Applied Financial Economics, 2006, **16**, 199–211

Routledge
Taylor & Francis Group

The real exchange rate and the Purchasing Power Parity puzzle: further evidence

Sofiane H. Sekioua[a,*] and Menelaos Karanasos[b]

[a]*Institut National de la Planification et de la Statistique (INPS) and Citibank N.A.*
[b]*Business School, Brunel University Uxbridge, Middlesex, UB8 3PH, UK*

This study presents additional evidence on the convergence speeds of real exchange rates. Using median unbiased estimation, impulse response analysis and long horizon data sampled annually and monthly, we estimate the speeds at which deviations from purchasing power parity (PPP) die out. Both monthly and annual data have been used since temporal aggregation has been proposed as a possible cause of the implausibly large half-lives reported in the literature. Moreover, since reporting only point estimates provides an incomplete picture of the speed of convergence towards PPP, median unbiased confidence intervals are also estimated. The results show that the confidence intervals for the half-lives are typically very wide. Interestingly, however, the intervals estimated using monthly data are tighter than those estimated with annual data, though, they do not help solve the PPP puzzle. In fact, it appears that the point estimates of the half-lives obtained with monthly data are much larger. Therefore, on the basis of the evidence reported in this study, the results on temporal aggregation by Taylor (2001) are unlikely to have a major role in empirically explaining the PPP puzzle.

I. Introduction

According to the Purchasing Power Parity (PPP) theory, deviations are attributed to transitory monetary shocks which translate into real exchange rate variability due to the stickiness of nominal prices. Consequently, while PPP is compatible with the observed short-term volatility of real exchange rates, it also means that deviations should be short-lived as they can only occur during a time-frame in which nominal prices and wages are sticky, between 1 to 2 years. However, Rogoff (1996) points out that the observed persistence of real exchange rates is far too high to be explained by existing theoretical models (e.g. Dornbusch, 1976). Although growing evidence in support of mean-reversion towards PPP has been documented, consensus estimates of the speed of reversion are remarkably slow with half-lives ranging from 3 to 5 years, implying a slow parity reversion rate of between 13 to 20% per annum.

This is the so-called PPP puzzle and the empirical work that Rogoff cites in support of his consensus

*Corresponding author. E-mail: sofiane@sekioua.com

Applied Financial Economics ISSN 0960–3107 print/ISSN 1466–4305 online
http://www.tandf.co.uk/journals
DOI: 10.1080/09603100500389945

mostly comes from univariate studies with long horizon data[1] (Abuaf and Jorion, 1990; Diebold *et al*., 1991; Lothian and Taylor, 1996). For instance, using annual data spanning two centuries for dollar-sterling and franc-sterling real exchange rates, Lothian and Taylor (1996) find strong evidence of mean-reverting real exchange rate behaviour. The econometric estimates imply half-lives of 5.78 years for dollar-sterling and 2.76 years for franc-sterling.

The Lothian and Taylor (1996) result for the dollar-sterling constitute the slowest speed of mean-reversion mentioned by Rogoff (1996). In other words, the half-life for the dollar-sterling represents the upper limit for the consensus of Rogoff (1996) which has become the benchmark in the empirical literature. Taylor (2001), however, claimed that temporal aggregation can yield substantial biases in estimates of the half-lives. Specifically, using a model in which the real exchange rate follows an autoregressive AR process of order 1 at a higher frequency than that at which the data are sampled, Taylor (2001) shows that the degree of upward bias in the half-life rises as the degree of temporal aggregation increases.[2] Thus, the half-life can be seriously over-estimated if adjustment towards equilibrium takes place during a time-frame that is shorter than the sampling frequency of the data. The time aggregation problem is, none the less, a difficult issue for researchers to deal with since long spans of data are required to have a reasonable level of statistical power when tests of a unit root in the real exchange rate are applied and long spans of high frequency data do not exist (Sarno and Taylor, 2003). Indeed, studies that tend to provide evidence against the unit root null hypothesis are normally based on long spans of data sampled annually and the fact that evidence in favour of PPP is found at all is encouraging given the biases introduced by temporal averaging in historical data (Taylor, 2002).

This study investigates the issue of temporal aggregation by applying uniform tests to monthly and annual data spanning three major US real exchange rates and at least 75 years of experience. With such a long span, there may be sufficient data to have a reasonably powerful test[3] whilst the frequency of the data can help shed light on the issue raised by Taylor (2001). This is an important contribution of this study since long spans of monthly data have only recently been made available from Global Financial Data. Further, the study brings a recent empirical innovation to the long span of historical data.[4] The innovation is the median unbiased estimation (MUE) method of Gospodinov (2004) which allows for the construction of confidence intervals for the half-life of shocks from parity based on impulse response analysis. This is particularly important in the present context for the following two reasons. First, Murray and Papell (2002) illustrate the existence of a substantial amount of sampling variability in measuring the half-life and, as a result, the point estimate alone does not provide a complete description of the persistence of deviations from PPP. Therefore, it needs to be supplemented with confidence intervals in order to gauge the precision of the estimates. Second, the commonly used estimate of the half-life, $\mathrm{H.L} = \ln(1/2)/\ln(\alpha)$, which is based on an autoregressive (AR) model of order 1, assumes that shocks decay monotically, but for higher order

[1] Actually, deviations from PPP are so persistent that even these consensus estimates for the half-life have been challenged recently (Kilian and Zha, 2002; Murray and Papell, 2002). These studies report that the interval estimators of several univariate measures of persistence provide little support of the hypothesis of 3 to 5 years half-lives. This is supported by Lopez *et al*. (2005) who claim that the consensus itself is problematic whilst Murray and Papell (2005) argued that the puzzle is worse.

[2] The other source of the puzzle discussed by Taylor (2001) is nonlinear adjustment. While inviting alternative explanations for the PPP puzzle, Rogoff (1996) posits that transactions costs may be a contributing factor. Because of goods market frictions, there is a band of inaction within which nominal exchange rates can move without eliciting a quick response in prices. With slow price adjustment, real exchange rates converge very slowly inside the band (Dumas, 1992). The transaction costs view has gained popularity and prompted numerous studies (Michael *et al*., 1997; Baum *et al*., 2001; Taylor *et al*., 2001). The transaction costs explanation is useful in highlighting the significance of goods market impediments to price adjustment (Cheung *et al*., 2004). However, this contribution concentrates on the effects of temporal aggregation on the estimation of the half-life only since many of the stylized facts in the PPP literature have been obtained within the linear framework.

[3] Long samples are necessary to achieve a reasonable level of power as conventional tests may fail to reject the unit root null hypothesis even when real exchange rates exhibit slow mean-reversion. This low power problem is magnified for small samples, such as the recent floating experience, because a mean-reverting series could be drifting away from its long-run equilibrium in the short-run. In response to this, Diebold *et al*. (1991) showed that long samples are important for identifying mean-reversion in slowly decaying processes whilst Frankel (1990) argued that, with slow convergence speeds, the autoregressive parameter might be very close to unity and one would need long spans of data to reject the unit root null. Hence, with the data there might be sufficient span to have a powerful test.

