

# PPP in OECD Countries: An Analysis of Real Exchange Rate Stationarity, Cross-Sectional Dependency and Structural Breaks

Mark J. Holmes · Jesús Otero ·  
Theodore Panagiotidis

Published online: 9 December 2011  
© Springer Science+Business Media, LLC 2011

**Abstract** The stationarity of OECD real exchange rates over the period 1972–2008 is tested using a panel of 26 member countries. The methodology followed stems from the need to meet several key concerns: (i) the identification of which panel members are stationary; (ii) the presence of cross-sectional dependence among the countries in the panel; and (iii) the identification of potential structural breaks that might have occurred at different points in time. To address these concerns, we employ a recent test that examines the time series properties of the data within a panel framework, namely the Hadri and Rao (Oxford Bulletin of Economics and Statistics 70: 245–269, 2008) panel stationarity test. The real exchange rates of the 26 OECD countries are found to be stationary when considered as a panel, but only after allowing for endogenously-determined structural breaks and cross section dependence. We also find that once these structural breaks are removed from the underlying series, the half-life of shocks to the real exchange rate is much shorter than has been calculated in earlier studies.

**Keywords** Heterogeneous dynamic panels · Purchasing power parity · Mean reversion · Panel stationarity test

**JEL** F31 · F33 · G15

---

M. J. Holmes (✉)  
Department of Economics, Waikato University, Hamilton, New Zealand  
e-mail: holmesmj@waikato.ac.nz

J. Otero  
Facultad de Economía, Universidad del Rosario, Bogotá, Colombia  
e-mail: jesus.otero@urosario.edu.co

T. Panagiotidis  
Department of Economics, University of Macedonia, Thessaloniki, Greece  
e-mail: tpanag@uom.gr

## 1 Introduction

Domingo de Báñez, a Spanish theologian, was probably unaware of the long-lasting impact that his statement in 1594 would have on the economic literature: “one party may lawfully agree to repay a large sum to another, corresponding to the amount required to buy the same parcel of goods that the latter might have bought it if he had not delivered his money in exchange”.<sup>1</sup> This statement represents the origins of the Purchasing Power Parity (PPP) hypothesis that holds that prices in different countries should move towards equality when measured in common currency terms. This is of interest for both policy makers and economists. On the one hand, PPP becomes a prediction model for exchange rates and a criterion for judging the over- and undervaluation of currencies and, on the other hand, exchange rate determination models employ some notion of PPP in their construction. Therefore, the quality of policy advice, insofar as it is based on these models, may depend on the validity of PPP.

For many EU countries, the combination of the move towards a common currency, along with explicit inflation convergence targets as specified in the Maastricht treaty, provides an ideal case study with which to test for long-run PPP. While PPP is one of the most widely tested economic hypotheses, the verdict with regard to its validity remains open.<sup>2</sup> Early studies such as Balassa (1964), Dornbusch (1980) and Frenkel (1981) find no evidence supporting PPP, while Hakkio (1984) finds in favor. More recently, mixed evidence on the PPP relationship is provided by studies such as Johnson (1990), MacDonald (1993), Pedroni (2001), Taylor (2002) and Ceglowski (2003).

While unit root testing of the real exchange rate has become a commonly used methodological approach adopted by the literature, it is well known that unit root tests applied to single series suffer from low power. In an attempt to overcome this deficiency, the more recent PPP literature has considered two main approaches. One strand of the literature uses data spanning over very long periods of time, as in, *inter alia*, Lothian and Taylor (1996) and Hegwood and Papell (1998) while the other considers the application of panel unit root techniques, as in, *inter alia*, Wu (1996) and O’Connell (1998). In this paper we follow the latter approach since panel data offer a way forward in terms of enhanced test power, as they combine both the time-series dimension with the cross-sectional dimension such that fewer time observations are required for these tests to have power. More specifically, we investigate real exchange rate stationarity and long-run PPP for a panel comprising 26 OECD members.

In our study, we address three problems that potentially can undermine the validity of any conclusions drawn concerning long-run PPP. These are the inability to determine which panel members are stationary, the existence of potential cross sectional dependencies between panel members and the presence of structural breaks. Our methodology is based on the application of the Hadri and Rao (2008) panel stationarity test which allows for structural breaks, and addresses cross sectional dependencies through a bootstrap procedure.

<sup>1</sup> See also Sarno and Taylor (2002).

<sup>2</sup> See, for example, Sarno and Taylor (2002) and Sarno (2005).

The most commonly used unit root tests applied to panels include Maddala and Wu (MW) (1999), Levin et al. (2002), Im et al. (2003), hereafter IPS, and Pesaran (2007) which test the joint null hypothesis of a unit root against the alternative of at least one stationary series in the panel. These tests are based on augmented Dickey–Fuller (ADF) (1979) statistics across the cross-sectional units of the panel. However, IPS (2003, p.73) warn that the heterogeneous nature of the alternative hypothesis in their test requires care when interpreting such results because the null hypothesis of a unit root in each cross section may be rejected when only a fraction of the series in the panel is stationary.<sup>3</sup> A further issue here is that the presence of cross-sectional dependencies can undermine the asymptotic normality of the IPS test and lead to over-rejection of the null hypothesis.

Recent panel data studies of OECD real exchange rates include Wu (1996), Papell and Theodoridis (1998), Luintel (2001), Koedijk et al. (2004), Harris et al. (2005), Im et al. (2005), Papell (2006), López and Papell (2007), Narayan (2008) and López (2008). These studies utilize a range of panel unit root tests and find evidence in favour of stationarity. While many of these investigations address cross-sectional dependency or structural breaks in the data, the question still remains as to which and how many of the individual panel members are responsible for rejecting the joint non-stationary null leading to the conclusion that long-run PPP holds. In sharp contrast to most of the existing literature on PPP, we examine the stationarity of OECD real exchange rates using the Hadri (2000) test of the null hypothesis that all individual series are stationary against the alternative of at least a single unit root in the panel. This test offers a key advantage insofar as we may conclude that all the real exchange rates in the panel are stationary if the joint null hypothesis is not rejected.

