1\}$ and for $z_{\bar{t}}$ for $T \in \{100, 1000\}$, $N \in \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30\}$ and $p_j \in \{2, 3, 4, 5, 6, 7, 8\}$ have been obtained and are quite close to the N-asymptotic normal critical values. As a result they are not reported but are available upon request.

3.1.2. Monte Carlo study

In this section we carry out a Monte Carlo investigation of our new method. We consider the following setup. Let

$$y_{j,t} = \phi_j y_{j,t-1} + \epsilon_{j,t}, j = 1, \ldots, N, \quad t = 1, \ldots, T \tag{10}$$

where $\epsilon_{j,t} \sim N(0,1)$. We investigate the new method along a number of different dimensions for the above model. Namely, we consider variations in N, T and ϕ_j. More specifically, we consider $T \in \{30, 50, 150, 400\}$ and $N \in \{5, 10, 15, 20, 25, 30\}$. For ϕ_j we consider the following setup: $\phi_j = 1$ with probability δ over j and $\phi_j \sim U(\gamma_1, \gamma_2)$ with probability $1 - \delta$. This is a general setup designed to address a number of issues not widely discussed in the literature. As this is a heterogeneous panel allowing variation in ϕ_j under the alternative hypothesis is of great importance. Further, the choice of δ is likely to affect the performance of the new method. We set $\delta \in \{0.05, 0.2, 0.5\}$.

Further we consider two overall experiment groups labeled experiment group A and experiment group B. For experiment group A, $\gamma_1 = 0.85$ and $\gamma_2 = 0.95$. For experiment group B, $\gamma_1 = 0.75$ and $\gamma_2 = 0.85$. Due to space constraints, we do not report results for Setup B in the Tables since the results suggest a better ability to separate stationary from nonstationary series for this setup, compared to Setup A, as expected. These results are available upon request from the authors. Finally, we carry out the whole analysis for $p_j = 0$[6]. We expect that our method will be able to identify the stationary series when δ is low since then there are many stationary series and therefore the power of the heterogeneous panel unit root test is likely to be higher. The performance measure we use is the estimated probability of classifying a series as stationary. This should tend to zero for nonstationary series and to one for stationary series. Denote the number of Monte Carlo replications by B. B is set to 1000. This probability is calculated as follows in our experiments $\hat{P}(\hat{\mathcal{I}}_{\mathbf{i}^u} = 1 | \mathcal{I}_{\mathbf{i}^u} = s) = \frac{1}{N_s B} \sum_{b=1}^{B} \sum_{\mathcal{I}_{\mathbf{i}^q} = s} \hat{\mathcal{I}}_{\mathbf{i}^q}^b$, where $N_s = N(1-\delta)s + N\delta(1-s)$ and u denotes a generic series. As an alternative method of determining the stationarity or not of the set of series we consider the standard DF test for each series. Results are presented in Tables 2 and 3.

A number of conclusions emerge from these Tables. Firstly, we note that the performance of SPSM in terms of classifying I(1) series as I(1) is in general satisfactory. The probability of misclassification never exceeds 15%. This is to be expected given that the method is based on a test whose null hypothesis is that of a set of series being I(1). On the other hand, as the number of observations increases we see that this probability falls especially for $\delta = 0.5$. This indicates that the result in (ii) of Theorem 2 holds. Moving on to the ability of SPSM to classify I(0) series as I(0) we see that the probability of that happening increases drastically with T and substantially with N as expected. It also decreases with respect to δ. This is expected as well. When there is a large proportion of I(1) series in the dataset, the panel unit root test is less powerful as the I(1) series cause a deterioration in power. Therefore, the method stops when I(0) series are still in the dataset causing the observed patterns for the estimated probability of finding an I(0) series to be I(0).

[6]Results for the case $p_j = 1$, keeping the rest of the Monte Carlo design as presented here, have also been obtained. They are very similar to the case $p_j = 0$ and are therefore not reported. They are, however, available from the authors upon request.

Table 1: Estimated $E(t_{j,T}(p_j,0)|\beta_j=0)$ and $Var(t_{j,T}(p_j,0)|\beta_j=0)$ for $p_j \in \{0,1\}$

	p=0						p=1					
	$E(t_{j,T})$			$Var(t_{j,T})$			$E(t_{j,T})$			$Var(t_{j,T})$		
T	DF 1	DF 2	DF 3	DF 1	DF 2	DF 3	DF 1	DF 2	DF 3	DF 1	DF 2	DF 3
10	-0.349	-1.501	-2.161	1.074	1.072	1.121	-0.396	-1.499	-2.150	1.032	1.030	1.052
15	-0.375	-1.508	-2.172	1.034	0.953	0.935	-0.386	-1.497	-2.169	1.017	0.976	0.934
20	-0.387	-1.508	-2.167	1.012	0.934	0.878	-0.386	-1.487	-2.179	1.009	0.948	0.906
25	-0.399	-1.516	-2.182	1.010	0.907	0.850	-0.414	-1.514	-2.172	1.003	0.918	0.868
30	-0.394	-1.517	-2.160	1.006	0.894	0.836	-0.406	-1.524	-2.180	1.013	0.906	0.841
40	-0.418	-1.540	-2.184	0.994	0.870	0.808	-0.403	-1.506	-2.171	1.005	0.900	0.828
50	-0.391	-1.517	-2.170	0.996	0.870	0.806	-0.434	-1.537	-2.179	0.996	0.886	0.816
60	-0.415	-1.524	-2.179	0.993	0.872	0.789	-0.420	-1.524	-2.177	0.997	0.880	0.793
70	-0.420	-1.514	-2.178	0.986	0.863	0.786	-0.416	-1.532	-2.186	0.991	0.878	0.794
80	-0.404	-1.527	-2.172	0.983	0.863	0.780	-0.416	-1.523	-2.183	0.995	0.872	0.787
90	-0.404	-1.530	-2.174	0.989	0.864	0.776	-0.421	-1.539	-2.188	0.999	0.864	0.780
100	-0.405	-1.517	-2.177	0.995	0.853	0.768	-0.427	-1.533	-2.176	0.975	0.859	0.784
150	-0.417	-1.531	-2.183	0.989	0.845	0.768	-0.421	-1.524	-2.173	0.980	0.847	0.767
200	-0.416	-1.523	-2.174	0.994	0.848	0.768	-0.407	-1.525	-2.184	0.988	0.854	0.754
400	-0.433	-1.537	-2.169	0.968	0.830	0.747	-0.426	-1.543	-2.182	0.983	0.839	0.749

As usual, SPSM based on DF 1 finds more series being I(0) compared to SPSM based on DF 2 or DF 3. When compared to DF we see that for low δ SPSM does better since it misclassifies fewer series on average. This can be seen by adding the probability of finding an I(1) to be I(0) and one minus the probability of finding an I(0) series to be I(0). So for $\delta = 0.05, 0.2$ SPSM does better than DF especially for samples of 150 observations which is a relevant sample size for econometric work. For samples of 400 observations both methods do well as expected. When we look at datasets with $\delta = 0.5$ DF does better. Again this is to be expected since the ability of SPSM to find an I(0) to be I(0) decreases with δ. Of course, δ does not affect the performance of DF.

We note a couple of things about this comparison here. Firstly, the DF test is not a consistent estimator of $\mathcal{I}_{\mathbf{i}^1,N}$ neither as N or T go to infinity. Even for infinite T it will reject the null even if it is true as long as the significance level is not 0. Of course it can be made consistent by making the significance level of the test depend on T. This may be problematic because we do not know the power performance of the DF in this case. In any case DF does not improve in performance when N increases. Here the importance of the panel dimension is clear. To make our analysis more concrete we have increased N to 200 and 400 and redid the $p = 0$, Setup A, $\delta = 0.5$ experiment for $T = 50$. Results are presented in Table 4. As we can see SPSM does clearly better than DF.

3.2. Separating poolable from nonpoolable series

This subsection gives a brief outline of the method suggested in Kapetanios (2003). To illus-

Table 2: SPSM, $p = 0$, Setup A[a]