[4] Of course, there is the potential problem of structural instability when using long data. This study follows the prior literature that utilizes long spans of data and assume that the dynamics of the real exchange rates are relatively stable over the sample period. In order to employ more powerful tests and long spans of data, one has to assume relatively stable dynamic processes over long periods (Rapach and Wohar, 2002).

AR processes this may not be the case. To remedy this, Cheung and Lai (2000) recommend using impulse response analysis.

The remainder of this paper is set as follows. Section II explains the econometrics of local-to-unity processes. Sections III and IV discuss the data and empirical results, respectively. Impulse response analysis is provided in Section V, robustness in Section VI. The last section concludes.

II. Empirical Methodology

Median unbiased estimation

The method employed in this study is due to Gospodinov (2004) and is based on inverting the likelihood ratio (LR) statistic of the largest root under a sequence of null hypotheses of possible values for the impulse response and the half-life. Starting from the following augmented Dickey–Fuller (ADF) regression which includes lagged first differences to account for serial correlation:

$$y_t = \alpha y_{t-1} + \sum_{i=1}^{k-1} \psi_i \, \Delta y_{t-i} + \varepsilon_t \tag{1}$$

where α is a measure of the persistence of the series (Andrews and Chen, 1994) and is cast as local-to-unity ($\alpha = 1 + c/T$ and holding c fixed as $T \to \infty$), $\varphi = (\alpha, \psi')' \in \Xi \subset R^p$ and the maximum likelihood estimator of φ is $\hat{\varphi}$. Suppose that one is interested in the null hypothesis that the impulse response function at horizon l, denoted by θ_l, is 0.5 (the half-life, defined as the number of periods it takes for deviations to subside permanently below 50% in response to a shock), versus the alternative $\theta_l \neq 0.5$, then this null or restriction can be written as $h(\varphi) = 0$, where $h \equiv \theta_l - 0.5$: $R^p \to R$ is a polynomial of degree l. Let $\tilde{\varphi}$ denote the restricted maximum likelihood estimator and LR_T the likelihood ratio statistic of the null. Gospodinov (2004) shows that the restricted estimator of α converges at a faster rate than the unrestricted estimator and this helps obtain a consistent estimate of the nuisance parameter c under the imposed restriction (null hypothesis). Moreover, the restricted estimation provides consistent estimates of the impulse response functions and thus the half-lives.[5]

The restricted LR estimator of Equation 1 under the null hypothesis $h(\varphi) = 0$ is:

$$LR_T \Rightarrow \frac{\left[\int_0^1 J_c^\tau(s) dW(s)\right]^2}{\int_0^1 J_c^\tau(s)^2 ds} \tag{2}$$

where $J_c^\tau(r) = J_c(r) - \int_0^1 J_c(s)ds$, $J_c(r) = \int_0^r \exp[(r-s)c]\, dW(s)$ is a homogenous Ornstein–Uhlenbeck process and $\Rightarrow$ denotes weak convergence. The limiting theory of LR is dominated by the near nonstationary component and is not affected by the presence of stationary components as measured by the second term in regression, $\sum_{i=1}^{k-1} \psi_i \, \Delta y_{t-i}$.

The employed method has many interesting features. First, contrary to standard asymptotic and bootstrap methods, which have been shown to have poor coverage properties, this method parameterizes α as a function of T and is expected to yield better small-sample and coverage performance. Second, the LR statistic does not require variance estimation for studentization. It is criterion function-based and is tracking closely the profile of the objective function. Also, the inversion of the LR statistic appears to shift the confidence intervals away from the nonstationarity region much more often compared to methods based on inverting the OLS estimator of α such as the grid bootstrap of Hansen (1999). Further, using a series of Monte Carlo experiments, Gospodinov (2004) shows that the inversion of the LR statistic appears to be controlling the coverage over a wide range of parameter configurations and across different forecasting horizons. This method is also expected to yield tight confidence intervals, which makes them highly informative.

Another statistic which takes into account the restricted and the unrestricted estimates is also proposed:

$$LR_T^{\pm} = \mathrm{sgn}[h(\hat{\varphi}) - h(\tilde{\varphi})]\sqrt{LR_T} \tag{3}$$

where $\mathrm{sgn}(\cdot)$ is the sign of $[h(\hat{\varphi}) - h(\tilde{\varphi})]$. This statistic can be used for constructing two-sided, equal-tailed confidence intervals and median unbiased estimates.

[5] The standard method for estimating Equation 1 is ordinary least squares (OLS) and the conventional asymptotic interval is based on the asymptotic $N(0, 1)$ approximation to the t-statistic which is valid only if $|\alpha| < 1$. This approximation is poor in practice especially when the persistence parameter $|\alpha|$ is close or equal to unity. Specifically, if the true persistence parameter is not unity, OLS estimates are biased downwards and confidence intervals based on asymptotic methods have poor coverage properties. When persistence is unity, the coverage problems of the asymptotic intervals stem from the fact that the asymptotic distribution of α is non-standard. Bootstrap methods are also poor. This is because the percentile-t bootstrap is based on the assumption that the bootstrap quantile functions are constant, which is false for the AR model. This nonconstancy persists in large samples if we cast α as local-to-unity as $\alpha = 1 + c/T$. In this case, the asymptotic distribution of the t-statistic depends on α through the nuisance parameter c that is not consistently estimable (Hansen, 1999). Thus, in the near unit setting, the interval does not properly control for Type I error (Basawa *et al.*, 1991).

Finally, the $100\eta\%$ confidence interval for the half-life, which is based on impulse response analysis, is: $C_\eta(l) = \{l \in L: LR_T \leq q_\eta(c)\}$, where $q_\eta(c)$ is the ηth quantile of the asymptotic distribution, l is the lead time of the impulse response function and $\tilde{\varphi} = \arg\max l_T(\varphi)$ subject to $\theta_l - 0.5 = 0$. The confidence interval for the half-life can be constructed using either $LR_T^{\pm}$ or LR_T.

III. Data

The data utilized in this study is extracted from the www.globalfindata.com database and includes monthly and annually sampled nominal exchange rate and consumer price index (CPI) series for the USA, the UK, Germany and Switzerland. The exact sample period for each country's real exchange rate *vis-à-vis* the US is: 1906 to 2002 for the UK, 1928 to 2002 for Germany and 1920 to 2002 for Switzerland. The start dates were in each case dictated by the availability of the monthly CPI series. In addition, the real exchange rate is defined according to the identity: $q_t \equiv s_t - p_t + p_t^*$ where s_t is the nominal exchange rate and p_t and p_t^* are the domestic and foreign price levels, respectively. Following the common practice in the literature, all the series are expressed in logarithms.[6]

IV. Empirical Results

Unit root test results

The study first tests for a unit root in the real exchange rates using the efficient generalized least squares (GLS) version of the Dickey–Fuller (DF) test due to Elliott *et al.* (1996) whose results are reported in Table 1. While most unit root tests are only concerned with testing the null that the largest root is unity against the alternative that it is less than one, the DF-GLS test tests the null against a specific alternative $H_1{:}\,\alpha < 1$ where $\alpha = 1 + c/T$. Further, using a sequence of tests of the null of a unit root against a set of stationary persistent alternatives, Elliott *et al.* (1996) showed substantial power gain from the DF-GLS method over the conventional ADF test (which has low power against close alternatives so that the unit root null can seldom be rejected for highly persistent variables). The lag lengths are chosen using the Akaike (AIC) and Schwarz (SIC) information criteria. For completeness, evidence is also reported for all possible values of k (up to 12 lags for monthly data and 6 lags for annual data) to determine if the results are sensitive to the number of lags included in the DF-GLS regression. Recently, Ng and Perron (2001) showed that criteria which choose too few lags for the DF-GLS test are badly sized and may produce inappropriate rejections of the unit root null. From Table 1, one can see that the DF-GLS test rejects the unit root null for Germany and the UK. For Switzerland, the null is hardly ever rejected so that long-run PPP did not hold over the last century between the USA and Switzerland.