In addition to this, an important feature of our analysis is that we allow for the presence of structural breaks, serial correlation, and cross-sectional dependency across the individuals in the panel. More specifically, we also apply the Hadri and Rao (2008) panel stationarity test with structural breaks, which allows the possibility of different endogenously determined breaking dates across the individuals in the panel. This procedure takes into account both serial correlation and cross-sectional dependency through the implementation of an AR-based bootstrap. Harris et al. (2005) test the null hypothesis of joint stationarity of 17 OECD US dollar real exchange rates. Their test is a nonparametric panel stationarity test that is robust to the presence of serial dependence within the panel and cross-sectional dependence (of an arbitrary nature) across the panel. After allowing for structural breaks, they reject the null and conclude against the validity of long-run PPP. By contrast, we apply Hadri-based panel procedures based on Kwiatkowski, Phillips, Schmidt and Shin (KPSS) (1992) stationarity testing. Monte Carlo simulations undertaken by Carrion-i-Silvestre and Sansó (2006) indicate that this approach offers reasonable power combined with less size distortion.

The outline of the paper is as follows. Section 2 briefly reviews the econometric approach adopted to test for stationarity in a heterogeneous panel of data. This corresponds to the Hadri-based approaches as applied to test for stationarity of OECD real exchange rates, allowing for the likely presence of endogenously

<sup>3</sup> See Koedijk et al. (2011) on the importance of allowing for heterogeneity in testing for PPP.

determined structural breaks and cross section dependence, where the latter is accounted for by means of a bootstrap procedure. Section 3 describes the data and presents the results of the empirical analysis, which offer support for the view that real exchange rates are stationary when considered as a panel of data, and only after allowing for structural breaks and cross section dependence. Therefore, long-run PPP holds across our entire panel of OECD countries. Section 4 presents a VAR-based analysis of impulse-response functions. This indicates that the half life of shocks to the real exchange rate is much shorter than the calculations provided elsewhere in the literature once structural breaks are taken into account. Section 5 concludes.

## 2 Methodology

### 2.1 Stationarity in Heterogeneous Panel Data in the Presence of Structural Breaks

Hadri (2000) proposes a Lagrange Multiplier (LM) procedure to test the null hypothesis that all individual series in a panel, denoted as  $y_{it}$ , are stationary (either around a mean or around a linear trend) against the alternative of at least a single unit root. The two LM tests proposed by Hadri are panel versions of the tests developed by KPSS to test for mean and trend stationarity in a univariate setup. More specifically, the model specifications considered by Hadri (2000) are:

$$\text{Model 1 : } y_{it} = \alpha_i + r_{it} + \varepsilon_{it}, \quad (1)$$

$$\text{Model 2 : } y_{it} = \alpha_i + r_{it} + \beta_i t + \varepsilon_{it}, \quad (2)$$

where  $r_{it}$  is a random walk,  $r_{it} = r_{it-1} + u_{it}$ ,  $\varepsilon_{it}$  and  $u_{it}$  are mutually independent normal distributions and  $\varepsilon_{it}$  and  $u_{it}$  are *i.i.d* across  $i$  and over  $t$ , with  $E[\varepsilon_{it}] = 0$ ,  $E[\varepsilon_{it}^2] = \sigma_{\varepsilon,i}^2 > 0$ ,  $E[u_{it}] = 0$  and  $E[u_{it}^2] = \sigma_{u,i}^2 \geq 0$ . In the models postulated above the number of time observations is denoted  $t=1, \dots, T$ , while the number of cross sections in the panel (in our case trade weighted multilateral real exchange rates) is denoted  $i=1, \dots, N$ . The null hypothesis that all cross sections in the panel are stationary is given by  $H_0 : \sigma_{u,i}^2 = 0$ ,  $i=1, \dots, N$  while under the alternative hypothesis there is at least one non-stationary cross section in the panel, that is  $H_1 : \sigma_{u,i}^2 > 0$ ,  $i=1, \dots, N_1$  and  $\sigma_{u,i}^2 = 0$ ,  $i=N_1+1, \dots, N$ .

In a recent paper, Hadri and Rao (2008) extend the Hadri stationarity tests to allow for the presence of a structural break. In particular, these authors consider the following four different types of models of structural break under the null hypothesis of stationarity (where it should be mentioned that the model labels used in this paper are not the same used by Hadri and Rao):

$$\text{Model 3 : } y_{it} = \alpha_i + r_{it} + \delta_i D_{it} + \varepsilon_{it}, \quad (3)$$

$$\text{Model 4 : } y_{it} = \alpha_i + r_{it} + \delta_i D_{it} + \beta_i t + \varepsilon_{it}, \quad (4)$$

$$\text{Model 5 : } y_{it} = \alpha_i + r_{it} + \beta_i t + \gamma_i DT_{it} + \varepsilon_{it}, \quad (5)$$

$$\text{Model 6 : } y_{it} = \alpha_i + r_{it} + \delta_i D_{it} + \beta_i t + \gamma_i DT_{it} + \varepsilon_{it}, \quad (6)$$

where, in addition to the terms already defined, the variables  $D_{it}$  and  $DT_{it}$  are dummy variables that help characterize the type of structural break in the models. These dummy variables are given by:

$$D_{it} = \begin{cases} 1, & \text{if } t > T_{B,i}, \\ 0 & \text{otherwise} \end{cases}$$

and

$$DT_{it} = \begin{cases} t - T_{B,i}, & \text{if } t > T_{B,i}, \\ 0, & \text{otherwise} \end{cases}$$

where  $T_{B,i}$  denotes the time of occurrence of the break, and  $T_{B,i} = \omega_i T$  with  $\omega_i \in (0,1)$  indicating the break fraction for the individual  $i$ , that is the position of the break point relative to the whole sample period. In turn, the parameters  $\delta_i$  and  $\gamma_i$  measure the magnitude of the structural break and allow for the possibility of different breaking dates across the individuals in the panel. Model 3 includes an intercept as deterministic component and allows for a shift in the level of the series. Model 4 contains an intercept and a linear trend term and allows for a shift in the level of the series. Model 5 includes an intercept and a linear trend and permits a change in the slope of the series. Lastly, Model 6 incorporates an intercept and a linear trend and permits a change in both the level and the slope of the series.<sup>4</sup>