%I(1)	(N,T)	DF 1				DF 2				DF 3			
		030	050	150	400	030	050	150	400	030	050	150	400
0.05	5	$\binom{0.099}{0.208}$	$\binom{0.090}{0.346}$	$\binom{0.053}{0.842}$	$\binom{0.057}{0.965}$	$\binom{0.032}{0.044}$	$\binom{0.052}{0.080}$	$\binom{0.077}{0.636}$	$\binom{0.053}{0.922}$	$\binom{0.021}{0.025}$	$\binom{0.029}{0.034}$	$\binom{0.108}{0.440}$	$\binom{0.048}{0.873}$
	10	$\binom{0.107}{0.321}$	$\binom{0.119}{0.568}$	$\binom{0.072}{0.832}$	$\binom{0.052}{0.975}$	$\binom{0.036}{0.049}$	$\binom{0.067}{0.166}$	$\binom{0.104}{0.587}$	$\binom{0.061}{0.943}$	$\binom{0.012}{0.016}$	$\binom{0.039}{0.063}$	$\binom{0.100}{0.390}$	$\binom{0.070}{0.916}$
	15	$\binom{0.111}{0.340}$	$\binom{0.144}{0.599}$	$\binom{0.081}{0.886}$	$\binom{0.049}{0.985}$	$\binom{0.034}{0.050}$	$\binom{0.077}{0.171}$	$\binom{0.141}{0.692}$	$\binom{0.046}{0.964}$	$\binom{0.007}{0.015}$	$\binom{0.032}{0.055}$	$\binom{0.121}{0.506}$	$\binom{0.053}{0.944}$
	20	$\binom{0.154}{0.398}$	$\binom{0.160}{0.630}$	$\binom{0.096}{0.906}$	$\binom{0.053}{0.978}$	$\binom{0.039}{0.052}$	$\binom{0.088}{0.176}$	$\binom{0.111}{0.737}$	$\binom{0.060}{0.947}$	$\binom{0.009}{0.013}$	$\binom{0.025}{0.048}$	$\binom{0.120}{0.566}$	$\binom{0.080}{0.914}$
	25	$\binom{0.147}{0.415}$	$\binom{0.164}{0.679}$	$\binom{0.075}{0.901}$	$\binom{0.035}{0.976}$	$\binom{0.038}{0.060}$	$\binom{0.116}{0.241}$	$\binom{0.112}{0.753}$	$\binom{0.033}{0.956}$	$\binom{0.012}{0.015}$	$\binom{0.047}{0.074}$	$\binom{0.133}{0.598}$	$\binom{0.055}{0.934}$
	30	$\binom{0.159}{0.472}$	$\binom{0.171}{0.682}$	$\binom{0.084}{0.911}$	$\binom{0.029}{0.978}$	$\binom{0.055}{0.073}$	$\binom{0.123}{0.230}$	$\binom{0.142}{0.757}$	$\binom{0.039}{0.960}$	$\binom{0.017}{0.016}$	$\binom{0.051}{0.062}$	$\binom{0.139}{0.598}$	$\binom{0.055}{0.938}$
0.20	5	$\binom{0.077}{0.161}$	$\binom{0.099}{0.441}$	$\binom{0.075}{0.711}$	$\binom{0.052}{0.979}$	$\binom{0.027}{0.039}$	$\binom{0.058}{0.106}$	$\binom{0.081}{0.416}$	$\binom{0.047}{0.953}$	$\binom{0.020}{0.024}$	$\binom{0.030}{0.044}$	$\binom{0.072}{0.241}$	$\binom{0.054}{0.920}$
	10	$\binom{0.086}{0.175}$	$\binom{0.107}{0.535}$	$\binom{0.054}{0.784}$	$\binom{0.034}{0.939}$	$\binom{0.021}{0.027}$	$\binom{0.067}{0.152}$	$\binom{0.079}{0.576}$	$\binom{0.034}{0.901}$	$\binom{0.011}{0.010}$	$\binom{0.037}{0.052}$	$\binom{0.084}{0.385}$	$\binom{0.047}{0.866}$
	15	$\binom{0.103}{0.318}$	$\binom{0.105}{0.518}$	$\binom{0.063}{0.809}$	$\binom{0.017}{0.943}$	$\binom{0.033}{0.047}$	$\binom{0.073}{0.128}$	$\binom{0.093}{0.598}$	$\binom{0.028}{0.911}$	$\binom{0.011}{0.014}$	$\binom{0.032}{0.039}$	$\binom{0.089}{0.416}$	$\binom{0.037}{0.884}$
	20	$\binom{0.105}{0.265}$	$\binom{0.120}{0.520}$	$\binom{0.062}{0.824}$	$\binom{0.014}{0.949}$	$\binom{0.021}{0.031}$	$\binom{0.069}{0.124}$	$\binom{0.092}{0.620}$	$\binom{0.022}{0.923}$	$\binom{0.009}{0.011}$	$\binom{0.024}{0.039}$	$\binom{0.091}{0.439}$	$\binom{0.027}{0.905}$
	25	$\binom{0.119}{0.386}$	$\binom{0.118}{0.578}$	$\binom{0.057}{0.865}$	$\binom{0.013}{0.951}$	$\binom{0.039}{0.058}$	$\binom{0.082}{0.181}$	$\binom{0.095}{0.698}$	$\binom{0.020}{0.925}$	$\binom{0.010}{0.013}$	$\binom{0.035}{0.051}$	$\binom{0.112}{0.539}$	$\binom{0.033}{0.898}$
	30	$\binom{0.130}{0.357}$	$\binom{0.137}{0.533}$	$\binom{0.056}{0.857}$	$\binom{0.011}{0.947}$	$\binom{0.038}{0.048}$	$\binom{0.073}{0.131}$	$\binom{0.106}{0.689}$	$\binom{0.028}{0.918}$	$\binom{0.011}{0.012}$	$\binom{0.024}{0.031}$	$\binom{0.107}{0.529}$	$\binom{0.046}{0.885}$
0.50	5	$\binom{0.032}{0.102}$	$\binom{0.042}{0.192}$	$\binom{0.019}{0.609}$	$\binom{0.018}{0.856}$	$\binom{0.026}{0.030}$	$\binom{0.035}{0.059}$	$\binom{0.031}{0.434}$	$\binom{0.026}{0.811}$	$\binom{0.021}{0.018}$	$\binom{0.016}{0.029}$	$\binom{0.040}{0.281}$	$\binom{0.021}{0.777}$
	10	$\binom{0.045}{0.099}$	$\binom{0.052}{0.297}$	$\binom{0.020}{0.709}$	$\binom{0.012}{0.850}$	$\binom{0.019}{0.017}$	$\binom{0.034}{0.074}$	$\binom{0.042}{0.512}$	$\binom{0.014}{0.787}$	$\binom{0.011}{0.011}$	$\binom{0.018}{0.027}$	$\binom{0.043}{0.379}$	$\binom{0.021}{0.730}$
	15	$\binom{0.041}{0.131}$	$\binom{0.051}{0.216}$	$\binom{0.027}{0.599}$	$\binom{0.010}{0.832}$	$\binom{0.016}{0.023}$	$\binom{0.022}{0.041}$	$\binom{0.044}{0.335}$	$\binom{0.010}{0.783}$	$\binom{0.012}{0.009}$	$\binom{0.011}{0.016}$	$\binom{0.038}{0.185}$	$\binom{0.015}{0.729}$
	20	$\binom{0.049}{0.115}$	$\binom{0.059}{0.248}$	$\binom{0.028}{0.707}$	$\binom{0.007}{0.900}$	$\binom{0.013}{0.015}$	$\binom{0.024}{0.041}$	$\binom{0.052}{0.471}$	$\binom{0.006}{0.879}$	$\binom{0.006}{0.005}$	$\binom{0.010}{0.014}$	$\binom{0.048}{0.312}$	$\binom{0.009}{0.846}$
	25	$\binom{0.056}{0.167}$	$\binom{0.065}{0.343}$	$\binom{0.026}{0.687}$	$\binom{0.008}{0.878}$	$\binom{0.014}{0.020}$	$\binom{0.033}{0.071}$	$\binom{0.044}{0.467}$	$\binom{0.009}{0.839}$	$\binom{0.006}{0.008}$	$\binom{0.016}{0.025}$	$\binom{0.046}{0.308}$	$\binom{0.013}{0.806}$
	30	$\binom{0.060}{0.165}$	$\binom{0.070}{0.353}$	$\binom{0.020}{0.753}$	$\binom{0.006}{0.880}$	$\binom{0.016}{0.020}$	$\binom{0.035}{0.070}$	$\binom{0.045}{0.574}$	$\binom{0.011}{0.841}$	$\binom{0.005}{0.007}$	$\binom{0.013}{0.022}$	$\binom{0.052}{0.415}$	$\binom{0.015}{0.798}$

[a]%I(1) denotes the proportion of series which are I(1). For the notation $\binom{a}{b}$ we have that a gives the probability that an I(1) series will be classified as I(0), whereas b gives the probability that an I(0) series will be classified as I(0).

trate the methodology, consider the following panel data model

$$y_{j,t} = \alpha_j + \beta_j x_{j,t} + \epsilon_{j,t}, \quad j = 1, \ldots, N, \quad t = 1, \ldots, T. \tag{11}$$

where $x_{j,t}$ is a k-dimensional vector of predetermined variables. This is a standard panel data model where we do not need to specify the nature of the cross sectional individual effect α_j. Our discussion carries through both for fixed and random effect models. The poolability test is concerned with the null hypothesis $H_0 : \beta_j = \beta, \ \ \forall j$. A test that $\beta_j = \beta$ for a given j may be based on the test statistic

$$S_{T,j} = (\hat{\beta}_j - \tilde{\beta})' Var(\hat{\beta}_j - \tilde{\beta})^{-1} (\hat{\beta}_j - \tilde{\beta}) \tag{12}$$

This is a Haussman type statistic. If the panel estimator, $\tilde{\beta}$, were efficient then, under the null hypothesis we know from Hausman (1978) that $Var(\hat{\beta}_j - \tilde{\beta}) = Var(\hat{\beta}_j) - Var(\tilde{\beta})$. However, the estimator is not assumed to be efficient and hence the variance is given by $Var(\hat{\beta}_j - \tilde{\beta}) =$

Table 3: DF, $p = 0$, Setup A[a]

		DF 1				DF 2				DF 3			
%I(1)	(N, T)	030	050	150	400	030	050	150	400	030	050	150	400
	5	$\binom{0.054}{0.201}$	$\binom{0.047}{0.310}$	$\binom{0.050}{0.973}$	$\binom{0.056}{1.000}$	$\binom{0.067}{0.096}$	$\binom{0.058}{0.130}$	$\binom{0.055}{0.665}$	$\binom{0.052}{0.996}$	$\binom{0.081}{0.093}$	$\binom{0.067}{0.096}$	$\binom{0.068}{0.441}$	$\binom{0.049}{0.969}$
	10	$\binom{0.043}{0.185}$	$\binom{0.042}{0.393}$	$\binom{0.045}{0.822}$	$\binom{0.052}{1.000}$	$\binom{0.053}{0.100}$	$\binom{0.043}{0.156}$	$\binom{0.049}{0.479}$	$\binom{0.059}{0.992}$	$\binom{0.061}{0.087}$	$\binom{0.056}{0.118}$	$\binom{0.056}{0.335}$	$\binom{0.058}{0.961}$
0.05	15	$\binom{0.036}{0.166}$	$\binom{0.056}{0.345}$	$\binom{0.052}{0.867}$	$\binom{0.049}{1.000}$	$\binom{0.056}{0.094}$	$\binom{0.048}{0.139}$	$\binom{0.055}{0.532}$	$\binom{0.043}{0.998}$	$\binom{0.057}{0.085}$	$\binom{0.056}{0.110}$	$\binom{0.061}{0.372}$	$\binom{0.050}{0.982}$
	20	$\binom{0.050}{0.170}$	$\binom{0.044}{0.310}$	$\binom{0.047}{0.874}$	$\binom{0.052}{1.000}$	$\binom{0.056}{0.091}$	$\binom{0.049}{0.128}$	$\binom{0.047}{0.574}$	$\binom{0.047}{0.972}$	$\binom{0.069}{0.087}$	$\binom{0.059}{0.100}$	$\binom{0.051}{0.412}$	$\binom{0.057}{0.912}$
	25	$\binom{0.041}{0.173}$	$\binom{0.053}{0.370}$	$\binom{0.047}{0.896}$	$\binom{0.051}{1.000}$	$\binom{0.053}{0.098}$	$\binom{0.063}{0.152}$	$\binom{0.045}{0.596}$	$\binom{0.049}{0.995}$	$\binom{0.073}{0.088}$	$\binom{0.067}{0.116}$	$\binom{0.053}{0.417}$	$\binom{0.052}{0.964}$
	30	$\binom{0.049}{0.180}$	$\binom{0.047}{0.327}$	$\binom{0.051}{0.878}$	$\binom{0.051}{1.000}$	$\binom{0.070}{0.094}$	$\binom{0.066}{0.130}$	$\binom{0.052}{0.569}$	$\binom{0.053}{0.991}$	$\binom{0.081}{0.088}$	$\binom{0.071}{0.106}$	$\binom{0.051}{0.392}$	$\binom{0.046}{0.957}$
	5	$\binom{0.055}{0.178}$	$\binom{0.047}{0.394}$	$\binom{0.053}{0.799}$	$\binom{0.052}{1.000}$	$\binom{0.057}{0.095}$	$\binom{0.054}{0.147}$	$\binom{0.039}{0.434}$	$\binom{0.047}{1.000}$	$\binom{0.069}{0.090}$	$\binom{0.047}{0.104}$	$\binom{0.056}{0.288}$	$\binom{0.055}{0.998}$
	10	$\binom{0.051}{0.142}$	$\binom{0.052}{0.402}$	$\binom{0.042}{0.870}$	$\binom{0.056}{1.000}$	$\binom{0.060}{0.087}$	$\binom{0.059}{0.163}$	$\binom{0.050}{0.544}$	$\binom{0.057}{0.990}$	$\binom{0.075}{0.079}$	$\binom{0.065}{0.122}$	$\binom{0.063}{0.363}$	$\binom{0.055}{0.955}$
0.20	15	$\binom{0.048}{0.193}$	$\binom{0.050}{0.338}$	$\binom{0.047}{0.852}$	$\binom{0.048}{1.000}$	$\binom{0.062}{0.096}$	$\binom{0.056}{0.130}$	$\binom{0.053}{0.524}$	$\binom{0.053}{0.993}$	$\binom{0.067}{0.089}$	$\binom{0.066}{0.102}$	$\binom{0.057}{0.361}$	$\binom{0.057}{0.959}$
	20	$\binom{0.045}{0.152}$	$\binom{0.047}{0.297}$	$\binom{0.053}{0.859}$	$\binom{0.045}{1.000}$	$\binom{0.065}{0.087}$	$\binom{0.056}{0.127}$	$\binom{0.049}{0.511}$	$\binom{0.048}{0.991}$	$\binom{0.070}{0.084}$	$\binom{0.063}{0.097}$	$\binom{0.051}{0.344}$	$\binom{0.049}{0.968}$
	25	$\binom{0.046}{0.193}$	$\binom{0.047}{0.349}$	$\binom{0.054}{0.915}$	$\binom{0.046}{1.000}$	$\binom{0.063}{0.101}$	$\binom{0.058}{0.143}$	$\binom{0.050}{0.589}$	$\binom{0.048}{0.994}$	$\binom{0.069}{0.089}$	$\binom{0.061}{0.107}$	$\binom{0.053}{0.400}$	$\binom{0.058}{0.968}$
	30	$\binom{0.051}{0.175}$	$\binom{0.048}{0.280}$	$\binom{0.049}{0.883}$	$\binom{0.046}{1.000}$	$\binom{0.063}{0.092}$	$\binom{0.055}{0.117}$	$\binom{0.053}{0.568}$	$\binom{0.053}{0.988}$	$\binom{0.072}{0.086}$	$\binom{0.060}{0.095}$	$\binom{0.049}{0.394}$	$\binom{0.057}{0.941}$
	5	$\binom{0.048}{0.246}$	$\binom{0.052}{0.445}$	$\binom{0.049}{0.996}$	$\binom{0.050}{1.000}$	$\binom{0.060}{0.114}$	$\binom{0.059}{0.162}$	$\binom{0.049}{0.817}$	$\binom{0.052}{1.000}$	$\binom{0.077}{0.099}$	$\binom{0.057}{0.123}$	$\binom{0.055}{0.607}$	$\binom{0.051}{1.000}$
	10	$\binom{0.051}{0.164}$	$\binom{0.050}{0.390}$	$\binom{0.047}{0.971}$	$\binom{0.047}{1.000}$	$\binom{0.063}{0.085}$	$\binom{0.056}{0.155}$	$\binom{0.053}{0.752}$	$\binom{0.046}{0.992}$	$\binom{0.068}{0.084}$	$\binom{0.064}{0.116}$	$\binom{0.055}{0.547}$	$\binom{0.051}{0.955}$
0.50	15	$\binom{0.043}{0.200}$	$\binom{0.050}{0.283}$	$\binom{0.048}{0.842}$	$\binom{0.048}{1.000}$	$\binom{0.056}{0.100}$	$\binom{0.056}{0.123}$	$\binom{0.046}{0.467}$	$\binom{0.052}{0.985}$	$\binom{0.073}{0.085}$	$\binom{0.059}{0.102}$	$\binom{0.053}{0.307}$	$\binom{0.052}{0.946}$
	20	$\binom{0.053}{0.149}$	$\binom{0.047}{0.258}$	$\binom{0.049}{0.907}$	$\binom{0.047}{1.000}$	$\binom{0.063}{0.088}$	$\binom{0.055}{0.113}$	$\binom{0.054}{0.564}$	$\binom{0.049}{1.000}$	$\binom{0.070}{0.082}$	$\binom{0.059}{0.090}$	$\binom{0.053}{0.379}$	$\binom{0.049}{0.999}$
	25	$\binom{0.053}{0.189}$	$\binom{0.051}{0.346}$	$\binom{0.051}{0.868}$	$\binom{0.050}{1.000}$	$\binom{0.061}{0.095}$	$\binom{0.055}{0.137}$	$\binom{0.050}{0.569}$	$\binom{0.051}{0.996}$	$\binom{0.071}{0.089}$	$\binom{0.058}{0.110}$	$\binom{0.054}{0.403}$	$\binom{0.052}{0.975}$
	30	$\binom{0.051}{0.170}$	$\binom{0.049}{0.311}$	$\binom{0.047}{0.929}$	$\binom{0.051}{1.000}$	$\binom{0.064}{0.094}$	$\binom{0.057}{0.125}$	$\binom{0.050}{0.663}$	$\binom{0.048}{0.995}$	$\binom{0.070}{0.084}$	$\binom{0.063}{0.100}$	$\binom{0.054}{0.475}$	$\binom{0.049}{0.964}$