The study also allowed for linear time trends in the DF-GLS regressions. While the rejection of the null in this case is not evidence of long-run PPP in the strict sense, one can ascribe a Harrod–Balassa–Samuelson (HBS) interpretation to a mean-reverting real exchange rate around a linear deterministic time trend. Specifically, the HBS effect is based on differential rates of productivity growth in traded and non-traded goods sectors. Evidently, the linear long-run trends may be purely deterministic (Obstfeld, 1993). None the less, allowing a time trend to be present does not seem to help very much as the evidence against the unit root null is rather weak. For the UK, the null is rarely rejected now. This finding is somewhat in contrast to what is reported in Lothian and Taylor (2000)[7] where the inclusion of a time trend in the autoregressive representation for the dollar/pound real exchange rate over 200 years is viewed as strengthening their claim for significant mean-reversion in the real exchange rate. Taylor (2002) also reports more rejections of the unit root null based on the

[6] Under the arbitrage view of PPP, the appropriate price index should cover only those goods that are traded internationally. It can be argued that the producer price index (PPI) is a better choice since it is heavily weighted towards tradable goods than the CPI. But data for the PPI do not exist in the form of monthly series spanning the last century.

[7] In their paper, Lothian and Taylor (2000) argue that given the economic history of the USA and the UK, if the time trend in an AR representation of the real exchange rate is proxying for HBS effects, it seems restrictive to assume that the effects are linear. In this study, however, a linear time trend is used for only two reasons. First, since a linear framework is adopted to study the real exchange rate, it seems appropriate to use a linear trend. Second, Gospodinov (2004) is followed by using a linear trend since the detrending method that is adopted in the theoretical framework of his paper is basically linear; OLS detrending to be specific. For these reasons, the present study will not explore alternative detrending methods; although, it is likely that more efficient detrending methods may increase the power of the tests and reduce the width of the confidence intervals for the parameter of interest (Elliott *et al.*, 1996). This nevertheless remains an interesting area for future research.

Table 1. Confidence intervals for the largest root and the half-life

k	DF-GLS	α_{OLS}	α_{MUE}	95_{lower}	95_{upper}	H.L	95_{lower}	95_{upper}
(a) UK, 1906–2002								
Monthly demeaned data								
2_{SIC}	[0.0880]	0.9905	0.9962	0.9916	1.0000	8.7645	3.3291	25.0076
3_{AIC}	[0.0595]	0.9899	0.9958	0.9914	1.0000	8.1233	3.3084	25.8368
4	[0.0745]	0.9903	0.9961	0.9916	1.0000	8.5706	3.3341	26.4218
5	[0.0860]	0.9903	0.9961	0.9915	1.0000	8.5510	3.3277	26.2556
6	[0.0995]	0.9902	0.9961	0.9915	1.0000	8.5197	3.3235	26.4223
7	[0.0820]	0.9905	0.9963	0.9915	1.0000	8.7797	3.3170	26.5893
8	[0.0760]	0.9902	0.9961	0.9914	1.0000	8.5444	3.3252	26.5888
9	[0.0755]	0.9901	0.9960	0.9915	1.0000	8.4731	3.3399	26.8530
10	[0.0740]	0.9900	0.9961	0.9913	1.0000	8.2789	3.3387	27.0211
11	[0.0875]	0.9902	0.9960	0.9914	1.0000	8.5605	3.3313	27.5040
12	[0.0710]	0.9898	0.9958	0.9911	1.0000	8.1503	3.3761	27.2575
Monthly detrended data								
2_{SIC}	[0.1356]	0.9889	0.9931	0.9843	1.0000	9.8730	2.7481	25.7693
3_{AIC}	[0.0900]	0.9880	0.9920	0.9834	1.0000	8.3180	2.7167	26.8436
4	[0.1261]	0.9885	0.9927	0.9837	1.0000	9.1472	2.7400	27.5112
5	[0.1336]	0.9885	0.9928	0.9839	1.0000	9.1533	2.7378	27.2635
6	[0.1341]	0.9884	0.9926	0.9837	1.0000	9.0956	2.7358	27.4302
7	[0.1681]	0.9888	0.9931	0.9842	1.0000	9.7946	2.7355	27.5879
8	[0.1181]	0.9884	0.9927	0.9838	1.0000	9.1588	2.7421	27.3992
9	[0.1201]	0.9882	0.9923	0.9836	1.0000	8.8387	2.7498	28.0123
10	[0.1061]	0.9880	0.9920	0.9830	1.0000	8.4881	2.7550	28.2617
11	[0.1281]	0.9883	0.9926	0.9834	1.0000	9.0083	2.7443	28.2749
12	[0.0995]	0.9876	0.9917	0.9829	1.0000	8.0942	2.7894	28.2924
Annual demeaned data								
$2_{AIC,SIC}$	[0.0260]	0.8607	0.8954	0.8028	0.9997	6.4563	2.7540	∞
3	[0.0460]	0.8681	0.9048	0.8103	1.0000	6.7630	1.8930	∞
4	[0.0460]	0.8661	0.9054	0.8098	1.0000	6.8558	1.8899	∞
5	[0.0790]	0.8731	0.9158	0.8142	1.0000	6.6137	1.8794	∞
6	[0.1256]	0.8852	0.9295	0.8274	1.0000	4.8605	1.8576	∞
Annual detrended data								
$2_{AIC,SIC}$	[0.0780]	0.8351	0.8816	0.7596	1.0000	6.3909	2.2812	∞
3	[0.1581]	0.8438	0.8994	0.7715	1.0000	7.1646	1.7700	∞
4	[0.1496]	0.8370	0.8921	0.7573	1.0000	6.9587	1.7667	∞
5	[0.2601]	0.8476	0.9135	0.7764	1.0000	8.1385	1.7612	∞
6	[0.4352]	0.8648	0.9535	0.7979	1.0000	∞	1.7395	∞
(b) Germany, 1928–2002								
Monthly demeaned data								
$2_{AIC,SIC}$	[0.0005]	0.9810	0.9814	0.9668	0.9941	3.7598	1.7753	∞
3	[0.0010]	0.9802	0.9805	0.9653	0.9932	3.6238	1.7685	15.4179
4	[0.0010]	0.9795	0.9797	0.9646	0.9930	3.5100	1.7640	12.5176
5	[0.0000]	0.9792	0.9795	0.9639	0.9923	3.4630	1.7624	11.8917
6	[0.0005]	0.9790	0.9793	0.9643	0.9919	3.4388	1.7632	11.6534
7	[0.0010]	0.9791	0.9794	0.9628	0.9922	3.4548	1.7612	12.1302
8	[0.0010]	0.9789	0.9792	0.9631	0.9922	3.4350	1.7638	11.9301
9	[0.0005]	0.9787	0.9789	0.9630	0.9918	3.4093	1.7667	11.6274
10	[0.0020]	0.9784	0.9787	0.9616	0.9919	3.3810	1.7752	11.1989
11	[0.0005]	0.9783	0.9784	0.9616	0.9918	3.3682	1.7774	11.0896
12	[0.0005]	0.9781	0.9783	0.9615	0.9916	3.3494	1.7853	10.7986
Monthly detrended data								
$2_{AIC,SIC}$	[0.0210]	0.9805	0.9837	0.9721	0.9998	4.3048	1.6823	∞
3	[0.0230]	0.9798	0.9828	0.9710	0.9987	4.0867	1.6791	∞
4	[0.0155]	0.9790	0.9821	0.9706	0.9972	3.9217	1.6759	30.7192
5	[0.0100]	0.9787	0.9815	0.9697	0.9970	3.8563	1.6736	31.8859
6	[0.0195]	0.9785	0.9813	0.9695	0.9961	3.8190	1.6734	32.6333
7	[0.0120]	0.9786	0.9815	0.9697	0.9973	3.8435	1.6731	33.3024
8	[0.0195]	0.9784	0.9813	0.9694	0.9960	3.8133	1.6756	33.0455