At this point, it is worth highlighting that the models postulated above permit us to test different versions of the PPP hypothesis. Model 1 tests the traditional version of PPP, according to which the real exchange rate is a stationary variable around a constant mean. Model 2 allows for a trend stationary real exchange rate under the null hypothesis. The inclusion of a time trend may, for example, enable an assessment of whether breakdowns in PPP are due to changes in productivity differentials. By contrast, the remaining models allow us to test what Hegwood and Papell (1998) refer to as *quasi PPP* according to which the real exchange rate revert to a changing mean. These different structural change possibilities are captured in Models 3 to 6 where we also allow for a changing trend. Throughout the period under investigation, numerous economic and political factors might have shifted only the mean (Models 3 and 4), only the slope of the series (Model 5) or both (Model 6). In a similar spirit, Hegwood and Papell (1998) analyze real exchange rate for a two century study period and discuss the importance of structural change in terms of a range of events that include changes in exchange rate regime, oil price shocks, changes in government, the Great Depression and so on. These types of events have the potential to impact on the equilibrium value of the real exchange

<sup>4</sup> In their study of GDP per capita, Carrion-i-Silvestre et al. (2005) analyse two of the models considered by Hadri and Rao (2008), namely the model with breaks in the level and no time trend, and the model with breaks in the level and in the time trend.

rate. What is clear is that a longer study period makes the possibility of structural change more likely.

Hadri and Rao (2008) use a systematic approach to select the appropriate model specification for each individual time series  $y_{it}$  among the six model specifications postulated above. The first stage in this systematic approach is that of endogenously determining the location of the unknown break point. To do this, Hadri and Rao suggest estimating for each individual in the panel and for each model, the break date  $\widehat{T}_{B,i,k}$ . This can be accomplished by minimizing with respect to  $0 < \omega_i < 1$ , the residual sum of squares (RSS) from the relevant regression under the null hypothesis, where  $i = 1, \dots, N$  denotes the cross section units in the panel, and  $k = 1, 2, \dots, 6$  indicates the six models stated in Eqs. (1) to (6) above, respectively. It should be noted that in practice Models 1 and 2 are estimated only once since they do not involve the dummy variables  $D_{it}$  and  $DT_{it}$ . Then, for each individual time series in the panel the break-type model is chosen by minimizing the Schwarz Information Criterion.

Let us denote  $\widehat{\varepsilon}_{it}$  the residuals obtained from the estimation of the chosen break-type model. Then, the individual univariate KPSS stationarity test is given by:

$$\eta_{i,T,k}(\widehat{\omega}_i) = \frac{\sum_{t=1}^T S_{it}^2}{T^2 \widehat{\sigma}_{\varepsilon_i}^2},$$

where  $S_{it}$  is the partial sum process of the residuals given by  $S_{it} = \sum_{j=1}^t \widehat{\varepsilon}_{ij}$ , and  $\widehat{\sigma}_{\varepsilon_i}^2$  represents a consistent estimator of the long-run variance of  $\widehat{\varepsilon}_{it}$  from the appropriate regression. In their original paper, KPSS favor the use of a nonparametric estimator of  $\widehat{\sigma}_{\varepsilon_i}^2$  based on a Bartlett window having a truncation lag parameter of  $l_q = \text{integer} [q(T/100)^{1/4}]$ , where  $q$  is set equal to 0, 4 and 12 in their Monte Carlo simulation results (the value of the test statistics turns out to be sensitive to the choice of  $q$ ). Nevertheless, in a subsequent paper, Caner and Kilian (2001) observe that stationarity tests, like the KPSS, exhibit very low power after correcting for size distortions. Thus, for the purposes of this study we follow the recent approach advocated by Sul et al. (2005), who propose a new boundary condition rule to obtain a consistent estimate of the long-run variance  $\widehat{\sigma}_{\varepsilon_i}^2$ , that improves the size and power properties of the KPSS stationarity tests. The idea in Sul et al. (2005) is to start by estimating the following AR model for the residuals:

$$\widehat{\varepsilon}_{it} = \rho_{i,1} \widehat{\varepsilon}_{i,t-1} + \dots + \rho_{i,p_i} \widehat{\varepsilon}_{i,t-p_i} + v_{it}, \tag{7}$$

where the lag length of the autoregression can be determined for example using the Schwarz Information Criterion (SIC) or the General-To-Specific (GETS) algorithm proposed by Hall (1994). Then, the long-run variance estimate of  $\widehat{\sigma}_{\varepsilon_i}^2$  is obtained with the boundary condition rule:

$$\widehat{\sigma}_{\varepsilon_i}^2 = \min \left\{ T \widehat{\sigma}_{v_i}^2, \frac{\widehat{\sigma}_{v_i}^2}{(1 - \widehat{\rho}_i(1))^2} \right\},$$

where  $\widehat{\rho}_i(1) = \widehat{\rho}_{i,1}(1) + \dots + \widehat{\rho}_{i,p_i}(1)$  denotes the autoregressive polynomial evaluated at  $L=1$ . Lastly,  $\widehat{\sigma}_{v_i}^2$  is the long-run variance estimate of the residuals in Eq. (7)

that is obtained using a quadratic spectral window Heteroscedastic and Autocorrelation Consistent (HAC) estimator.<sup>5</sup>

The Hadri and Rao (2008) panel stationarity test statistic is constructed by calculating the simple average of the individual univariate KPSS stationarity tests, that is:

$$LM_{T,N,k}(\hat{\omega}_i) = \frac{1}{N} \sum_{i=1}^N \eta_{i,T,k}(\hat{\omega}_i).$$