[a] %I(1) denotes the proportion of series which are I(1). For the notation $\binom{a}{b}$ we have that a gives the probability that an I(1) series will be classified as I(0), whereas b gives the probability that an I(0) series will be classified as I(0).

Table 4: A comparison of SPSM and DF for large N[a]

	SPSM		
N	DF 1	DF 2	DF 3
200	$\binom{0.083}{0.703}$	$\binom{0.103}{0.344}$	$\binom{0.068}{0.161}$
400	$\binom{0.089}{0.737}$	$\binom{0.116}{0.380}$	$\binom{0.083}{0.191}$
	DF		
	DF 1	DF 2	DF 3
200	$\binom{0.050}{0.546}$	$\binom{0.056}{0.213}$	$\binom{0.061}{0.147}$
400	$\binom{0.050}{0.552}$	$\binom{0.056}{0.214}$	$\binom{0.061}{0.149}$

[a] %I(1) denotes the proportion of series which are I(1). For the notation $\binom{a}{b}$ we have that a gives the probability that an I(1) series will be classified as I(0), whereas b gives the probability that an I(0) series will be classified as I(0).

$Var(\hat{\beta}_j) - 2Cov(\hat{\beta}_j, \tilde{\beta}) + Var(\tilde{\beta})$. However, as argued by Kapetanios (2003), the covariance term is asymptotically negligible for the test as $N \to \infty$. An appropriate estimate of $Var(\hat{\beta}_j - \tilde{\beta})$ may then be based on a consistent estimate of the variance of $\hat{\beta}_j$. Then, it follows from the assumption of asymptotic normality of the estimators, made in Kapetanios (2003), that as $T \to \infty$, $S_{T,j} \xrightarrow{d} \chi^2_k$, for each unit j.

The poolability test is based on the $S_{T,j}$ statistics. In particular Kapetanios (2003) suggests that $S^s_T = sup_j S_{T,j}$ be used as a test statistic for the test of the null hypothesis H_0. As before, let $\mathbf{Y_i} = (\mathbf{y}_{j_1}, \ldots, \mathbf{y}_{j_M})$, $\mathbf{i} = \{j_1, \ldots, j_M\}$ and $\mathbf{i}^j = \{j\}$, $\{1, \ldots, N\} \equiv \mathbf{i}^{1,N}$ and $\mathbf{i}^{-j}$ such that $\mathbf{i}^{-j} \cup \mathbf{i}^j = \mathbf{i}$. We now define the object we wish to estimate. To simplify the analysis we assume that there exists one cluster of series with equal $\beta_j = \beta$. If all series have different β_j then without loss of generality we assume that $\beta_1 \equiv \beta$. The more general case, of multiple clusters, is straightforward to deal with and is discussed in Kapetanios (2003). For every series $y_{j,t}$ (and associated set of predetermined variables $x_{j,t}$) define the binary object $\mathcal{I}_j$ which takes the value 0 if $\beta_j = \beta$ and 1 if $\beta_j \neq \beta$. Then, $\mathcal{I}_\mathbf{i} = (\mathcal{I}_{j_1}, \ldots, \mathcal{I}_{j_M})'$. We wish to estimate $\mathcal{I}_{\mathbf{i}^{1,N}}$. We denote the estimate by $\hat{\mathcal{I}}_{\mathbf{i}^{1,N}}$.

To do so we consider the following procedure.

1. Set $j = 1$ and $\mathbf{i}_j = \{1, \ldots, N\}$.

2. Calculate the S^s_T-statistic for the set of series $\mathbf{Y}_{\mathbf{i}_j}$. If the test does not reject the null hypothesis $\beta_i = 0$, $i \in \mathbf{i}_j$, stop and set $\hat{\mathcal{I}}_{\mathbf{i}_j} = (0, \ldots, 0)'$. If the test rejects go to step (3).

3. Set $\hat{\mathcal{I}}_{\mathbf{i}^l} = 1$ and $\mathbf{i}_{j+1} = \mathbf{i}_j^{-l}$, where l is the index of the series associated with the maximum $S_{T,s}$ over s. Set $j = j + 1$. Go to step (2).

In other words, we estimate a set of binary objects that indicate whether a series is poolable or not. We do this by carrying out a sequence of poolability tests on a reducing dataset where the reduction is carried out by dropping series for which there is evidence of nonpoolability. A large individual $S_{T,j}$-statistic is used as such evidence. Note that we do not need to use the poolability test based on S^s_T. The method can be equally applied using any available poolability test in Step 2 of the algorithm. The asymptotic properties of this method are discussed in detail in Kapetanios (2003). This methodology has been shown to apply to stationary processes. As a result we consider it only for the series that are found to be stationary in Section .

4. Data

We construct the bilateral real exchange rate q against the *i-th* currency at time t as $q_{i,t} = s_{i,t} + p_{j,t} - p_{i.t}$, where $s_{i,t}$ is the corresponding nominal exchange rate (*i-th* currency units per one unit of the *j-th* currency), $p_{j,t}$ the price level in the *j-th* country, and $p_{i,t}$ the price level of the *i-th* country. That is, a rise in $q_{i,t}$ implies a real appreciation of the *j-th* country's currency against the *i-th* country's currency.

The 26 currencies considered are those of Australia, Austria, Belgium, Canada, Cyprus, Denmark, Finland, France, Germany, Greece, Italy, Japan, Korea, Luxembourg, Malta, Mexico, Netherlands, New Zealand, Norway, Portugal, South Africa, Spain, Sweden, Switzerland, the United Kingdom, and the United States. All data are quarterly, spanning from 1957Q1 to 1998Q4 and the bilateral nominal exchange rates against the currencies other than the US dollar are cross-rates computed using the US dollar rates. More specifically we consider two different panels each one of which consists of up to 25 country pairs and corresponds to a different numeraire currency (US dollar, DM). We use the average quarterly nominal exchange rates and the price levels are consumer price indices (not seasonally adjusted[7]). All variables are in logs. All data are from the International Monetary Fund's *International Financial Statistics* in CD-ROM.[8]

We confine our focus on the $US and the DM only as numeraires since the Yen's behavior has been considered as exceptional in the post WWII era. The yen experienced trend like appreciation and it is likely that tests that allow for structural break or nonlinearities may be better equipped to capture the corresponding real exchange rate dynamics.[9] The length of the data was dictated by the availability of the IFS/IMF data. We stop at 1998 when a number of the countries in our sample joined the European Monetary Union (EMU) and shared a common currency. Considering more recent data would result either in using a much smaller number of cross-sectional units or using only the relative price levels for the EMU member countries.[10] To better convey the main point of our paper we produce results that are easily comparable with well-studied data sets and focus on the period before the introduction of the euro.

We therefore, focus out analysis on two sample periods: one combines the Bretton Woods era from 1957 to 1974 and the post Bretton Woods era from 1974 till 1998, while the other considers only the post Bretton Woods era. It is reasonable to claim that the first sample period may incorporate a structural break in the dataset. We consider, however, the problem of cross-sectional dependence to be more important. To the best of our knowledge, there is only one paper (Basher and Carrion-i-Silvestre (2007)) proposing panel unit root tests that combine robustness to structural breaks and cross-sectional dependence. The treatment of cross-dependence used in this test, however, is based on Bai and Ng (2005) which has some unsatisfactory features (as our

[7]Although the data are not seasonally adjusted, visual inspection suggests no apparent seasonal patterns. Note that the usual practice of adding dummies to a regression to account for seasonality is problematic for the panel unit root tests we consider in Section 5. For the IPS and the test by Pesaran (2003), adding extra deterministic terms such as seasonal dummies changes the asymptotic distributions and requires the derivation of new critical values. Further, the properties of the tests under these conditions have not been explored in the literature. For the test by Chang (2002) there are further issues arising out of the fact that deterministic terms in the context of this test are not simply added to the regression but are dealt with prior to running the test regression. For more details see Chang (2002).