(*Continued*)

Table 1. Continued

k	DF-GLS	α_{OLS}	α_{MUE}	95_{lower}	95_{upper}	H.L	95_{lower}	95_{upper}
9	[0.0185]	0.9782	0.9811	0.9690	0.9966	3.7796	1.6795	33.9651
10	[0.0150]	0.9779	0.9808	0.9684	0.9967	3.7300	1.6877	34.4198
11	[0.0140]	0.9777	0.9805	0.9685	0.9963	3.7080	1.6909	37.8755
12	[0.0130]	0.9775	0.9803	0.9674	0.9958	3.6706	1.6990	39.7725
Annual demeaned data								
2_{SIC}	[0.0005]	0.7165	0.7330	0.5667	0.9036	2.7815	1.3653	7.7190
3	[0.0010]	0.6867	0.7001	0.5112	0.8779	2.8773	1.3619	6.4572
4	[0.0240]	0.7652	0.7901	0.5854	0.9955	2.6917	1.5222	∞
5	[0.0065]	0.7165	0.7287	0.5119	0.9360	2.7005	1.6816	19.7964
6_{AIC}	[0.0105]	0.7392	0.7540	0.5194	0.9638	2.7271	1.7053	∞
Annual detrended data								
2_{SIC}	[0.0105]	0.7084	0.7602	0.5984	0.9553	2.9513	1.2124	∞
3	[0.0175]	0.6762	0.7248	0.5452	0.9328	2.9979	1.1750	13.8596
4	[0.1696]	0.7574	0.8450	0.6443	1.0000	2.8097	1.3397	∞
5	[0.0675]	0.7057	0.7671	0.5640	1.0000	2.7958	1.5529	∞
6_{AIC}	[0.1001]	0.7193	0.7938	0.5779	1.0000	2.8072	1.5693	∞
(c) Switzerland, 1920–2002								
Monthly demeaned data								
2	[0.3057]	0.9960	0.9983	0.9931	1.0000	18.9942	5.8025	22.8373
3_{SIC}	[0.2611]	0.9960	0.9982	0.9930	1.0000	15.6264	5.9889	23.4687
4	[0.2386]	0.9959	0.9980	0.9929	1.0000	14.7123	5.9264	23.9206
5	[0.2546]	0.9959	0.9980	0.9929	1.0000	14.7941	5.9132	24.1715
6	[0.2446]	0.9959	0.9980	0.9929	1.0000	14.7267	5.9274	24.2757
7	[0.2541]	0.9960	0.9981	0.9930	1.0000	15.1473	5.9322	24.2910
8	[0.2206]	0.9956	0.9978	0.9928	1.0000	14.0213	5.8552	24.4427
9	[0.2326]	0.9956	0.9976	0.9927	1.0000	14.1014	5.8971	24.6703
10_{AIC}	[0.1931]	0.9953	0.9974	0.9924	1.0000	13.3505	5.7821	24.7535
11	[0.1906]	0.9952	0.9973	0.9923	1.0000	12.8579	5.7278	25.8374
12	[0.1551]	0.9949	0.9971	0.9920	1.0000	12.0740	5.6308	26.0038
Monthly detrended data								
2	[0.2306]	0.9902	0.9946	0.9857	1.0000	17.9301	2.9143	25.2694
3_{SIC}	[0.1576]	0.9893	0.9937	0.9850	1.0000	12.2742	2.8846	26.1803
4	[0.1721]	0.9891	0.9934	0.9847	1.0000	11.7908	2.8796	26.8452
5	[0.1606]	0.9893	0.9937	0.9848	1.0000	12.6906	2.8899	27.0951
6	[0.1456]	0.9889	0.9932	0.9843	1.0000	11.1551	2.8690	27.0958
7	[0.1761]	0.9895	0.9940	0.9849	1.0000	12.9473	2.8929	27.5961
8	[0.1276]	0.9887	0.9928	0.9844	1.0000	10.8713	2.8933	27.1765
9	[0.1116]	0.9881	0.9924	0.9837	1.0000	9.4670	2.8799	28.0246
10_{AIC}	[0.0865]	0.9877	0.9917	0.9828	1.0000	8.6481	2.8929	28.2715
11	[0.0715]	0.9877	0.9915	0.9828	1.0000	8.6248	2.8983	29.1142
12	[0.0580]	0.9870	0.9906	0.9822	1.0000	7.7939	2.9265	29.1785
Annual demeaned data								
2_{SIC}	[0.1931]	0.9360	0.9619	0.8911	1.0000	∞	4.8360	∞
3_{AIC}	[0.2446]	0.9368	0.9665	0.8958	1.0000	∞	2.9569	∞
4	[0.3217]	0.9400	0.9744	0.9015	1.0000	∞	2.5717	∞
5	[0.2731]	0.9402	0.9705	0.8951	1.0000	∞	2.6752	∞
6	[0.3257]	0.9381	0.9736	0.8953	1.0000	∞	2.6635	∞
Annual detrended data								
2_{SIC}	[0.0415]	0.8139	0.8558	0.7302	1.0000	5.7085	2.4476	∞
3_{AIC}	[0.1106]	0.8315	0.8833	0.7495	1.0000	6.9435	1.9768	∞
4	[0.2206]	0.8402	0.9037	0.7604	1.0000	8.3301	1.9798	∞
5	[0.0850]	0.8138	0.8620	0.7273	1.0000	6.8727	2.1971	∞
6	[0.1866]	0.8293	0.8901	0.7416	1.0000	6.9883	2.2642	∞

Note: The median unbiased estimates and confidence intervals for the largest root are constructed with the grid bootstrap of Hansen (1999) using the efficiently demeaned DF-GLS statistic; 1999 bootstrap replications at each of 51 grid-points. The optimal lag lengths for the unit root test statistics are set according to the Akaike (AIC) and Schwarz (SIC) information criteria. Figures in square brackets are p-values. The half-lives (H.L) estimated from the impulse response functions are measured in years.