After a suitable standardization, using appropriate moments, it can be shown that the statistic given in the previous equation follows a standard normal limiting distribution. More formally:

$$Z_k(\hat{\omega}_i) = \frac{\sqrt{N}(LM_{T,N,k}(\hat{\omega}_i) - \bar{\xi}_k)}{\bar{\zeta}_k} \Rightarrow N(0, 1), \tag{8}$$

where  $\bar{\xi}_k = \frac{1}{N} \sum_{i=1}^N \xi_{i,k}$  and  $\bar{\zeta}_k^2 = \frac{1}{N} \sum_{i=1}^N \zeta_{i,k}^2$  are respectively the mean and variance required for standardization. It is important to mention that in the case of Models 3 to 6 these two moments depend upon the estimated position of the break relative to the whole sample period, that is  $\hat{\omega}_i$ . Thus, for example, for Model 3 the mean required for standardization is given by  $\xi_{i,3} = \frac{1}{6}(2\omega_i^2 - 2\omega_i + 1)$  while the variance is  $\zeta_{i,3}^2 = \frac{1}{45}(2\omega_i^4 - 4\omega_i^3 + 6\omega_i^2 - 4\omega_i + 1)$ . The interested reader is referred to Hadri and Rao (2008), Theorem 3, for the corresponding mean and variance of Models 4 to 6.

A critical assumption underlying the Hadri and Rao (2008) test is that of cross section independence among the individual time series in the panel.<sup>6</sup> To allow for the presence of cross-sectional dependency, Hadri and Rao recommend implementing an autoregressive (AR) bootstrap, which we outline below for the reader’s convenience. First, we correct for serial correlation using the AR model for the residuals given in (7), and obtain  $\hat{v}_{it}$  which is centered around zero. Second, following Maddala and Wu (1999), the residuals  $\hat{v}_{it}$  are re-sampled with replacement with the cross-section index fixed, so that their cross-correlation structure is preserved; the resulting bootstrap innovation  $\hat{v}_{it}$  is denoted  $\hat{v}_{it}^*$ . Third,  $\hat{\varepsilon}_{it}^*$  is generated recursively using the mechanism:

$$\hat{\varepsilon}_{it}^* = \hat{\rho}_{i,1}\hat{\varepsilon}_{i,t-1} + \dots + \hat{\rho}_{i,p_i}\hat{\varepsilon}_{i,t-p_i} + v_{it}^*$$

where, in order to ensure that initialization of  $\hat{\varepsilon}_{it}^*$  becomes unimportant, a large number of  $\hat{\varepsilon}_{it}^*$  are generated, let us say  $T + Q$  values and then the first  $Q$  values are

<sup>5</sup> Additional Monte Carlo evidence reported by Carrion-i-Silvestre and Sansó (2006) also indicates that the proposal in Sul et al. (2005) is to be preferred since the KPSS statistics exhibit less size distortion and reasonable power.

<sup>6</sup> Giulietti et al. (2009) examine the effect of cross sectional dependency in the Hadri (2000) panel stationarity tests in the absence of structural breaks and with no serial correlation. They find that even for relatively large  $T$  and  $N$  the Hadri (2000) tests suffer from severe size distortions, the magnitude of which increases as the strength of the cross-sectional dependence increases. To correct the size distortion caused by cross-sectional dependence, Giulietti et al. (2009) apply the bootstrap method and find that the bootstrap Hadri tests are approximately correctly sized.

discarded; this approach is advocated by Chang (2004); for the purposes of our estimations we set  $Q=30$ . Finally, the bootstrap samples of  $y_{it}^*$  are calculated by adding  $\hat{\varepsilon}_{it}^*$  to the deterministic component of the corresponding chosen model, and the Hadri LM statistic is calculated for each  $y_{it}^*$ . The results later presented and discussed in the paper are based on 2,000 bootstrap replications used to derive the empirical distribution of the LM statistic.

### 3 Data and Empirical Analysis

The data set, obtained from the OECD Main Economic Indicators database, consists of 146 quarterly observations on the trade weighted multilateral real exchange rate for 26 OECD members. This sample includes 12 EU countries that have so far adopted the Euro as their currency (Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal and Spain); three EU members currently not Euro members (Denmark, Sweden and United Kingdom); and 11 non-EU countries (Australia, Canada, Iceland, Japan, Korea, Mexico, New Zealand, Norway, Switzerland, Turkey and United States). The sample period runs from 1972q1 to 2008q2 and the real exchange rates are expressed in natural logarithm form.

This approach is in contrast to most studies in the PPP literature that employ bilateral real exchange rates against a base country. While it could be argued that this implies our findings may not be directly comparable to existing results, the use of multilateral rates are very useful from an economic policy standpoint given that the real exchange rate is a measure of international competitiveness and, more often than not, countries have more than one significant trading partner. A second data issue associated with the use of multilateral rates is that cross sectional dependence might be introduced in the analysis thereby requiring a careful econometric handling; see for example O'Connell (1998). In line with the earlier discussion, the issue of cross section dependence can be accounted for by means of the outlined bootstrap method in the previous section.

Table 1 presents the results from applying the KPSS stationarity test to the real exchange rate series of the countries listed above, based on the model with intercept only.<sup>7</sup> In order to correct for possible serial correlation, the autoregressive processes in (7) are estimated for up to  $p=12$  lags where the optimal number of lags chosen is based on the SIC and GETS algorithm. This algorithm involves testing whether the last autoregressive coefficient is statistically different from zero (say, at the 10% significance level). If it is insignificant, the order of the autoregression then is reduced by one until the last coefficient is statistically significant. As to be expected, the SIC mostly favors a more parsimonious specification than that obtained from applying the GETS algorithm. The results based on the SIC see the null hypothesis of stationarity rejected at the 10% significance level for four countries, for four more countries rejection is at the 5%

<sup>7</sup> Qualitatively similar findings are obtained when using the model with intercept and trend. In the interests of brevity, these results are not reported here.