[8]It should be noted that the use of CPI indexes is not an exact measure of price levels and this can make the discussion of PPP and half-lives subtler. This is, however, the approach followed in the vast majority of the literature. Another approach focuses on data from commodities which are price level data. We follow the first approach not only because it is the conventional one and allows comparability of our results with the main relevant contributions in the literature but also for practical purposes pertaining to the calculation of appropriate price levels for such a large sample. Various other approaches have been proposed recently such as focusing on an inflation measure extracted from financial markets (Chowdhry, Doll, and Xia (2005)) or on data transformations to achieve consistency between real exchange rates and real exchange rate indexes (Wagner (2007)).

[9]For a recent analysis of the yen real exchange that uncovers evidence of stationarity taking into account nonlinear behavior see Chortareas and Kapetanios (2004). Sarno and Valente (2006) is another example of work that finds support for PPP using nonlinear models, both for the yen and more generally.

[10]One difference between the Bretton Woods system and the euro area is that although the former had fixed exchange rates some degree of nominal exchange rate flexibility was available through realignments and small margins of fluctuations while the countries that form the euro area constitute a monetary union with irrevocably fixed rates. A second difference is that the distinction between pre and post Bretton Woods periods is covering all countries in the sample while the distinction between the pre and post euro periods covers a subset of countries.

discussion in the next section discusses in detail) and therefore we do not consider it.

5. Empirical Results

We present results from two of the standard univariate unit-root test specifications, i.e., a model with constant only, and a model with constant and trend. Then, we present results from the SPSM methodology, outlined in section 3, as applied to the panel data in order to obtain the country specific stationarity results.

Our benchmark test is the IPS test, and its univariate counterpart, the ADF test, as discussed in Section 3. One potential weakness of standard panel tests of unit roots is that linkages among the units may exist. For example, O'Connell (1998) suggests that non-zero covariances of the errors across the units imply short-run linkages among the units. If such dependency exists among cross sections then many panel unit tests including the IPS test are invalid. In order to avoid this problem we also use two tests that are designed to be correctly sized even in the presence of cross-sectional dependency since the empirical literature on PPP strongly suggests the presence of cross-sectional dependence (see, e.g., Basher and Carrion-i-Silvestre (2007), Lyhagen (2000) and Wagner (2007)).

First, we consider the panel unit root test by Chang (2002) which makes use of nonlinear instrumental variables in a Dickey-Fuller unit root regression context. These instrumental variables make the test robust to cross-sectional dependence. The nonlinear function used to construct the instrumental variables is chosen to be a Hermite polynomial following Chang (2002). We also consider a further modification of the test, as discussed in Chang and Song (2002) where the instrumental variables are uncorrelated even in the presence of cointegration in the series. In the case of serial correlation, the test regression is augmented with lags of $\Delta y_{i,t}$, in a similar fashion to the ADF test.

The second test we consider has been proposed by Pesaran (2003). This test corrects for cross-sectional dependence by augmenting the standard Dickey-Fuller regression by the cross sectional averages of $y_{i,t-1}$ and $\Delta y_{i,t}$. These cross-sectional averages proxy a single unobserved factor. The resulting Dickey-Fuller test statistics are then averaged in a similar fashion to the averaging carried out when applying the IPS panel unit root test. In the case of serial correlation, Pesaran (2003) suggests further augmenting the individual Dickey-Fuller regression with both the cross sectional average of $\Delta y_{i,t-j}$, $j = 1, \ldots, p$ and $\Delta y_{i,t-j}$ itself. Pesaran (2003) suggests using either an unadjusted or a truncated version of the Dickey-Fuller statistics. We choose to use the truncated version[11]. Both these tests have similar structure to the IPS test and so the application of the methodology in Section 3 is extremely straightforward. For both these tests we propose to choose the order of the augmentation in the serially correlated case using sequential testing following Ng and Perron (1995). We set the maximum allowable lag order to 6. Note that, for all the panel unit root tests we consider, we allow, in the case where the lag structure is data-determined, different lag orders for different cross-sectional units.[12]

We report both results from the panel unit root tests that are robust to cross-sectional depen-

[11] It is useful to note the assumptions underlying the error terms $\epsilon_{j,t}$ for the tests by Chang (2002) and Pesaran (2003). Both these tests, in their simple version, basically assume that the errors are i.i.d processes, as in Assumption 1. Chang (2002) allows for the errors to be cross-sectionally correlated, whereas Pesaran (2003) adds a factor structure to the errors. In the case where serial correlation needs to be modelled both tests assume that the errors

Table 5: Stationary Series: \$US, Full sample[a][b].

4 Lags											
IPS				Chang				Pesaran			
Univ.		Panel		Univ.		Panel		Univ.		Panel	
NZ	Mal	Bg	Cyp	Bg	Cyp	Bg	Aut	Sp	Fr	.	Aus
SAf	SAf	Gr	Fr	Cyp	Fr	Cyp	Bg	UK	Jap		Aut
		Ita	Ita	Fin	Ita	Fin	Cyp		Sp		Can
		NZ	Jap	Fr	Jap	Fr	Fin				Fin
		SAf	Mal	Gr	Mal	Gr	Fr				Fr
			NZ	Ita	NZ	Ita	Ger				Ger
			SAf	Lux	SAf	Lux	Gr				Ita
			Sp	NZ	UK	NZ	Ita				Jap
			UK	SAf		SAf	Jap				Lux
				UK		UK	Lux				Mal
							Mal				NZ
							NZ				Por
							SAf				SAf
							Sp				Sp
							Swe				Swe
							Swi				Swi
							UK				UK
Automatic Lag Selection											
IPS				Chang				Pesaran			
Univ.		Panel		Univ.		Panel		Univ.		Panel	
NZ	Mal	Aus	Cyp	Aus	Fr	Aus	Aut	NZ	Fr	NZ	Aus
	NZ	Bg	Gr	Bg	Ita	Bg	Fin	Sp	Sp		Aut
	UK	Gr	Ita	Fr	Mal	Fr	Fr				Bg
		Ita	Jap	Ita	NZ	Ger	Ita				Can
		NZ	Mal	NZ	Sp	Gr	Jap				Cyp
			NZ	UK	UK	Ita	Lux				Fin
			SAf			NZ	Mal				Fr
			UK			UK	NZ				Ger
							SAf				Gr
							Sp				Ita
							Swi				Jap
							UK				Lux
											Mal
											Neth
											NZ
											SAf
											Sp
											Swe
											Swi
											UK

[a] Aus: Australia, Aut: Austria, Bg: Belgium, Can: Canada, Cyp: Cyprus, Den: Denmark, Fin: Finland, Fr: France, Ger: Germany, Gr: Greece, Ita: Italy, Jap: Japan, Kor: Korea, Lux: Luxembourg, Mal: Malta, Mex: Mexico, Neth: Netherlands, NZ: New Zealand, Nor: Norway, Por: Portugal, SAf: South Africa, Sp: Spain, Swe: Sweden, Swi: Switzerland, UK: United Kingdom, US: United States

[b] The columns denoted 'Univ.' in the Table refer to the two versions of the Dickey-Fuller test depending on the deterministic terms included (the first column is for constant only and the second for constant and trend.). The columns denoted 'Panel' refer to the sequential testing procedure discussed in section . Again, the first column is for constant only and the second for constant and trend. A dot in any given column denotes that no series was found to be stationary in a given dataset.

Table 6: Stationary Series: $US, Post- Bretton Woods[a]

4 Lags											
IPS				Chang				Pesaran			
Univ.		Panel		Univ.		Panel		Univ.		Panel	
Fin	.	Aut	Mal	Aut	.	Aut	Neth	.	Aut	Aut	.
Neth		Bg	Neth	Bg		Bg	SAf		Swi	Cyp	
		Cyp	SAf	Den		Cyp				Den	
		Den		Fin		Den				Fin	
		Fin		Fr		Fin				Fr	
		Fr		Ger		Fr				Ger	
		Ger		Gr		Ger				Mex	
		Ita		Ita		Gr				Neth	
		Mex		Lux		Ita				NZ	
		Mal		Mal		Kor				Nor	
		Neth		Neth		Lux				Sp	
		NZ		NZ		Mal				Swe	
		SAf		Swi		Neth				Swi	
		Swi		UK		NZ					
		UK				SAf					
						Sp					
						Swi					
						UK					
Automatic Lag Selection											
IPS				Chang				Pesaran			
Univ.		Panel		Univ.		Panel		Univ.		Panel	
Neth	NZ	Bg	NZ	Bg	.	Aus	.	NZ	Aus	Cyp	.
NZ		Cyp		Fr		Aut		Swe	NZ	Fin	
		Fin		Ger		Bg			Swi	Fr	
		Ger		Ita		Cyp				Ger	
		Ita		Kor		Fr				Kor	
		Kor		Neth		Ger				Mex	
		Mex		NZ		Ita				NZ	
		Mal		Swi		Kor				Nor	
		Neth		UK		Lux				Swe	
		NZ				Neth				Swi	
		Swi				NZ					
		UK				Nor					
						Sp					
						Swi					
						UK					

[a]See footnotes in Table 5

Table 7: Stationary Series: DM, Full sample [a]

4 Lags											
IPS				Chang				Pesaran			
Univ.		Panel		Univ.		Panel		Univ.		Panel	
Por	Aus	.	Aus	Fin	Aus	Nor	Aus	UK	SAf	.	SAf
	Por		Por	Nor	Can	Por	Can				
	SAf			Por	Jap	UK	Jap				
				UK	Por		Nor				
					SAf		Por				
							SAf				
							UK				
							US				
Automatic Lag Selection											
IPS				Chang				Pesaran			
Univ.		Panel		Univ.		Panel		Univ.		Panel	
Por	Aus	.	Aus	Fin	Aus	Nor	Aus	NZ	Fr	.	Fr
	Por		Por	Nor	Can	Por	Can	UK	NZ		NZ
	SAf			Por	Jap	UK	Jap		Sp		Sp
				UK	Por		Nor				
					SAf		Por				
							SAf				
							UK				
							US				

[a]See footnotes in Table 5

dence and the IPS test. Although, the IPS test is not valid under cross-sectional dependence it provides a natural benchmark with which to compare the results of the other tests, both within our empirical work and the existing literature more generally.

In presenting our results we arrange each table block so that the first two columns (labelled 'Univ.' for univariate) correspond to the two different specifications of the corresponding/benchmark univariate unit-root test, i.e., a model with constant only, and a model with constant and trend. The next two columns (labelled 'Panel') provide the results from the SPSM methodology outlined in section 3 as applied to the panel data, for each of the panel unit root tests (IPS, Chang (2002) and Pesaran (2003)), in order to obtain the country specific stationarity results. Again, the two different specifications of a model with constant only, and a model with constant and trend, are used. We report results both for a typical four-lag structure since the data are quarterly, and for a data-dependent lag structure, using the sequential testing approach of Ng and Perron (1995), as discussed above. We feel that reporting results for two different lag structures provides some robustness to our analysis, especially noting that for the panel unit root tests of Chang (2002) and Pesaran (2003) there is no evidence on the performance of any automatic lag selection procedure in small samples. Results are presented in Tables 5-9.