DF-GLS test with the inclusion of a deterministic time trend. However, this has recently been challenged by Lopez *et al.* (2005) who employ the same data set as Taylor (2002) and argue that it is the use of sub-optimal lag selection which leads Taylor to find so many rejections in the presence of linear time trends, not the (trend) stationarity of the data. Indeed, Lopez *et al.* (2005) showed that the rejection of the unit root null hypothesis, even with long horizon data, is very sensitive to the number of lags considered. This is why k is chosen with both SIC and AIC and evidence reported for all the values of the lag length to side step the seemingly arcane topic of lag selection.

Finally, note that the rejection or lack thereof of the unit root null does not depend too much on the frequency of the data. This is not altogether surprising since Shiller and Perron (1985), Perron (1989) and Pierse and Snell (1995) demonstrate that the asymptotic local power of unit root tests with the same data span is independent of sampling frequency.

Confidence intervals for the largest root and the half-life

Although some of the evidence presented in the previous section supports the validity of long-run PPP, it offers little information about the speed at which deviations die out. To obtain such information, computation of persistence is needed and both the largest root and the half-life are used to quantify persistence. Table 1 reports the median unbiased estimates of α and the 95% MUE confidence intervals for this measure of persistence. The intervals are constructed by inverting the acceptance region of the powerful DF-GLS test of Elliott *et al.* (1996). Whilst the methodology in Section II ('Median unbiased estimation') is based on an ADF regression, the extension of this method to the DF-GLS test is simple. Instead of working with the data in levels as in Equation 1, one simply works with the GLS demeaned or detrended data in the DF-GLS regression. Moreover, the finite-sample distribution of the DF-GLS test is obtained using the grid bootstrap of Hansen (1999).

The median unbiased estimates of the largest root are indicative of strong persistence in real exchange rates.[8] For monthly data, the MUE estimates are seldom below 0.98 whilst for annual data the estimate is as low as 0.7. Most of the confidence intervals are found to contain unity as an upper bound with the notable exception of Germany. Although, the upper limits are in this case near the unit root boundary. Thus, Germany's real exchange rate is mean-reverting, albeit highly persistent. It also displays near-unit-root behaviour, precisely the type of behaviour that will be difficult for standard tests to detect for short samples. For the UK and Switzerland, the confidence intervals are not inconsistent with a unit root in the real exchange rates. It is interesting to note that the lower bounds are close to the point estimates and are never below 0.96 for monthly data. These bounds which can be re-interpreted as upper bounds for the fastest speed of mean-reversion are therefore consistent with the view that the variables under scrutiny are slow to mean-revert. Furthermore, the confidence intervals constructed with annual data are very wide and this may make it difficult to make definitive statements one way or another regarding the unit root/stationarity question. Overall, while the monthly confidence intervals from the powerful DF-GLS test appear to be quite tight, compared to their annual counterpart, and thus demonstrate the potential for sharper inference, they clearly imply that deviations from PPP are extremely persistent.

The MUE point estimates and confidence intervals for the half-life based on impulse response analysis are shown in Table 1. Starting with monthly data, the point estimates are around 8.12 to 9.87 years for the UK, 3.34 and 4.30 years for Germany and 7.79 and 18.99 years for Switzerland. For this latter country, notice the correspondence between the outcome of the unit root test and the half-life. Given that the null cannot be rejected for most lag lengths, the point estimate can be as high as 18.99 years. Interestingly, the point estimates are considerably larger than would be expected based on the consensus of Rogoff (1996). This corroborates Murray and Papell's (2002) claim that the literature surveyed by Rogoff does not accurately represent the behaviour of real exchange rates. In fact, it is only the point estimates for Germany that provide support for the consensus of Rogoff. If indeed it is considered a range of likely point estimates. But, although one cannot reject the idea that the half-life can be high, these

[8] The OLS estimate of the largest root, α_{OLS}, is also reported. This estimate is based on ADF regressions and thus does not optimally exploit the sample information in terms of power whereas α_{MUE}, based on the DF-GLS test, does. Besides, α_{OLS} is normally treated cautiously as it is biased downwards in small samples. However, given the large size of the samples, the bias disappears almost completely though it is still present for annual data. An idea we will keep in mind.

results do not determine how low the lower bound or high the upper bound can be.[9] To answer this question, the study now looks at the constructed confidence intervals which are robust to the presence of highly persistent data. The upper limits are consistent with high persistence with the exception of the demeaned German real exchange rate. Though, an upper limit of 10.79 years (lag 12) is hardly in line with the theory of PPP. However, for the lag lengths chosen with SIC and AIC, the upper bounds for the German data are infinite. Essentially, the evidence in the case of Germany is sensitive to the number of lags in the DF-GLS regressions. In any case, the uncertainty over the half-life is so big that a (lower bound) half-life of 1.77 years is also compatible with Germany's real exchange rate. For the UK and Switzerland, the lower bounds are outside the theoretical range of 1 to 2 years. Therefore, aside from Germany's real exchange rate, even the lower bounds are not compatible with PPP holding in the long-run.[10]

Looking now at the annual data, one can see that the confidence intervals are extremely wide. Accordingly, while the monthly intervals do not solve the puzzle, they at least provide one with much better inference than confidence intervals constructed with annual data. This is at least one benefit of using monthly data. The lower bounds for the UK and Germany are consistent with the theory of PPP. At the same time, however, one cannot rule out the possibility of no convergence in the data, that is one cannot reject that the half-life can be infinity. As a result, it is possible to interpret the intervals from annual data as simply not informative at all since they are consistent with virtually anything. Overall, regardless of the frequency of the data, the half-life estimates are always indicative of very strong persistence in PPP deviations.[11]

Notice also that the MUE point estimates of the half-lives from monthly data are even higher than from annual data. That is monthly real exchange rates converge substantially slower than annual real exchange rates. This differing-speed finding is quite intriguing. A possible explanation for this result, which obviously runs against Taylor's intuition, is that for monthly data tests detect some evidence of conditional heteroscedasticity in the residuals of the AR representations of the real exchange rates. Although the method of Gospodinov (2004) is robust to the presence of heterogeneity, persistence in the conditional variance may possibly lead to some distortions in the inference procedures.