**Table 1** Individual stationarity tests (model with constant)

Countries	Lag length based on:			
	SIC		GETS	
	Lag	Statistic	Lag	Statistic
EU countries in Euro zone				
Austria	1	1.033 <sup>c</sup>	6	1.423 <sup>c</sup>
Belgium	2	0.371 <sup>a</sup>	4	0.419 <sup>a</sup>
Finland	2	1.108 <sup>c</sup>	4	1.461 <sup>c</sup>
France	2	0.408 <sup>a</sup>	10	0.286
Germany	2	0.194	10	0.328
Greece	2	0.172	6	0.169
Ireland	1	0.483 <sup>b</sup>	10	0.629 <sup>b</sup>
Italy	2	0.101	4	0.095
Luxembourg	2	0.543 <sup>b</sup>	4	0.700 <sup>b</sup>
Netherlands	2	0.125	9	0.168
Portugal	1	0.867 <sup>c</sup>	5	1.103 <sup>c</sup>
Spain	2	0.429 <sup>a</sup>	8	0.521 <sup>b</sup>
EU countries not in Euro zone				
Denmark	1	0.497 <sup>b</sup>	9	0.724 <sup>b</sup>
Sweden	2	1.853 <sup>c</sup>	12	2.215 <sup>c</sup>
United Kingdom	2	0.101	2	0.101
Non-EU countries				
Australia	1	0.855 <sup>c</sup>	6	1.192 <sup>c</sup>
Canada	2	2.301 <sup>c</sup>	7	2.402 <sup>c</sup>
Iceland	2	0.210	11	0.093
Japan	2	1.840 <sup>c</sup>	6	1.986 <sup>c</sup>
Korea	2	0.375 <sup>a</sup>	4	0.412 <sup>a</sup>
Mexico	1	0.057	4	0.133
New Zealand	1	0.035	9	0.218
Norway	1	0.057	2	0.097
Switzerland	2	0.830 <sup>c</sup>	11	1.360 <sup>c</sup>
Turkey	1	0.582 <sup>b</sup>	3	0.527 <sup>b</sup>
United States	5	0.198	5	0.198

<sup>a</sup>, <sup>b</sup> and <sup>c</sup> indicate 10, 5 and 1% levels of significance, respectively, based on finite sample critical values calculated from the response surfaces in Sephton (1995). The long-run variance required to calculate the KPSS statistic is consistently estimated using the new boundary condition rule put forward by Sul et al. (2005)

significance level, and for eight additional countries rejection is at the 1% level. There are ten countries for which we fail to reject the null hypothesis of stationarity, thus supporting the view that PPP holds for less than half the sample. A similar picture emerges when inspecting the results based on GETS lag length

selection, thereby not providing a clear indication in favor of PPP. Overall, the data provide only weak support for stationarity using the traditional version of PPP with a constant only.

We now consider the application of the Hadri panel stationarity test to the real exchange rate series. Table 2 reports Hadri (2000) test statistics along with corresponding  $p$ -values given in brackets. The  $p$ -values are based on the standard normal distribution (one-sided test statistic). The results indicate that the null hypothesis of panel stationarity is clearly rejected (at the 1% significance level) independently of the country groupings considered.

However, failure to account for potential cross section dependence will result in severe size distortion of the Hadri test statistics. We therefore apply an AR-based bootstrap to the Hadri tests as outlined in the previous section, with the aim of correcting not only for serial correlation but also for cross-section dependence. In addition, the analysis thus far has made no consideration for the possibility of structural breaks. Hence, in what follows we apply the Hadri and Rao (2008). This allows us to test for panel stationarity in the presence of an unknown break point that is endogenously determined for each individual country in the panel, and for the four models postulated in Eqs. (3) to (6).<sup>8</sup>

The results reported in Table 3 indicate that all the multilateral real exchange rates under consideration exhibit evidence of a structural break. It is noticeable that the model specifications postulated in Eqs. (1) and (2), for mean and trend stationarity, respectively, are never selected. The most favored functional form is that given in Eq. (6), which includes both intercept and linear trend terms, and permits a change in both the level and slope of the series. As to the timing of the breaks, the results reported in the third column indicate that for the 12 Euro members the break dates range from 1978 to 2001, with several breaks occurring during various stages of the 1990s.<sup>9</sup> In their study of PPP in the Eurozone, López and Papell (2007) also find that the process of convergence towards PPP begins prior to the creation of the single currency, following the currency crises of 1992 and 1993, the adoption of the Maastricht Treaty, and the official completion of the Single Market. Harris et al. (2005), in testing a panel of US dollar real exchange rates for stationarity, find a predominance of break dates in 1980s. In our study, the presence of some breaks dates prior to the 1990s coincide with events such as the creation of the exchange rate mechanism in 1979 and the various exchange rate realignments that took place in the 1980s.<sup>10</sup>

<sup>8</sup> The empirical results reported in the paper were implemented using the computer software RATS and are based on a GAUSS code which was kindly provided to the authors by Yao Rao.

<sup>9</sup> We also considered the post 1988 period in order to take into account the rise and fall in the dollar in the early 1980s and the fact that central bank intervention fell continuously after 1988. This approach is consistent with the view that structural breaks are of such magnitude that may affect all individuals in the panel at the same time. Results not reported here confirm that the multilateral real exchange rates are mean stationary processes when considered as a panel, but only after allowing for the presence of cross-section dependence. This conclusion reflects the importance of accounting for the presence of structural breaks either exogenously or endogenously determined. We are grateful to an anonymous referee for raising this point.

<sup>10</sup> Of course, at this point it may well be argued that there exists the possibility of additional structural breaks in the multilateral real exchange rates. However, visual inspection of the resulting residuals from the chosen break-type models does not reveal evidence of the presence of additional structural breaks.