Table 5 provides the results from applying our procedure to the three panel unit root tests we consider, on a panel of 22 bilateral real exchange rates against the US dollar from 1957Q

are finite autoregressions similarly to the treatment used for the IPS test.

[12]An alternative panel unit root test that is robust to cross-sectional dependence, is that proposed by Bai and Ng (2005). This test assumes that cross-sectional dependence arises out of the presence of unobserved factors similarly to Pesaran (2003). Unlike Pesaran (2003), Bai and Ng (2005) estimate these factors and remove their effects from the panel dataset. However, we choose not to consider it because theoretical work by Westerlund (2007) and Monte Carlo evidence from Kapetanios (2007), who also considered a number of extensions, suggest that the performance of this test is not satisfactory. This also motivates out choice not to consider the panel unit root tests of Basher and Carrion-i-Silvestre (2007) which are robust to structural breaks and cross-sectional dependence, since cross-sectional dependence is dealt with either by demeaning or via the methods discussed in Bai and Ng (2005).

Table 8: Stationary Series: DM, Post-Bretton Woods [a]

4 Lags											
IPS				Chang				Pesaran			
Univ.		Panel		Univ.		Panel		Univ.		Panel	
Aut	Aut	Aus	Aus	Aus	Kor	Aus	Aus	Aut	Aut	Aus	Aus
Cyp	Bg	Aut	Aut	Bg	Por	Bg	Fin	Cyp	Bg	Aut	Aut
Den	Cyp	Bg	Bg	Can		Can	Kor	Fr	Can	Bg	Bg
Fin	Fr	Cyp	Cyp	Fin		Cyp	NZ	Swe	Fr	Can	Can
Fr	Por	Den	Den	Fr		Fin	Por		Swi	Cyp	Den
Swe	Swi	Fin	Fr	Kor		Fr	Swi			Den	Fr
Swi		Fr	Por	NZ		Gr				Fin	Lux
		Mex	Swi	Nor		Kor				Fr	NZ
		NZ		Por		NZ				Ita	Swi
		Nor		UK		Nor				Jap	
		Por		US		Por				Kor	
		Swe				SAf				Lux	
		Swi				Sp				Mex	
		UK				UK				Neth	
		US				US				NZ	
										Nor	
										Por	
										SAf	
										Sp	
										Swe	
										Swi	
										UK	
										US	
Automatic Lag Selection											
IPS				Chang				Pesaran			
Univ.		Panel		Univ.		Panel		Univ.		Panel	
Fr	Fr	Aus	Fr	Fr	NZ	Fin	.	Fr	Can	Aut	Swi
NZ	Por	Bg	Por	NZ		Fr		NZ	NZ	Bg	
Swe	Swi	Cyp	Swi	Sp		Gr		Swe	Swi	Cyp	
Swi		Fin		US		Kor				Fin	
		Fr				NZ				Fr	
		Kor				Nor				Jap	
		Mex				Sp				Kor	
		NZ				US				NZ	
		Por								Nor	
		Sp								Swe	
		Swe								Swi	
		Swi								US	
		US									

[a]See footnotes in Table 5

to 1998Q4. Starting with the IPS test and a lag order of 4 lags, we see that the univariate test specifications provide up to two rejections of the null hypothesis out of the 22 series in our sample. The panel unit root test suggests stationarity of the panel and applying the new methodology we show that up to nine out of the 22 series are stationary. Those are the real exchange rates of four large European countries (France, Italy, Spain, and the UK), two small European economies (Cyprus and Malta), and those of New Zealand, South Africa and Japan. Next, we look at the results of the Chang test. The univariate version of the test provides more evidence of stationarity than the ADF tests with up to 10 series found to be stationary. The panel test combined with the SPSM methodology, though, finds more evidence for stationarity with 17 series found stationary in one case. In addition to the 4 major European countries found stationary using the IPS test, the Chang test suggests that the German real exchange rate is stationary. Moving on to the Pesaran test we see little evidence of stationarity using the univariate tests. On the contrary, the panel test combined with the SPSM methodology finds 17 countries to be stationary. Finally, we consider the tests when the lag order is chosen automatically. Results are comparable to the 4 lag case.

Table 6 considers a panel of real exchange rates against the US dollar in the Post Bretton Woods era. The panel now includes 25 countries (Denmark, Korea and Mexico have been added in the sample). We again start with the IPS test and a lag augmentation of 4 lags. The test finds little evidence of stationarity in the univariate case. When the panel methodology is applied we see that up to 15 countries are found to be stationary. The countries whose bilateral real exchange rate with the US dollar is stationary include the large European economies (France, Germany, Italy, UK). The univariate Chang test again finds evidence of stationarity, but the panel SPSM methodology finds even more evidence with up to 18 countries found stationary. The Pesaran test finds stationarity evidence, using the panel methodology, as well but only for 13 countries. Results for the case where the lag order is chosen automatically are similar to those described above.

We repeat the analysis for the bilateral real exchange rates using the German Mark as the numeraire currency and we provide the results in Tables 7 and 8. Table 7 reports results for the full sample period. The evidence in support of the PPP hypothesis is scant, in this case regardless of the tests and specification used. Most of the evidence in favour of stationarity is produced by the Chang test. The univariate tests provide evidence of stationarity for up to five series whereas the panel methodology provides evidence of stationarity for eight series including the UK and the US which are in fact found nonstationary using the univariate tests.

Moving on to the results for the post-Bretton Woods era with the DM as the numeraire we see a completely different picture. The ADF tests with a lag augmentation of four lags find up to seven series stationary. The panel methodology, based on the IPS test, however, finds up to 15 series stationary. These include those found stationary using the univariate tests but additionally the UK and the US. The Chang and Pesaran tests find a similarly large extent of stationarity evidence. Results for a data dependent lag order provides slightly less evidence of stationarity.

The results are consistent across the various tests in the sense that all the real exchange rates which are found stationary with the univariate tests are also found stationary with the panel tests. Moreover, the results for specific real exchange rates appear consistent across the various tests. That is, some real exchange rates emerge as stationary and some as nonstationary regardless of the test used. For most of the real exchange rates we consider, however, the use of the panel methodology is decisive in uncovering evidence for PPP.

Consider the real exchange rates against the dollar for example. Some of them appear consistently stationary regardless of whether one uses univariate or multivariate tests. Such are the real exchange rates of small open economies, such as Denmark, Finland, Malta, the Netherlands and New Zealand. Clearly for those real exchange rates it does not make a great difference whether one considers their stationarity using univariate or multivariate methods. Another set of exchange rates appears almost invariably as nonstationary and includes those of Canada, Korea, Norway, and Portugal. The typical panel unit root tests which show that the null for the panel can be rejected are therefore misleading. Another set of real exchange rates includes those where the choice to use univariate or multivariate approaches affects the results. This is a "gray area" where the usefulness of distinguishing between stationary and non stationary series in a given panel becomes critical. We find that the ability to reject the nonstationarity null is enhanced in eleven countries. For some of them (Cyprus, France, Germany and Italy), the evidence suggests that the real exchange rate is on balance stationary with the panel tests further strengthening the case. In other cases, however, (for example the UK) the use of multivariate methodologies becomes crucial in obtaining evidence of stationarity on balance. The use of panel tests allows for a more dramatic overturn of the results in the real exchange rates of Austria and Spain, where the multivariate tests indicate stationarity while the majority of the univariate would lead someone to accept the null. The panel tests are critical in obtaining the stationarity results.

Banerjee, Marcellino, and Osbat (2003) suggest that since the panel unit root tests assume away the presence of cross-section cointegrating relationships, if this assumption is violated the tests become oversized. Such relationships/linkages can emerge because of common factors or omitted variables. The test by Chang (2002) is itself not valid under cointegration. To correct for this possibility we employ a test introduced by Chang and Song (2002) that takes into account the possibility that cointegrating relationships between the cross-sectional units may exist. Again, Hermite polynomials are used (in this case different ones for every cross-sectional unit). We provide the results of this test in Table 9. The results are not identical but in general the results point to the same direction as those of Tables 5-8. That is, using multivariate tests produces significantly more evidence of real exchange rate stationarity. One problem with the Chang and Song (2002) tests with cointegration, however, is that their results may be sensitive to the ordering of the series. This is because different functional forms are used for different cross-sectional units. Indeed if we run the same test where the series are introduced in reverse order (Results appear under the headings 'Cointegration Robust (reverse)' in Table 9) the results are slightly affected. Thus, we use those results only as indicative and not as definitive.[13]

6. How Bad is the PPP Puzzle?

PPP is not inconsistent with temporary deviations from equilibrium. Theory suggests that the predominant causes for such departures from PPP should be sought in monetary and financial shocks when price stickiness exists. The observed high degree of short-term volatility in exchange rates would be also be consistent with such nominal stickiness. Consequently the real exchange rate persistence that one expects to observe should more or less match the period of price (and/or wage) adjustment to shocks. In reality, however, the degree of persistence in real exchange rates exceeds the magnitudes that would be consistent with adjustment to nominal shocks and seems to be more easily reconcilable with real shocks (e.g., shocks to productivity and tastes). This, however, is not consistent with the high degree of short-term exchange rate volatility. This inconsistency has been termed the "PPP puzzle" by Rogoff (1996).

The measure of persistence that dominates the literature is the half-life of PPP deviations which indicates how long it takes for the impact of a unit shock to dissipate by half. Half-lives are typically estimated from autoregressive processes. The standard formula for the half-life is given by $H = -\log 2/(\log \beta)$, where β is the speed of adjustment parameter (autoregressive coefficient) in an autoregressive process of order one. There are two issues with this definition. Firstly, assuming an $AR(1)$, instead of an $AR(p)$ with a data-dependent lag length, p, is restrictive. So, we consider both the standard $AR(1)$ specification and an $AR(p)$ specification where we choose p

[13] An anonymous referee has suggested that we use a modification suggested in a revised version of Chang and Song (2002) that proposes a rule for the ordering problem. The rule is to order the cross-sectional units by the ascending order of magnitudes of the long-run variances of their first differences. We have done that and the results do not change significantly in terms of the extent of the evidence for stationarity. Due to space constraints we do not report these results which are available upon request.