Finally, apart from Germany, the point estimates of the half-lives are generally much higher than what is reported in the literature. This finding has important implications for Rogoff's consensus. Indeed, studies that are referred to in order to arrive at this consensus calculate OLS point estimates of the half-lives from univariate versions of the ADF regression (AR(1) processes, specifically) and annual data. Nonetheless, the OLS estimate is significantly downward biased in models that contain either an intercept or an intercept and a trend (Murray and Papell, 2005). The bias becomes worse as α gets larger which has particular relevance for the case of PPP persistence. Given that the OLS estimates are biased downwards, this will tend to provide an inaccurate picture of the speed of mean-reversion. In the present context, the problem is rather clear; although it is more severe for annually sampled data. For these data, the OLS estimate is always significantly downward biased by between 0.0122 and 0.0887. Indeed, using median unbiased estimation and impulse response analysis, it is evident that the MUE point estimate of the largest root is much higher than its OLS counterpart for every lag length and the half-life is closer to 6.46 years than 5 years for the dollar/pound real exchange rate (lag 2). If one considers Switzerland's demeaned real exchange rate,

[9] Notice that for the UK and Switzerland, the upper bounds for the confidence intervals of the half-lives are finite when the intervals of the largest roots contain unity as an upper limit. As pointed out by the referee, it is reasonable to expect that when there is a unit root, the half-life is infinite. The referee is absolutely right that if $\alpha = 1$, the half-life should be infinity. But computationally, one needs to impose an upper bound on the possible values over which the search is done. In the GAUSS code used in this paper, the maximum value is set equal to 40 years and if at this point the value of the LR statistic is less than the critical value, the value of the half-life is interpreted as infinity.

[10] It is worth pointing out that for all three real exchange rates, the upper bound is always very distant from the half-life point estimate while the lower bound is always relatively close to it. This asymmetry of the confidence intervals for the half-lives, or impulse responses, comes from two sources. First, the asymptotic distribution of the half-life is a function of the Dickey–Fuller distribution which is asymmetric. As a result, the confidence interval of the underlying AR parameter is asymmetric. Second, further asymmetry is generated by the nonlinearity of the impulse response function or half-life. Suppose that one has an autoregressive model of order 1 with point estimate for the AR parameter equal to 0.93. The asymmetry in the asymptotic distribution will give rise to an asymmetric confidence interval, say $[0.9, 0.98]$. Now, computing the half-life as $\ln(0.5)/\ln(\alpha)$ gives a point estimate for the half-life equal to 9.6 and an interval of $[6.6, 34.3]$.

[11] In general, both sets of confidence intervals highlight a high level of imprecision with which the half-lives are estimated. However, the fact that the monthly intervals are tighter is encouraging and demonstrates the potential to extract more information than has been previously available.

on the other hand, then Rogoff's consensus should have an infinite upper bound, not even 5 years. Thus, if one were to base one's inference on the OLS estimate, as Lothian and Taylor (1996) did, then this will inevitably underestimate the half-lives and overestimate the speed of reversion. Therefore, we re-emphasize Murray and Papell's (2005) point: the PPP puzzle is worse than you think. The higher frequency of the data does not make it better either.

V. Further Evidence from Impulse Response Analysis

As a final exercise, the study constructs the 95% MUE confidence intervals of the impulse response functions derived from the inversion of the *LR* statistic for the demeaned real exchange rate data.[12] The impulse response functions, reported in Fig. 1, are selected based on the Schwarz information criterion and they are found to pass residual tests for serial correlation. According to the point estimates of the impulse responses, the real exchange rates have zero persistence in the very long-run, confirming the existence of mean-reversion for the UK and Germany. None the less, the upper limits of the confidence intervals of the impulse response functions suggest a high degree of persistence. Taylor's main point is that, because of temporal aggregation in the data, estimates of mean-reversion are upward biased. That is if time aggregation were a significant source of bias, estimates of reversion using monthly should be systematically smaller than estimates using annual data, but that does not seem to be the case. From Fig. 1 one can see that reversion is torpid regardless of data frequency.

VI. Robustness

This section examines if the results reported in Section IV are affected by the existence of structural breaks due to different exchange rate regimes during the period examined. As recommended by the referee, the study checks whether the results remain the same after the episodes of major turmoil, i.e. the interwar period, are dropped from the sample. The sample period that is considered runs from 1939 to 2002. The results are reported in Table 2 where, to save space, the focus is on the half-life only. It is evident from this table that re-estimating the half-life for the shorter sample leads to quantitatively and qualitatively almost identical results in terms of the magnitude of this measure of persistence for Switzerland. For the UK, there is a clear reduction in the point estimates and the lower limits of the confidence intervals, although the upper limits remain nearly unchanged and again indicate a high degree of persistence in deviations from PPP. The results for Germany, on the other hand, depend on whether a linear time trend is incorporated in Equation 1. Without a trend, one can notice the significant decrease in the upper limits of the confidence intervals. At the same time, when a trend is present, the 95% confidence intervals reveal that the uncertainty associated with the speed of mean-reversion is still substantial; a result that is similar to the one for the original sample. In sum, the omission of the interwar period does not lead one to overturn the conclusions reached earlier.

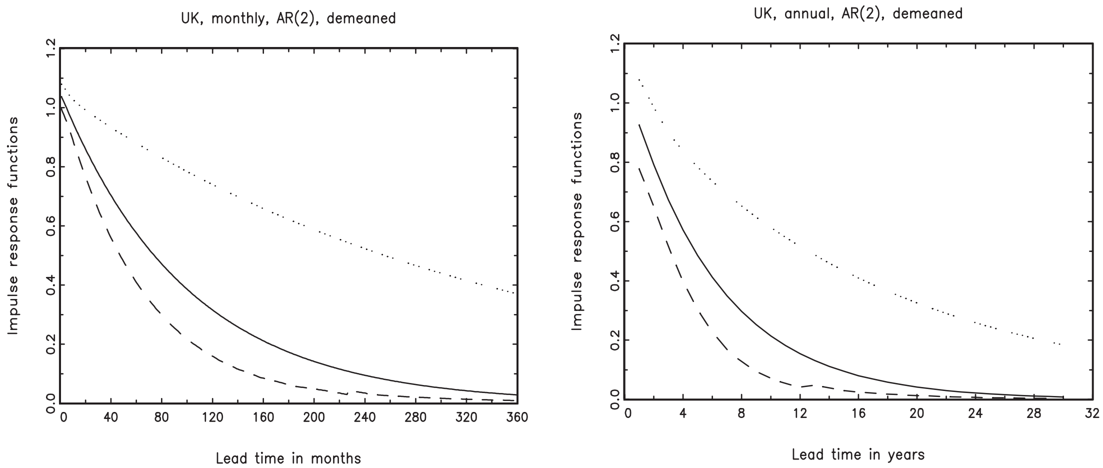
VII. Conclusion

PPP embodies the hypothesis that the nominal exchange rate and relative prices share a common trend so that the real exchange rate is a mean-reverting stationary process. However, this seemingly simple and intuitive concept has found limited empirical support in the literature. Though growing evidence in support of mean-reversion has been found, consensus estimates of the reversion speed are slow with half-lives ranging from 3 to 5 years (Rogoff, 1996). These speeds of adjustment are problematic for models with nominal stickiness which imply convergence to PPP of 1 to 2 years.