**Table 2** Panel stationarity tests (model with constant)

Countries	Lag length based on:			
	SIC		GETS	
	Statistic	<i>p</i> -value	Statistic	<i>p</i> -value
All	14.942	[0.000]	19.368	[0.000]
Euro12	7.462	[0.000]	10.331	[0.000]
EU15	10.076	[0.000]	13.670	[0.000]
11 non-EU	11.205	[0.000]	13.813	[0.000]
G7	10.144	[0.000]	10.794	[0.000]

Euro12 comprises Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal and Spain; EU15 comprises EU12 plus Denmark, Sweden and the United Kingdom; 11 Non-EU comprises Australia, Canada, Iceland, Japan, Korea, Mexico, New Zealand, Norway, Switzerland, Turkey and the United States; and G7 comprises Canada, France, Germany, Italy, Japan, UK and the US. The *p*-values of the Hadri test are based on the standard normal distribution

The residuals from the chosen break-type model for each country can be used to construct the KPSS statistics. These in turn can be used as in Eq. (8) to compute the Hadri and Rao (2008) panel stationarity test. As can be seen in Table 4, we are unable to reject the joint null hypothesis of panel stationarity, independently of the method used to select the optimal lag length of the autoregressive processes in (7). This finding is in sharp contrast to Harris et al. (2005) who reject the null of joint stationarity and conclude that long-run PPP does not hold. The only exception to this in our study is in the case of the panel comprising the 11 non-EU countries when the GETS procedure is employed. It is worth noting that if we wrongly assume cross-sectional independence among the countries in the panel, and use the standard normal distribution for the purposes of inference, then the null hypothesis of panel stationarity is clearly rejected for the panels comprising all 26 countries and the 11 Non-EU countries when using the GETS procedure. This finding highlights the importance of allowing for the possibility of potential cross-sectional dependencies among the individual countries in the panel. While we offer evidence of long-run PPP, it should be noted that it is evidence that is consistent with quasi PPP insofar as the results are conditioned against models with changing constants and/or trends. These statistical models are inconsistent with traditional approaches to PPP because economic factors have most likely driven changes in the equilibrium real exchange rates and these subsequent breaks need to be accounted for.<sup>11</sup>

<sup>11</sup> Another strand of the literature focuses on nonlinearities instead of structural breaks; see, for instance, Jiménez and Robles-Fernandez (2010) and Norman (2010). For completeness, we also employed the recently proposed Ucar and Omay (2009) unit-root test for non-linear heterogeneous panels. Within this framework, we tested the null hypothesis that all series in the panel are linear non-stationary against the alternative that some are non-linear stationary, as described by exponential smooth transition (ESTAR) models. The results, which are not reported here but are available upon request, reject the joint non-stationarity null throughout the different panels.

**Table 3** Individual stationarity tests with endogenously determined structural break

Countries	Model	Break date	Lag length based on:			
			SIC		GETS	
			<i>p</i>	Statistic	<i>p</i>	Statistic
EU countries in Euro zone						
Austria	6	1995Q1	1	0.022	11	0.066
Belgium	6	1979Q1	1	0.027	1	0.027
Finland	6	1990Q4	1	0.014	9	0.085
France	6	1998Q1	2	0.041	10	0.073
Germany	3	1978Q3	1	0.030	10	0.046
Greece	6	1983Q4	2	0.040	6	0.037
Ireland	6	1997Q2	1	0.050	1	0.050
Italy	4	1991Q1	1	0.062	11	0.088
Luxembourg	4	1979Q2	1	0.032	1	0.032
Netherlands	4	2001Q1	1	0.017	2	0.027
Portugal	6	1988Q2	1	0.021	5	0.049
Spain	4	1991Q3	1	0.017	11	0.048
EU countries not in Euro zone						
Denmark	6	1978Q1	1	0.030	2	0.044
Sweden	4	1986Q2	2	0.023	12	0.080
United Kingdom	3	1977Q2	1	0.043	2	0.068
Non-EU countries						
Australia	5	1999Q2	1	0.014	12	0.163
Canada	5	2000Q4	2	0.073	7	0.078
Iceland	6	1991Q3	2	0.028	9	0.082
Japan	6	1991Q2	2	0.017	10	0.064
Korea	6	1995Q4	1	0.016	8	0.053
Mexico	6	1981Q1	4	0.052	4	0.052
New Zealand	6	1997Q4	1	0.017	9	0.160
Norway	6	1991Q1	1	0.014	9	0.069
Switzerland	6	1977Q2	2	0.094	4	0.126
Turkey	6	1984Q2	1	0.054	9	0.062
United States	4	1985Q4	1	0.015	5	0.038

The long-run variance required to calculate the KPSS statistic in the presence of a structural break is consistently estimated using the new boundary condition rule put forward by Sul et al. (2005)

#### 4 Speed of Adjustment and Half-Life

Following Rogoff (1996), the PPP literature has keenly debated whether the likely values for the half-life of shocks to the real exchange rate under the recent float are of the order 3–5 years. The panel-based studies of PPP have provided comparable calculations, for example Wu (1996) finds a half life of 2.5 years. In

**Table 4** Hadri and Rao panel stationarity test with endogenously determined structural break and allowing for cross-sectional dependence

Countries	Lag length based on:			
	SIC		GETS	
	Statistic	<i>p</i> -value	Statistic	<i>p</i> -value
All	-2.242	[0.806]	2.808	[0.321]
Euro12	-1.778	[0.770]	0.351	[0.861]
EU15	-2.075	[0.865]	0.138	[0.948]
11 non-EU	-0.866	[0.370]	5.907	[0.004]
G7	-1.256	[0.750]	-0.052	[0.978]

The *p*-values are bootstrap based on 2,000 replications

this section, we examine the speed at which real exchange rates adjust to exogenous shocks or innovations. For this, we use half-life estimates based on the impulse response functions that result from estimating a Vector Autoregressive (VAR) model.<sup>12</sup> The VAR model itself consists of the residuals that are obtained from the estimation of the corresponding chosen break-type model for each of the 26 OECD countries under consideration. Since we are unable to reject the hypothesis that all the (de-trended) real exchange rates in our sample are stationary, they are considered as suitable for modeling in a VAR framework.

It is well known that in the case of a simple AR(1) process, the estimated value of the autoregressive coefficient, say  $\theta$ , can be used to calculate the approximate half-life of a shock to the dependent variable using the formula  $-\ln(2)/\ln(\theta)$ . However, for more complicated processes this formula is no longer valid. This provides a case for the preferred use of impulse response functions, as recommended for example by Goldberg and Verboven (2005), footnote 11. For the specific purpose of this paper, we consider the generalized impulse response (GIR) functions developed by Pesaran and Shin (1998). GIR functions offer the advantage of being invariant to the way shocks in the underlying VAR model are orthogonalised, and therefore provide an important extension to the traditional impulse response analysis, which has been found to be sensitive to the ordering of the variables in the VAR; see for example Lütkepohl (2005).