Table 9: Stationary Series according to the Cointegration Robust Chang Test[a]

4 Lags								Automatic Lag Selection							
Cointegration Robust				Cointegration Robust (reverse)				Cointegration Robust				Cointegration Robust (reverse)			
Univ.		Panel		Univ.		Panel		Univ.		Panel		Univ.		Panel	
Dataset 1															
Bg	Cyp	Bg	Aut	UK	UK	NZ	UK	Bg	Fr	Bg	Aut	NZ	SAf	NZ	Sp
Cyp	Fr	Fin	Bg	NZ	SAf	Gr	Swi	Fr	Ita	Swi	Fr	Bg	Mal		SAf
Fin	Ita	Gr	Cyp	Lux	Mal	Bg	Swe	NZ	Lux		Ita		Lux		Mal
Fr	Lux	Swi	Fin	Gr	Lux		Sp	Swi	Mal		Lux		Ita		Lux
Gr	Mal		Fr	Fin	Ita		SAf				Mal		Fr		Ita
Lux			Ger	Bg	Fr		NZ								Fr
NZ			Gr		Cyp		Mal								
Swi			Ita				Lux								
			Jap				Jap								
			Lux				Ita								
			Mal				Ger								
			Sp				Fr								
							Cyp								
Dataset 2															
Aut	.	Aut	.	UK	.	UK	.	Bg	.	Aut	.	Swi	.	Swi	.
Bg		Bg		Swi		Swi		Fr		Bg		Neth		NZ	
Den		Cyp		Neth		Sp		Lux		Fr		Lux		Neth	
Fin		Den		Mal		Neth		Neth		Ita		Fr		Lux	
Fr		Fin		Lux		Mal				Lux		Cyp		Fr	
Ger		Fr		Ger		Lux				Mex		Bg		Cyp	
Lux		Ger		Fr		Ger				Neth				Bg	
Mal		Gr		Fin		Fr									
Neth		Lux		Cyp		Fin									
		Mex		Bg		Cyp									
		Mal				Bg									
		Neth													
Dataset 3															
Fin	Aus	.	Aus	UK	SAf	.	US	.	Aus	.	Can	UK	SAf	.	SAf
Nor	Can		Can	Por	Por		UK		Can		Por		Por		Por
Por	Jap		Jap	Nor	Aus		SAf		Jap						
	Por		Lux	Fin			Por		Por						
			Nor				Nor								
			Por				Jap								
							Aus								
Dataset 4															
Aus	Por	Aus	Por	US	Por	US	Por	Bg	.	Bg	.	Nor	.	Nor	.
Bg		Bg		UK		UK	Kor	Fin		Cyp		Fr		NZ	
Can		Can		Por		Sp		Fr		Fin		Fin		Gr	
Fin		Cyp		Nor		Por		Nor		Fr		Bg		Fr	
Fr		Fin		NZ		Nor		Swe		Gr				Fin	
NZ		Fr		Fr		NZ				Nor				Bg	
Nor		Gr		Fin		Gr				Sp					
Por		NZ		Bg		Fr				Swe					
Swe		Nor				Fin									
		Por				Cyp									
		Sp				Bg									
		Swe													
		UK													

[a]Dataset 1: \$US, Full sample; Dataset 2: \$US, Post Bretton Woods; Dataset 3: DM, Full sample; Dataset 4: DM, Post Bretton Woods.

using the Akaike information criterion using a maximum lag order of 6. Since, the half life measure does not have a closed form solution for $AR(p)$, $p > 1$, models, we calculate it numerically. The second problem relates to OLS estimation of AR models. Simple OLS, is downward biased, in small samples, implying a downward bias to the half-life estimate. Median unbiased estimators have been suggested by Andrews (1991) and Andrews and Chen (1994). We, therefore, report half life measures for both $AR(1)$ and $AR(p)$ models based on the estimation method of Andrews and Chen (1994). We also consider panel estimation of the AR model below. To the best of our knowledge there is no fully articulated and tested extension of estimators such as that suggested in Andrews and Chen (1994), to a panel dataset context. Further, the bias, we are faced with, is a small sample phenomenon. Given that our panel estimation deals with datasets which for all but one case have more than 300 observations, and in the overwhelming majority of cases over

Table 10: Poolable Series[a]

IPS Test				Chang Test				Pesaran Test			
D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12
Cyp	Bg	Aus	Aus	Aut	Aus	Aus	Fin	Aus	Cyp	Fr	Aut
Gr	Cyp	Por	Bg	Fin	Aut	Can	Fr	Aut	Fin	NZ	Bg
Ita	Fin		Cyp	Fr	Bg	Jap	Gr	Bg	Fr	Sp	Cyp
Jap	Ger		Fin	Ita	Cyp	Nor	Kor	Can	Ger		Fin
Mal	Ita		Fr	Jap	Fr	Por	NZ	Cyp	Kor		Fr
NZ	Kor		Kor	Lux	Ger	SAf	Nor	Fin	Mex		Jap
SAf	Mex		Mex	Mal	Ita	UK	Sp	Fr	NZ		Kor
UK	Mal		NZ	NZ	Kor	US	US	Ger	Nor		NZ
	Neth		Por	SAf	Lux			Gr	Swe		Nor
	NZ		Sp	Sp	Neth			Ita	Swi		Swi
	Swi		Swi	Swi	NZ			Jap			US
	UK		US	UK	Nor			Lux			
					Sp			Mal			
					Swi			Neth			
					UK			NZ			
								SAf			
								Sp			
								Swe			
								Swi			
								UK			

[a] D1: \$US, Full sample, IPS test; D2: \$US, Post Bretton Woods, IPS test; D3: DM, Full sample, IPS test; D4: DM, Post Bretton Woods, IPS test; D5: \$US, Full sample, Chang test; D6: \$US, Post Bretton Woods, Chang test; D7: DM, Full sample, Chang test; D8: DM, Post Bretton Woods, Chang test; D9: \$US, Full sample, Pesaran test; D10: \$US, Post Bretton Woods, Pesaran test; D11: DM, Full sample, Pesaran test; D12: DM, Post Bretton Woods, Pesaran test;

1000 observations, OLS estimation emerges as the appropriate procedure for the panel case[14]. Our discussion focuses on OLS estimation to allow comparability of individual and panel half life estimates.

Studies of PPP typically find a high degree of persistence in real exchange rates with half-lives usually ranging between three to five years (see Rogoff (1996)).[15] Frankel and Rose (1996), for example, in a study covering 150 countries find a half-life of four years. The multivariate approach of Abuaf and Jorion (1991) indicates half-lives of 3.3 years. Those results typically refer to the average half-life estimates based on autoregressive models of *all* real exchange rates. That is, both the stationary and nonstationary ones are considered. Including the half-lives of the nonstationary real exchange rates, however, may be misleading since one cannot expect their persistence to die out. The nonstationary real exchange rates do not revert to their PPP values and therefore the estimated half-lives for those process are or little relevance. So, it is more meaningful to focus only on the half-lives of the stationary real exchange rates when of assessing the speed of adjustment to PPP.

Existing PPP studies that use multivariate methodologies are not able to identify the individual real exchange rates that make it possible to reject the null. Therefore it has not been feasible to obtain half-lives estimates of the stationary series. Our analysis, however, allows us to do so, and as we show below the results are striking. Results on individual half-life estimates are reported in Table 11 whereas panel estimates and average half-lives are reported in Table 12.

[14] Experimentation suggests that for the sample sizes we have for the panel case, OLS is not biased.

[15] Studies exist, nevertheless that either exceed or fall short of those bounds. For example, Lothian and Taylor (1996) find that the half-life for the \$/£ real exchange rate is 5.9 years and Papell (1997) finds that the half-lives of the real exchange rates in Europe can be as low as 1.9 years. Also Cumby (1996) puts this number close to 1 but the methodology he uses is different, focusing on Big Mac indices.

We compare the average half-lives for all real exchange rates within a given panel with the average half-lives for the stationary-only real exchange rates. We consider the sets of stationary series that emerge from applying our methodology to the IPS, Chang (2002) and Pesaran (2003) tests. The results in upper section of Table 12 indicate that when only stationary series are considered the half-lives of adjustment to PPP become shorter by up to almost one year for the $US real exchange rate and by up to 1.5 years for the DM real exchange rate. These results are for half-lives calculated using an $AR(p)$ model. Results obtained using the $AR(1)$ model suggest even greater differences. The gains in the speed of mean reversion for both the $US and the DM real exchange rates are more pronounced when the full period is considered. They are also more pronounced when the IPS test is used as compared to the Chang and Pesaran tests, except in the case of the post-Bretton Woods DM real exchange rates.

A similar pattern emerges when we consider the half-life of the series estimated as a panel. To do this we fit an $AR(p)$ model to the whole panel, where p is determined in a data dependent way using the Akaike information criterion. We assume that the whole AR coefficient structure is homogeneous across processes[16]. First, we estimate the half-life when all series are included in the panel. Then, we estimate the model when the panel contains only stationary series which have also been found to be poolable following the methodology of section 3.2. The results of the poolability methodology are presented in Table 10. Note that for the purposes of constructing the stationary dataset we use a version of the test that has both a constant and a trend in the case of the full sample but only a constant for the Post Bretton Woods sample. The half-lives that emerge are relatively close to the average half-lives when the AR processes for real exchange rates are estimated individually. Table 12 shows that when we include all the real exchange rates the resulting half-lives for the four datasets under consideration vary from 3.66 to 3.87 years in the case of an $AR(1)$ model and from 1.75 to 2.29 for an $AR(p)$ model. The $AR(1)$ results are consistent with the surveying of the literature by Mark (2001) which shows an average half-life of 3.7. When we consider the panels that include only the stationary real exchange rates the adjustment process becomes substantially faster. The degree to which this happens depends on the test used with the IPS providing the biggest changes, and the Pesaran test the smallest. The Chang test suggests substantial changes as well. Note that although reductions in half-lives are smaller in absolute terms when an $AR(p)$ model is used, they are quite large in percentage terms. Thus, we find that the persistence of deviations from PPP may have been overstated in previous research and that, on balance, the so-called PPP-puzzle is less pronounced when one focuses only on the stationary real exchange rates.

7. Conclusion

We consider the stationarity of real exchange rates in up to 25 OECD economies in order to assess the case for PPP focusing on the recent float and using the $US and the DM as numeraires. We implement a new set of procedures that allows us to identify the mean-reverting series within a panel. This procedure is applied to panel unit root tests, both conventional as well as recently developed ones that account for cross-sectional dependence. In addition we introduce

[16]For completeness and comparability to the results obtained for individual half life measures we also consider an $AR(1)$ model.

Table 11: Individual Half Lifes (HL)[a]

Dataset 1	HL1	HL2	HL3	HL4	Dataset 2	HL1	HL2	HL3	HL4
Aus	6.154	3.078	14.739	7.370	Aus	5.583	2.792	10.529	5.265
Aut	4.161	2.330	13.999	3.645	Aut	3.112	2.072	4.064	5.207
Bg	4.011	2.315	17.531	3.752	Bg	4.289	2.515	6.436	8.181
Can	74.051	7.788	172.326	8.345	Can	-	2.515	-	-
Cyp	3.352	1.677	8.629	4.315	Cyp	3.557	1.779	5.180	2.591
Fin	2.546	2.068	4.498	2.670	Den	3.386	2.162	5.214	5.934
Fr	3.282	1.930	8.154	2.545	Fin	2.668	1.714	3.455	3.121
Ger	4.649	2.577	7.337	4.322	Fr	3.186	1.914	4.374	3.297
Gr	3.338	1.671	8.583	4.292	Ger	3.244	1.947	4.429	4.311
Ita	2.952	1.717	6.143	2.169	Gr	3.400	1.701	4.598	2.300
Jap	4.167	1.983	13.673	2.613	Ita	3.069	1.810	4.333	3.161
Lux	5.003	2.810	8.944	4.917	Jap	4.238	3.182	7.796	5.331
Mal	2.302	1.426	3.733	1.707	Kor	3.297	1.588	4.565	2.498
Neth	5.631	2.816	10.154	5.077	Lux	4.341	2.482	6.634	8.702
NZ	3.128	1.755	7.064	2.200	Mex	4.191	2.173	6.562	3.091
Nor	6.571	3.286	17.338	8.669	Mal	3.016	1.693	4.139	2.967
Por	4.850	2.426	7.478	3.739	Neth	1.359	0.683	2.204	1.104
SAf	1.733	0.870	2.392	1.197	NZ	3.107	1.791	4.481	3.115
Sp	3.882	2.243	15.028	3.603	Nor	2.327	1.166	2.829	1.416
Swe	4.545	2.513	7.440	4.078	Por	5.148	2.575	9.291	4.646
Swi	4.165	2.278	21.808	3.289	SAf	3.035	1.518	4.122	2.062
UK	2.103	1.053	3.205	1.604	Sp	4.041	2.684	6.403	8.921
					Swe	5.997	2.999	15.034	7.517
					Swi	2.497	1.774	3.352	3.253
					UK	2.282	1.143	2.763	1.384