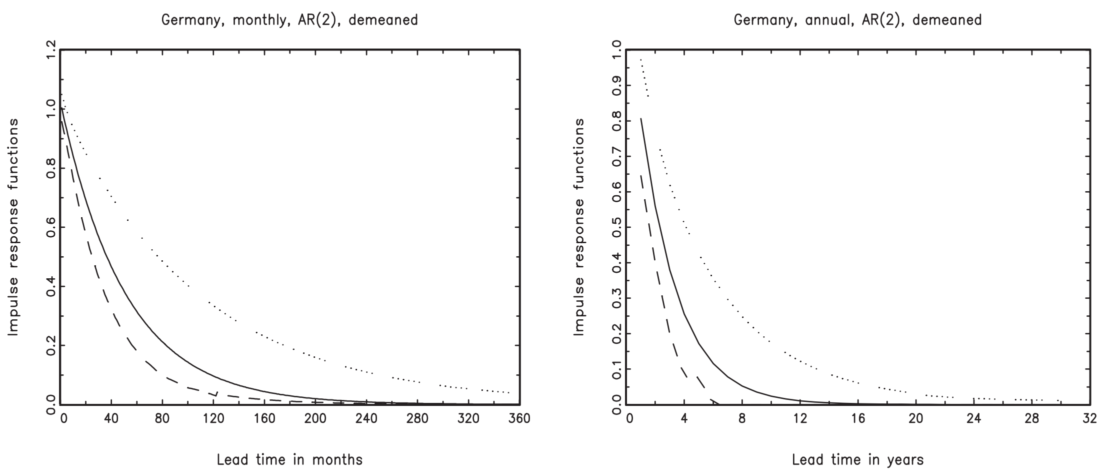
Recognizing the practical and theoretical importance of the real exchange rate, the study has examined the time-series properties of three major US real exchange rates using samples that range from 75 to 97 years. First, the expectation is that tests based on long span data series will have more power to reject a unit root than those using short samples. Second, the half-life estimates that form the basis of Rogoff's consensus have been typically obtained with long spans of annual data. Taylor (2001), however, argued that: 'sampling the data at low frequencies will never allow one to identify a high frequency adjustment process. Instead, a large bias could be introduced towards the finding of a long half-life, and the bias grows larger the greater the degree

[12] The graphs for the detrended data are not reported to preserve space. They are, however, available from the authors upon request.

UK, 1906-2002.



Germany, 1928-2002.



Switzerland, 1920-2002.

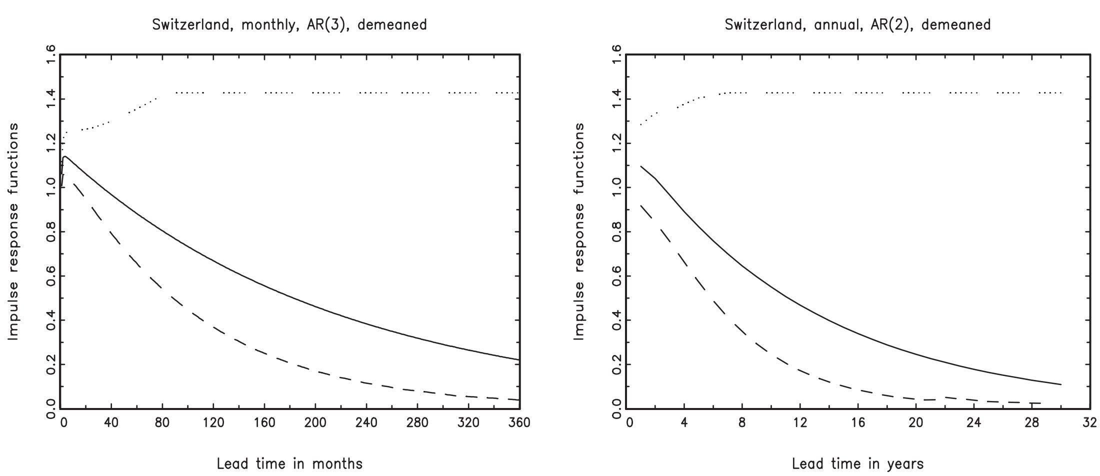


Fig. 1. Median unbiased impulse response functions estimated from the DF-GLS regressions. The unbroken line indicates the point estimates of the impulse responses. The dashed and dotted lines give the corresponding confidence intervals

Table 2. Robustness check, sample period: 1939–2002

	UK			Germany			Switzerland		
k	H.L	95_{lower}	95_{upper}	H.L	95_{lower}	95_{upper}	H.L	95_{lower}	95_{upper}
Monthly demeaned data									
2	5.2270	2.0810	26.7006	3.1935	1.6325	∞	17.5477	5.4371	22.7572
3	5.0603	2.0795	28.1084	3.0943	1.6268	∞	15.0681	5.4037	23.4236
4	5.1471	2.0601	28.4351	3.0030	1.6195	6.5572	15.2847	5.4138	23.8402
5	5.1094	2.0528	28.1075	2.9666	1.6156	6.3647	15.5304	5.4239	24.0856
6	4.9462	2.0272	28.4392	2.9442	1.6125	6.2832	15.1382	5.4182	24.2188
7	5.0310	2.0061	28.6077	2.9416	1.6055	6.3461	16.3187	5.4376	24.5896
8	5.8263	2.1059	28.4438	2.9351	1.6083	6.3286	15.3100	5.4901	24.1945
9	5.8408	2.1404	28.2733	2.9308	1.6139	6.3093	14.8869	5.4528	24.6945
10	5.4561	2.1509	28.6762	2.8957	1.6169	6.1083	14.1884	5.4695	25.1024
11	5.7413	2.0858	29.3524	2.8832	1.6160	6.0554	14.6013	5.4848	25.7505
12	5.3068	2.1554	28.7227	2.8651	1.6214	5.9327	14.0082	5.4635	23.9397
Monthly detrended data									
2	5.7554	2.0015	27.0452	2.8501	1.3747	∞	20.9940	3.0444	24.9362
3	5.5091	2.0012	28.5335	2.7445	1.3693	∞	17.6543	3.0385	25.9328
4	5.6021	1.9752	28.9470	2.6506	1.3646	∞	14.8797	3.0226	26.3497
5	5.5336	1.9655	28.6138	2.6019	1.3603	∞	15.0808	3.0221	26.5984
6	5.2728	1.9359	28.8631	2.5665	1.3556	∞	15.2353	3.0114	26.5960
7	5.3802	1.9110	29.1128	2.5469	1.3437	∞	15.2819	3.0189	27.0972
8	6.8552	2.0051	28.9505	2.5334	1.3453	∞	13.3911	3.0445	26.4181
9	6.8918	2.0437	28.6985	2.5243	1.3499	∞	13.0696	3.0567	27.4227
10	6.0789	2.0547	29.1080	2.4804	1.3520	∞	12.4616	3.0607	27.5857
11	6.6450	1.9795	29.9399	2.4606	1.3468	∞	12.6298	3.0508	28.6908
12	5.7774	2.0558	29.1741	2.4385	1.3492	∞	12.1019	3.0866	28.3480
Annual demeaned data									
2	4.3031	1.9058	22.3242	2.4023	1.2869	6.0735	∞	5.0126	∞
3	4.4661	1.7674	40.0000	2.5731	1.1971	5.1172	∞	2.6338	∞
4	4.7816	1.7138	31.7445	2.5185	1.3931	∞	∞	2.5227	∞
5	4.4770	1.7784	40.0000	2.5139	1.5897	7.9015	∞	2.6222	∞
6	4.2936	1.7668	40.0000	2.5087	1.5702	∞	∞	2.6710	∞
Annual detrended data									
2	3.9281	1.6099	26.9044	1.9028	1.4675	∞	7.3385	2.6581	∞
3	3.9453	1.4675	40.0000	2.0803	1.2435	8.6674	11.1699	1.9509	∞
4	4.2923	1.4157	27.7673	2.1422	1.1137	7.2789	8.4628	1.9057	∞
5	3.9388	1.4811	40.0000	1.9736	1.3875	∞	6.6541	1.9402	∞
6	3.7837	1.4471	40.0000	1.7763	1.3630	∞	6.7370	1.9300	∞