For the specification of the VAR model, an important initial stage in the analysis is the selection of the optimal lag length, which involves selecting an order high enough such that one can be reasonably confident that the optimal order will not exceed it. In the case of the present application we set 4 lags as the maximum order of the VAR model, and use both the SIC or Akaike Information Criterion (AIC) to select the optimal order of the models. The optimal order of the VAR model is one

<sup>12</sup> Alternatively, Jordá (2005) proposes a more a-theoretical approach, where impulse responses are approximated via local projections. Lo (2008), employing non-linear models from the STAR family, produces half-lives that are within the 3 to 5 years interval. Also within a non-linear framework, Norman (2010) develops a method for estimating all half lives associated with shocks of many different sizes and finds evidence of half lives shorter than the 3 to 5 years often reported in the literature.

and four lags using the SIC and the AIC, respectively. The associated generalized impulse responses plot the time profile of the effect of an own unit shock in the real exchange rate of a country, measured by one standard deviation. Then, the resulting statistically significant lag weights for each country are normalized so that they add up to one, and the half-life is calculated as the number of quarters required for 50% (or the first half) of the adjustment to take place.

The results, reported in Table 5, indicate that the estimated half life to own shocks is considerably less than the typical half lives computed in previous studies. For example, using a VAR(4) model based on the AIC, the half lives vary from 9

**Table 5** Half-life estimates (in quarters) from generalized impulse response functions

Countries	Lag length based on:	
	SIC VAR(1)	AIC VAR(4)
EU countries in Euro zone		
Austria	2	2
Belgium	3	2
Finland	2	2
France	2	2
Germany	3	4
Greece	1	7
Ireland	3	9
Italy	2	7
Luxembourg	4	4
Netherlands	3	5
Portugal	2	3
Spain	5	4
EU countries not in Euro zone		
Denmark	3	3
Sweden	4	2
United Kingdom	3	2
Non-EU countries		
Australia	2	1
Canada	4	6
Iceland	2	2
Japan	3	2
Korea	2	2
Mexico	3	2
New Zealand	3	2
Norway	2	4
Switzerland	3	4
Turkey	2	2
United States	6	6

quarters in the case of Ireland down to 1 quarter in the case of Australia, with an average half life over the 26 countries of 4 quarters. Our finding of a comparatively short half life is confirmed if instead we use a VAR(1) model based on the SIC. The relatively quick rate of adjustment has resulted from removing the broken-time deterministic trends from the underlying series. Indeed, these findings are consistent with the literature on structural breaks and unit root testing that goes back to Perron (1989) and others insofar as the presence of breaks in the data can induce spurious non-stationarity and longer times for mean reversion. This includes Hegwood and Papell (1998) who use six bilateral real exchange rates for industrialized countries over a substantially longer study period, and find substantially shorter half-lives for deviations from PPP once structural breaks are accounted for.<sup>13</sup>

## 5 Concluding Remarks

This paper examines panel data evidence on real exchange rate stationarity and long-run PPP for 26 OECD countries. In doing so, we have applied a testing procedure which collectively enables us to address problems associated with identifying individual stationarity within the panel, cross-sectional dependency, as well as structural breaks in the data. The standard tests that have been applied in the literature are of a joint non-stationary null, the rejection of which may be attributable to the stationary behavior of as little as one panel member. In sharp contrast to this approach, we employ the Hadri and Rao (2008) test based on the null hypothesis of joint stationarity. This approach also enables us to address problems that have bedeviled earlier panel data studies of PPP: the possibility of structural breaks in the data as well as cross-sectional dependence among panel members. We address the latter problem through pursuing an AR bootstrap approach to the Hadri and Rao test.

Our empirical analysis starts off by applying the KPSS test of univariate stationarity as well as the Hadri test of panel stationarity, both of which support the view that not all the real exchange rates under consideration are mean stationary. However, once the latter is extended along the lines of Hadri and Rao, so that allowance is made for the potential effect of cross-sectional dependencies and endogenously-determined breaks across the sample, the null hypothesis that the real exchange rates are jointly stationary in the long run is no longer rejected. This applies to the full sample of OECD countries as well as sub-groups based on Euro, EU and G7 membership. We therefore offer qualified support for those earlier panel studies of PPP that reject a joint null of non-stationarity because our results depend on a particular times series model that is restrictive on economic grounds. By allowing for structural breaks, however, we find that the half-life of a shock to long-run PPP is in fact much shorter than hitherto thought by much of the empirical literature that sidesteps issues associated with structural breaks.

**Acknowledgment** We are most grateful to two anonymous referees and to the Editor-in-Chief for constructive comments and suggestions that helped to improve our paper. The usual disclaimer applies.

<sup>13</sup> A different perspective is provided by Imbs et al. (2005) who find that the estimated persistence of real exchange rates falls when product heterogeneity is taken into account.