Dataset 3	HL1	HL2	HL3	HL4	Dataset 4	HL1	HL2	HL3	HL4
Aus	1.651	0.828	2.228	1.116	Aus	2.683	1.343	3.400	1.701
Aut	2.655	1.734	4.893	5.029	Aut	5.678	2.840	10.382	5.191
Bg	7.018	3.509	17.803	8.902	Bg	3.301	1.679	4.723	2.379
Can	3.441	2.009	9.436	2.659	Can	3.892	2.316	5.831	5.456
Cyp	15.676	7.838	173.180	86.590	Cyp	2.535	2.622	-	-
Fin	2.446	2.373	4.168	4.284	Den	2.636	1.319	3.369	1.686
Fr	2.016	1.009	2.986	1.493	Fin	4.398	2.423	7.285	5.397
Gr	5.123	2.562	10.213	5.107	Fr	1.216	-	1.842	-
Ita	6.210	3.485	15.128	9.084	Gr	0.987	0.872	-	-
Jap	1.642	0.990	2.213	1.100	Ita	4.176	2.588	5.656	12.817
Lux	2.823	1.640	5.488	2.163	Jap	4.354	2.178	6.922	3.461
Mal	11.647	5.703	166.480	7.163	Kor	1.520	0.896	2.692	1.088
Neth	6.302	3.152	14.405	7.203	Lux	3.604	2.603	5.169	8.998
NZ	2.399	1.201	4.015	2.008	Mex	3.359	1.988	5.021	4.629
Nor	3.745	2.004	12.356	2.872	Mal	4.801	3.769	8.563	60.230
Por	2.327	1.166	3.788	1.896	Neth	4.350	2.176	6.969	3.485
SAf	1.688	0.847	2.307	1.156	NZ	1.494	0.747	2.605	1.304
Sp	1.815	0.910	2.561	1.282	Nor	3.432	1.927	4.627	3.598
Swe	8.756	5.828	34.099	7.074	Por	4.250	2.182	7.613	5.038
Swi	2.513	1.257	4.389	2.195	SAf	2.837	1.420	3.948	1.974
UK	4.108	2.249	19.345	3.536	Sp	3.023	1.512	4.061	2.031
US	4.658	2.582	8.692	4.331	Swe	9.143	4.572	29.836	14.918
					Swi	3.111	1.557	4.452	2.226
					UK	3.250	1.568	4.450	3.115
					US	3.253	1.949	4.471	4.316

[a]Dataset 1: \$US, Full sample; Dataset 2: \$US, Post Bretton Woods; Dataset 3: DM, Full sample; Dataset 4: DM, Post Bretton Woods. HL1: Half Life estimate based on an $AR(1)$ estimated using OLS; HL2: Half Life estimate based on an $AR(p)$ estimated using OLS; HL3: Half Life estimate based on an $AR(1)$ estimated using the method based on Andrews and Chen (1994); HL4: Half Life estimate based on an $AR(p)$ estimated using the method based on Andrews and Chen (1994). Dashes in the Tables indicate cases where the half life could not be computed, either because the AR coefficient is above 1 (for HL1) or because the numerical solution failed to converge, (for HL2-HL4)

Table 12: Panel and Average Half-Life Measures[a]

Average Half-Life Measures								
	IPS Test		Chang Test		Pesaran Test		All Series	
Dataset	HL1	HL2	HL1	HL2	HL1	HL2	HL1	HL2
1	2.884	1.519	3.285	1.871	7.257	2.344	7.117	2.390
2	3.048	1.717	3.286	1.876	3.407	1.880	3.515	1.994
3	1.989	0.997	2.907	1.584	2.076	1.040	4.575	2.494
4	2.845	1.718	2.415	1.475	3.117	1.880	3.491	2.043
Panel Half-Life Measures								
	IPS Test		Chang Test		Pesaran Test		All Series	
Dataset	HL1	HL2	HL1	HL2	HL1	HL2	HL1	HL2
1	2.796	1.591	3.231	1.836	3.667	2.054	3.797	2.119
2	3.503	1.981	3.163	1.910	4.040	2.397	3.875	2.299
3	1.833	0.919	2.692	1.542	2.088	1.046	3.514	1.757
4	3.043	1.767	2.235	1.120	2.839	1.421	3.662	1.876

[a]The first part of the Table, titled 'Panel Half-Life Measures', presents Half-Life measures as estimated from the panels of stationary poolable series given in Table 10. The second part of the Table, titled 'Average Half-Life Measures', presents averages of individual Half-Life measures, presented in Table 11, using OLS estimation, for the stationary poolable series given in Table 10.

a methodology that formally evaluates the legitimacy of pooling particular sets of real exchange rates together.

Our results show increased evidence of mean-reversion in real exchange rates and therefore strengthen the case in favour of PPP for the real exchange rates of 25 OECD countries. Those results are particularly strong for the $US real exchange rates during the current float -a period for which earlier work has typically failed to find support for PPP. One novelty of our work is that we are able to identify the stationary real exchange rates in the panels while retaining the advantages of panel unit root tests. Moreover, when we perform tests for cross-sectional dependence our results remain robust.

Being able to identify the stationary real exchange rates in our panels, allows us to focus only on the half-lives of the mean-reverting series. We show that when one focuses on these, the half-lives become shorter. The PPP-puzzle becomes less pronounced and the resulting half-lives estimates become more compatible with those predicted by typical sticky-price models. We conclude that the so-called "PPP-puzzle" may have been overstated. Further issues remain open, however, pertaining to further explaining and understanding the stylized facts of the empirical literature such as those related to the sources of the deviations from PPP.

References

Abadir, K. (1995): "The Limiting Distribution of the t Ratio under a Unit Root," *Econometric Theory*, 11, 775–793.

Abuaf, N., and P. Jorion (1991): "Purchasing Power Parity in the Long Run," *Journal of Finance*, 45, 157–174.

Andrews, D. W. K. (1991): "Exactly Median-unbiased Estimation of First Order Autoregressive/Unit Root Models," *Econometrica*, 61, 139–165.

Andrews, D. W. K., and H. Y. Chen (1994): "Approximately Median-Unbiased Estimation of Autoregressive Models," *Journal of Business and Economic Statistics*, 12, 187–204.

Bai, J., and S. Ng (2005): "A PANIC Attack on Unit Roots and Cointegration," *Econometrica*, 72, 1127–1177.

Baltagi, B. H. (2001): *Econometric Analysis of Panel Data*. Wiley.

Baltagi, B. H., J. Hidalgo, and Q. Li (1996): "A Nonparametric Test for Poolability Using Panel Data," *Journal of Econometrics*, 75(1).

Banerjee, A., M. Marcellino, and C. Osbat (2003): "Some Cautions on the Use of Panel Methods for Integrated Series of Macro-Economic Data," *Econometrics Journal*, Forthcoming.

Basher, S. A., and J. L. Carrion-i-Silvestre (2007): "Another Look at the Null of Stationary Real Exchange Rates: Panel Data with Structural Breaks and Cross-section Dependence," *Unpublished Manuscript, University of Barcelona.*

Boucher Breuer, J. (1994): "An Assessment of the Evidence on Purchasing Power Parity," in *Estimating Equilibrium Exchange Rates*, ed. by J. Williamson, pp. 245–277. Institute for International Economics, Washington, DC.

Camba-Mendez, G., and G. Kapetanios (2001): "Testing the Rank of the Hankel Covariance Matrix: A Statistical Approach," *IEEE Transactions on Automatic Control*, 46(2), 331–336.

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

Chang, Y., and W. Song (2002): "Panel Unit Root Tests with Cross-Sectional Dependency and Heterogeneity," *Mimeo, Rice University.*

Choi, C. Y., N. Mark, and D. Sul (2004): "Unbiased Estimation of the Half-Life to PPP Convergence in Panel Data," *NBER Working Paper No. 10614.*

Chortareas, G. E., and G. Kapetanios (2004): "The Yen Real Exchange Rate may be Stationary After All: Evidence from Nonlinear Unit-Root Tests," *Oxford Bulletin of Economics and Statistics*, 65, 113–131.

Chow, A. (1960): "Tests of Equality Between Sets of Coefficients in two Linear Regressions," *Econometrica*, 28, 591–605.

Chowdhry, B., R. Doll, and Y. Xia (2005): "Extracting Inflation from Stock Returns to Test Purchasing Power Parity," *American Economic Review*, 95, 255–276.

Cumby, R. E. (1996): "Forecasting Exchange Rates and Relative Prices with the Hamburger Standard: Is What You Want What You Get With McParity?," *NBER Working Papers 5675.*

Frankel, J. A., and A. K. Rose (1996): "A Panel Project on Purchasing Power Parity: Mean Reversion Within and Between Countries," *Journal of International Economics*, 40, 209–224.

Froot, K., and K. Rogoff (1995): "Perspectives on PPP and Long-Run Real Exchange Rates," in *The Handbook of International Economics*, ed. by G. Grossman, and K. Rogoff, pp. 1647–1688. Elsevier Press.

Hakkio, C. (1984): "A Re-examination of Purchasing Power Parity, A multi-country and multi-period study," *Journal of International Economics*, 17, 265–277.

Harris, R. D., and E. Tzavalis (1999): "Inference for Unit Roots in Dynamic Panels Where the Time Dimension is Fixed," *Journal of Econometrics*, pp. 201–226.

Harvey, A. C., and D. Bates (2002): "Multivariate Unit Root Tests and Testing for Convergence," *University of Cambridge DAE Working Paper No. 0301.*

Hausman, J. A. (1978): "Specification Tests in Econometrics," *Econometrica*, 46(1).

Im, K. S., M. H. Pesaran, and Y. Shin (2003): "Testing for Unit Roots in Heterogeneous Panels," *Journal of Econometrics*, 115(1), 53–74.

Kapetanios, G. (2003): "Determining the Poolability Properties of Individual Series in Panel Datasets," *Queen Mary, University of London Working Paper No. 499.*

——— (2007): "Dynamic Factor Extraction of Cross-Sectional Dependence In Panel Unit Root Tests," *Forthcoming in Jurnal of Applied Econometrics.*

Levin, A., and C. F. Lin (1992): "Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties," *University of California, San Diego, Discussion Paper No. 92-93.*

Lothian, J. R., and M. P. Taylor (1996): "Real Exchange Rate Behavior: The Recent Float from the Perspective of the Past Two Centuries," *Journal of Political Economy*, 104, 488–509.

Lyhagen, J. (2000): "Why not use standard panel unit root test for testing PPP," *Working Paper Series in Economics and Finance 413, Stockholm School of Economics.*

Macdonald, R. (1995): "Long-run Exchange Rate Modelling: a Survey of the Recent Evidence," *IMF Stff Papers*, 42, 437–489.

MacDonald, R. (1996): "Panel Unit Root Tests and Real Exchange Rates," *Economics Letters*, 50, 7–11.

Mark, N. C. (2001): *International Macroeconomics and Finance.* Blackwell.