## References

- Balassa B (1964) The purchasing power doctrine: a reappraisal. *J Polit Econ* 72:584–596
- Caner M, Kilian L (2001) Size distortions of tests of the null hypothesis of stationarity: evidence and implications for the PPP debate. *J Int Money Finance* 20:639–657
- Carrion-i-Silvestre JL, Sansó A (2006) A guide to the computation of stationarity tests. *Empir Econ* 31:433–448
- Carrion-i-Silvestre JL, Del Barrio T, López-Bazo E (2005) Breaking the panels: an application to the GDP per capita. *Econ J* 8:159–175
- Ceglowski J (2003) The law of one price: intranational evidence for Canada. *Can J Econ* 36:373–400
- Chang Y (2004) Bootstrap unit root tests in panels with cross-sectional dependency. *J Econ* 120:263–293
- Dickey DA, Fuller WA (1979) Distribution of the estimators for autoregressive time series with a unit root. *J Am Stat Assoc* 74:427–431
- Dornbusch R (1980) Exchange rate economics: where do we stand? *Brookings Paper Econ Activ* 1:143–185
- Frenkel J (1981) The collapse of purchasing power parities during the 1970s. *Eur Econ Rev* 16:145–165
- Giulietti M, Otero J, Smith J (2009) Testing for stationarity in heterogeneous panel data in the presence of cross section dependence. *J Stat Comput Simul* 79:195–203
- Goldberg P, Verboven F (2005) Market integration and convergence to the law of one price: evidence from the European car market. *J Int Econ* 65:49–73
- Hadri K (2000) Testing for stationarity in heterogeneous panels. *Econ J* 3:148–161
- Hadri K, Rao Y (2008) Panel stationarity test with structural breaks. *Oxf Bull Econ Stat* 70:245–269
- Hakkio C (1984) A re-examination of purchasing power parity: a multi-country and multi-period study'. *J Int Econ* 17:265–277
- Hall A (1994) Testing for a unit root in time series with pretest data-based model selection. *J Bus Econ Stat* 12:461–470
- Harris D, Leybourne S, McCabe B (2005) Panel stationarity tests for purchasing power parity with cross-sectional dependence. *J Bus Econ Stat* 23:395–409
- Hegwood ND, Papell DH (1998) Quasi purchasing power parity. *Int J Finance Econ* 3:279–289
- Im K, Pesaran MH, Shin Y (2003) Testing for unit roots in heterogeneous panels. *J Econ* 115:53–74
- Im K, Lee J, Tieslau M (2005) Panel LM unit-root tests with level shifts. *Oxf Bull Econ Stat* 67:393–419
- Imbs J, Mumtaz H, Ravn MO, Rey H (2005) PPP strikes back: aggregation and the real exchange rate. *Q J Econ* 120:1–43
- Jiménez JA, Robles-Fernandez MD (2010) PPP: delusion or reality? evidence from a nonlinear analysis. *Open Econ Rev* 21:679–704
- Johnson D (1990) Co-integration, error and purchasing power parity between Canada and the United States. *Can J Econ* 23:839–855
- Jordá O (2005) Estimation and inference of impulse responses by local projections. *Am Econ Rev* 95:161–182
- Koedijk KG, Tims B, van Dijk MA (2004) Purchasing power parity and the Euro area. *J Int Money Finance* 23:1081–1107
- Koedijk KG, Tims B, van Dijk MA (2011) Why panel tests of purchasing power parity should allow for heterogeneous mean reversion. *J Int Money Finance* 30:246–267
- Kwiatkowski D, Phillips PCB, Schmidt P, Shin Y (1992) Testing the null hypothesis of stationarity against the alternative of a unit root. *J Econ* 54:159–178
- Levin A, Lin C-F, Chu C-S (2002) Unit root tests in panel data: asymptotic and finite-sample properties. *J Econ* 108:1–24
- Lo MC (2008) Nonlinear PPP deviations: a Monte Carlo investigation of their unconditional half-Life. *Stud Nonlinear Dynam Econometrics* 12: No. 4, Article 5
- López C (2008) Evidence of purchasing power parity for the floating regime period. *J Int Money Finance* 27:156–164
- López C, Papell DH (2007) Convergence to purchasing power parity at the commencement of the Euro. *Rev Int Econ* 15:1–16
- Lothian JR, Taylor MP (1996) Real exchange rate behavior: the recent float from the perspective of the past two centuries. *J Polit Econ* 104:488–510
- Luintel KB (2001) Heterogeneous panel unit root tests and purchasing power parity. *Manch Sch* 69:42–56
- Lütkepohl H (2005) New introduction to multiple time series analysis. Springer Verlag, Heidelberg
- MacDonald R (1993) Long-run purchasing power parity: is it for real? *Rev Econ Stat* 75:690–695

- Maddala GS, Wu S (1999) A comparative study of unit root tests with panel data and a new simple test. *Oxf Bull Econ Stat* 61:631–652
- Narayan PK (2008) The purchasing power parity revisited: new evidence for 16 OECD countries from panel unit root tests with structural breaks. *J Int Financ Market Inst Money* 18:137–146
- Norman S (2010) How well does nonlinear mean reversion solve the PPP puzzle? *J Int Money Finance* 29:919–937
- O’Connell PGJ (1998) The overvaluation of purchasing power parity. *J Int Econ* 44:1–19
- Papell DH (2006) The panel purchasing power parity puzzle. *J Money, Credit, Bank* 38:447–467
- Papell DH, Theodoridis H (1998) Increasing evidence of purchasing power parity over the current float. *J Int Money Finance* 17:41–50
- Pedroni P (2001) Purchasing power parity tests in cointegrated panels. *Rev Econ Stat* 83:727–731
- Perron P (1989) The great crash, the oil price shock, and the unit root hypothesis. *Econometrica* 57:1361–1401
- Pesaran MH (2007) A simple panel unit root test in the presence of cross section dependence. *J Appl Econ* 22:265–312
- Pesaran MH, Shin Y (1998) Generalized impulse response analysis in linear multivariate models. *Econ Lett* 58:17–29
- Rogoff K (1996) The purchasing power parity puzzle. *J Econ Lit* 34:647–668
- Sarno L (2005) Viewpoint: towards a solution to the puzzles in exchange rate economics: where do we stand? *Can J Econ* 38:673–708
- Sarno L, Taylor MP (2002) *The economics of exchange rates*. Cambridge University Press, Cambridge
- Sephton PS (1995) Response surface estimates of the KPSS stationarity test. *Economics Letters* 47:255–261
- Sul D, Phillips PCB, Choi CY (2005) Prewhitening bias in HAC estimation. *Oxf Bull Econ Stat* 67:517–546
- Taylor A (2002) A century of purchasing power parity. *Rev Econ Stat* 84:139–150
- Ucar N, Omay T (2009) Testing for unit root in nonlinear heterogeneous panels. *Econ Lett* 104:5–8
- Wu Y (1996) Are real exchange rates nonstationary? evidence from a panel-data test. *J Money, Credit, Bank* 28:54–63