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

Murphy, A., and Y. Zhu (2008): "Unraveling the complex interrelationships between exchange rates and fundamentals," *Journal of Banking and Finance*, 32(6), 1150–1160.

Ng, S., and P. Perron (1995): "Unit Root Tests in ARMA Models with Data Dependent Methods for the Selection of the Truncation Lag," *Journal of the American Statistical Association*, 90, 268–281.

Nikolaou, K. (2008): "The behaviour of the real exchange rate: Evidence from regression quantiles," *Journal of Banking and Finance*, 32(5), 664–679.

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

Oh, K. Y. (1996): "Purchasing Power Parity and Unit Root Tests Using Panel Data," *Journal of International Money and Finance*, 15, 405–410.

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

Papell, D. H., and H. Theodoridis (1998): "Increasing Evidence of Purchasing Power Parity over the Current Float," *Journal of International Money and Finance*, 17, 41–50.

——— (2001): "The Choice of Numeraire Currency in Panel Tests of Pourchasing Power Parity," *Journal of Money, Credit and Banking*, 33, 790–803.

Pesaran, M. H. (2003): "A Simple Panel Unit Root Test in the Presence of Cross section Dependence," *Cambridge University DAE Working Paper, number 0346.*

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

PHILLIPS, P. C. B., AND D. SUL (2002): “Dynamic Panel Estimation and Homogeneity Testing Under Cross Sectional Dependence,” *Cowles Foundation Discussion Paper No. 1362.*

QUAH, D. (1995): “Exploiting Cross-Sectional Variations for Unit Root Inference in Dynamic Data,” *Economics Letters*, 44, 9–19.

ROGOFF, K. (1996): “The Purchasing Power Parity Puzzle,” *Journal of Economic Literature*, XXXIV, 647–668.

SARNO, L., AND G. VALENTE (2006): “Deviations from purchasing power parity under different exchange rate regimes: Do they revert and, if so, how?,” *Journal of Banking and Finance*, 30(11), 3147–3169.

SERLETIS, A., AND P. GORAS (2004): “Long-horizon regression tests of the theory of purchasing power parity,” *Journal of Banking and Finance*, 28(8), 1961–1985.

SHIN, Y., AND A. SNELL (2003): “Mean Group Tests for Stationarity in Heterogenous Panels,” Mimeo, University of Edinburgh.

TAYLOR, M. P., AND L. SARNO (1998): “The Behaviour of Real Exchange Rates During the Post-Bretton Woods Period,” *Journal of International Economics*, 46, 281–312.

WAGNER, M. (2007): “On PPP, Unit Roots and Panels,” *Forthcoming in Empirical Economics.*

WESTERLUND, J. (2007): “A Note on the Pooling of Individual PANIC Unit Root Tests,” *Unpublished Manuscript, Lund University.*

WU, Y. (1996): “Are Real Exchange Rates Nonstationary? Evidence from a Panel Data Test.,” *Journal of Money, Credit and Banking*, 28, 54–63.

ZIEMER, R. F., AND M. E. WETZSTEIN (1983): “A Stein Rule Method for Pooling Data,” *Economics Letter*, 11(1).

Appendix

Proof of Theorem 1

Denote the α_T-critical value of the $z_{\bar{t}}$ test by c_T. Note that this is a one-sided test which rejects if the test statistic is smaller than c_T. In order to prove the theorem we have to prove three statements.

(I) For all N, $\lim_{T\to\infty} Pr(z_{\bar{t}} < c_T) = 1$, if $z_{\bar{t}}$ has been constructed on the set of series $\mathbf{Y_i}$ where $\mathbf{Y_i}$ contains at least one stationary series.

(II) For all N, $\lim_{T\to\infty} Pr(z_{\bar{t}} < c_T) = 0$, if $z_{\bar{t}}$ has been constructed on the set of series $\mathbf{Y_i}$, where $\mathbf{Y_i}$ contains no stationary series.

(III) If a set of series $\mathbf{Y_i}$ contains at least one stationary series, then the minimum DF test will correspond to a stationary series, with probability approaching 1.

By (I), the algorithm will not stop, with probability approaching 1, as long as there exist any stationary series in the panel dataset. By (III), stationary series will be removed first from the panel. Finally, by (II) the algorithm will stop as soon as no stationary series are contained in the panel dataset and hence the theorem holds. We prove (I) first. For a stationary series $t_{j,T} = O_p(T^{1/2})$. This implies that $\bar{t}_T$ is, at least, $O_p(T^{1/2}/N)$, and hence $z_{\bar{t}}$, is, at least, $O_p(\sqrt{T/N})$ for all N, even if $\mathbf{Y_i}$ contains only one stationary series. Since $N/T \to 0$, it follows that $\sqrt{T/N} \to \infty$. This implies that the test based on $z_{\bar{t}}$ is consistent even for one stationary series. It is, then, clear that $\frac{c_T}{\sqrt{T/N}} \to 0$ is sufficient for (I) to hold. We start with the case $N \to \infty$. As $N \to \infty$, $z_{\bar{t}}$ will tend to a normal distribution and hence normal critical values will be used. We, then, have that

$$\alpha_T = \int_{-\infty}^{c_T} c(x)dx \leq exp(c_T/d) \tag{13}$$

for some $d > 0$, where $c(x)$ is the pdf of a standard normal random variable. Thus, the following inequality holds: $\frac{-\ln \alpha_T}{\sqrt{T/N}} \geq \frac{-c_T}{\sqrt{T/Nd}}$, and by condition (ii) of the Theorem statement, (I) holds. For finite N, we note that each $t_{j,T}$ converges to a functional of a Brownian motion, as $T \to \infty$. Various authors, (see, e.g., Abadir (1995, (3.3) and (3.4))) have shown that either tail of the pdf of $t_{j,T}$ is proportional to the standard normal pdf. As (13) still holds for finite sums of independent standard normal random variables, it follows that both the above inequality and (I) hold. (II) is easily seen to hold since, if $\mathbf{Y_i}$ contains no stationary series, then $z_{\bar{t}} = O_p(1)$ but $\alpha_T \to 0$ and so, $c_T \to -\infty$. Finally, we know that with probability approaching 1, $t_{l,T} < t_{m,T}$ asymptotically if $\mathcal{I}_{\mathbf{i}^l} = 1$ and $\mathcal{I}_{\mathbf{i}^m} = 0$ for all l, m, since $t_{l,T} \xrightarrow{p} -\infty$ and $t_{m,T} = O_p(1)$. This implies (III). Hence, the Theorem is proven.

Proof of Theorem 2

We start by noting that with probability approaching 1 all series for which $\mathcal{I}_{\mathbf{i}^l} = 1$ will be detected by the sequential test before any series for which $\mathcal{I}_{\mathbf{i}^l} = 0$ and before the algorithm stops. This follows from (I) and (III) in the proof of Theorem 1 which hold for fixed α. So for all series for which $\mathcal{I}_{\mathbf{i}^j} = 1$ it follows that $\hat{\mathcal{I}}_{\mathbf{i}^j} = 1$. This proves (i). We then need to consider series for which $\mathcal{I}_{\mathbf{i}^j} = 0$.

We then need to examine the behaviour of the sequence of panel unit root tests on a set of nonstationary series. Thus, without loos of generality we assume that, for all series in the panel, $\mathcal{I}_{\mathbf{i}^j} = 0$. We have that $\lim_{T\to\infty} Pr\left(\left\{\hat{\mathcal{I}}_{\mathbf{i}^1} = 1\right\}\right) = \alpha$, where $\{.\}$ denotes an event. Using the rules of conditional probability, $\lim_{T\to\infty} Pr\left(\left\{\hat{\mathcal{I}}_{\mathbf{i}^2} = 0\right\} \middle| \left\{\hat{\mathcal{I}}_{\mathbf{i}^1} = 1\right\}\right) = 1 - \alpha$. Then,

$$\lim_{T\to\infty} Pr\left(\left\{\hat{\mathcal{I}}_{\mathbf{i}^2} = 0\right\} \cap \left\{\hat{\mathcal{I}}_{\mathbf{i}^1} = 1\right\}\right) = \lim_{T\to\infty} Pr\left(\left\{\hat{\mathcal{I}}_{\mathbf{i}^2} = 0\right\} \middle| \left\{\hat{\mathcal{I}}_{\mathbf{i}^1} = 1\right\}\right) Pr\left(\left\{\hat{\mathcal{I}}_{\mathbf{i}^1} = 1\right\}\right) = \alpha(1-\alpha) \quad (14)$$

Further,

$$\lim_{T\to\infty} \frac{Pr\left(\left\{\hat{\mathcal{I}}_{\mathbf{i}^3} = 0\right\} \cap \left\{\hat{\mathcal{I}}_{\mathbf{i}^2} = 1\right\} \cap \left\{\hat{\mathcal{I}}_{\mathbf{i}^1} = 1\right\}\right)}{Pr\left(\left\{\hat{\mathcal{I}}_{\mathbf{i}^2} = 1\right\} \cap \left\{\hat{\mathcal{I}}_{\mathbf{i}^1} = 1\right\}\right)} = \lim_{T\to\infty} Pr\left(\left\{\hat{\mathcal{I}}_{\mathbf{i}^3} = 0\right\} \middle| \left\{\hat{\mathcal{I}}_{\mathbf{i}^2} = 1\right\} \cap \left\{\hat{\mathcal{I}}_{\mathbf{i}^1} = 1\right\}\right) = 1 - \alpha \quad (15)$$

and

$$\lim_{T\to\infty} Pr\left(\left\{\hat{\mathcal{I}}_{\mathbf{i}^2} = 1\right\} \cap \left\{\hat{\mathcal{I}}_{\mathbf{i}^1} = 1\right\}\right) = \lim_{T\to\infty} Pr\left(\left\{\hat{\mathcal{I}}_{\mathbf{i}^2} = 1\right\} \middle| \left\{\hat{\mathcal{I}}_{\mathbf{i}^1} = 1\right\}\right) Pr\left(\left\{\hat{\mathcal{I}}_{\mathbf{i}^1} = 1\right\}\right) = \alpha^2 \quad (16)$$

Thus,

$$\lim_{T\to\infty} Pr\left(\left\{\hat{\mathcal{I}}_{\mathbf{i}^3} = 0\right\} \cap \left\{\hat{\mathcal{I}}_{\mathbf{i}^2} = 1\right\} \cap \left\{\hat{\mathcal{I}}_{\mathbf{i}^1} = 1\right\}\right) = \alpha^2(1-\alpha) \quad (17)$$

By recursion, it follows that

$$\lim_{N_2\to\infty} \lim_{T\to\infty} Pr\left(\left\{\hat{\mathcal{I}}_{\mathbf{i}^j} = 1\right\} \cap \left\{\hat{\mathcal{I}}_{\mathbf{i}^{j-1}} = 1\right\} \cap \ldots \cap \left\{\hat{\mathcal{I}}_{\mathbf{i}^2} = 1\right\} \cap \left\{\hat{\mathcal{I}}_{\mathbf{i}^1} = 1\right\}\right) = \lim_{N_2\to\infty} \left(\alpha^{N_2-1}(1-\alpha)\right) = 0 \quad (18)$$

Thus, the theorem follows.

Queen Mary
University of London

This working paper has been produced by
the Department of Economics at
Queen Mary, University of London

Department of Economics
Queen Mary, University of London
Mile End Road
London E1 4NS
Tel: +44 (0)20 7882 5096
Fax: +44 (0)20 8983 3580
Web: www.econ.qmul.ac.uk/papers/wp.htm