of temporal aggregation.' For this reason, the study has employed monthly as well as annually sampled data since with monthly data the bias is expected to be minimal. Further, it has investigated the persistence of real exchange rates through the computation of median unbiased point estimates and confidence intervals for the half-lives of deviations for DF-GLS regressions. The results indicate that although the confidence intervals for median-unbiased estimators are much tighter for monthly data than for annual data, they remain rather wide and, thus, do not help solve the PPP puzzle. Moreover, the point estimates of the half-lives are paradoxically much higher than those estimated from annual data. That is, estimating the half-lives with higher frequency data moves one even further away from resolving the puzzle. In addition, the upper limits of the confidence intervals, though not infinite as in the case of annual data, still contain long half-lives. Still, this can be viewed as one benefit of using monthly data. The lower bounds can be low, but are still not less than 2 years except for Germany. The confidence intervals from annual data, on the other hand, are so much wider as to be of very little use. On the whole, given the high level of imprecision with the half-life, great caution should thus be taken in making inferences based on the point estimates alone (Cheung and Lai, 2000).

Recently, Murray and Papell (2002) and Rossi (2004) have used different methodologies to construct confidence intervals for the half-lives using post-Bretton Woods data. The lower bounds of the confidence intervals reported in these papers are mostly less than 2 years, while the upper bounds are generally infinite. Whereas the lower bounds are consistent with the theory of PPP, the upper bounds are consistent with no convergence to long-run PPP. The present results are slightly different. The upper bounds are still high (though not always infinite) but the lower bounds are much higher than 2 years particularly when monthly data are used, which incidentally allows one to obtain sharper inference. Hence, in line with Murray and Papell (2005), it appears that the PPP puzzle is even more problematic.

Finally, besides temporal aggregation bias, Imbs *et al*. (2002) demonstrate that cross-sectional aggregation bias raises the persistence of real exchange rate shocks. Cross-sectional aggregation bias arises from the failure to take account of cross-sectoral heterogeneity in the dynamic properties of the typical components of aggregate price indices. This failure to allow for the persistence of relative prices to vary across sectors induces an upward bias in aggregate half-life measures, with the bias rising with the extent of cross-sectoral heterogeneity in the speed of parity reversion. Engel and Chen (2004), on the other hand, use monthly relative prices of goods at the sectoral level for 16 categories of goods for the USA and nine European countries and find that sector heterogeneity is not a quantitatively important source of bias. A common feature of the two studies is that, due to missing observations and data availability, both focus on short samples; 1981 to 1995 for the former and 1981 to 1996 for the latter. This unavailability of data for different sectors of the economy spanning long periods, particularly the last century, means that this study has focused on measuring the speed of mean-reversion of the aggregate real exchange rate and not disaggregated relative prices. Nonetheless, constructing and collecting long span data for different sectors and at different frequencies is, in the authors' opinion, an exciting area for future research since it will allow them to link a number of factors that have been proposed in the literature as reasons for the slow reversion of the real exchange rate towards equilibrium.

Acknowledgments

We thank Lucio Sarno, Nikolay Gospodinov, Karim Ait-Chabane and an anonymous referee for helpful comments and feedback. The usual disclaimer applies.

References

Abuaf, N. and Jorion, P. (1990) Purchasing Power Parity in the long-run, *Journal of Finance*, **45**, 157–74.

Andrews, D. W. K. and Chen, H. Y. (1994) Approximately median-unbiased estimation of autoregressive models, *Journal of Business and Economic Statistics*, **12**, 187–204.

Basawa, I. V., Mallik, A. K., McCormick, W. P., Reeves, J. H. and Taylor, R. L. (1991) Bootstrapping unstable first-order autoregressive processes, *The Annals of Statistics*, **19**, 1098–101.

Baun, C. F., Barkoulas, J. J. and Caglayan, M. (2001) Nonlinear adjustment to Purchasing Power Parity in the post-Bretton Woods era, *Journal of International Money and Finance*, **20**, 379–99.

Cheung, Y. W. and Lai, K. S. (2000) On the Purchasing Power Parity puzzle, *Journal of International Economics*, **52**, 321–30.

Cheung, Y. W., Lai, K. S. and Bergman, M. (2004) Dissecting the PPP puzzle: the unconventional roles of nominal exchange rate and price adjustments, *Journal of International Economics*, **64**, 135–50.

Diebold, F. X., Husted, S. and Rush, M. (1991) Real exchange rates under the Gold Standard, *Journal of Political Economy*, **99**, 1252–71.

Dornbusch, R. (1976) Expectations and exchange rate dynamics, *Journal of Political Economy*, **84**, 1161–76.

Dumas, B. (1992) Dynamic equilibrium and the real exchange rate in spatially separated world, *Review of Financial Studies*, **5**, 153–80.

Elliott, G., Rothenberg, T. and Stock, J. H. (1996) Efficient tests for an autoregressive unit root, *Econometrica*, **64**, 813–36.

Engel, C. and Chen, S.-S. (2004) Does aggregation bias explain the PPP puzzle?, *Pacific Economic Review*, **10**, 49–72.

Frankel, J. A. (1990) Zen and the art of modern macroeconomics: the search for perfect nothingness, in *Monetary Policy for a Volatile Global Economy* (Eds) W. Haraf and T. Willet, American Enterprise Institute, Washington, DC.

Gospodinov, N. (2004) Asymptotic confidence intervals for impulse responses of near-integrated processes, *Econometrics Journal*, **7**, 505–27.

Hansen, B. (1999) The grid bootstrap and the autoregressive model, *The Review of Economics and Statistics*, **81**, 594–607.

Imbs, J., Mumtaz, H., Ravn, K. and Rey, H. (2002) PPP strikes back: aggregation and the real exchange rate, NBER Working Paper No. 9372.

Kilian, L. and Zha, T. (2002) Quantifying the uncertainty about the half-life of deviations from PPP, *Journal of Applied Econometrics*, **17**, 107–25.

Lopez, C., Murray, C. J. and Papell, D. H. (2005) State of the art unit root tests and the PPP puzzle, *Journal of Money, Credit and Banking*, **37**, 361–9.

Lothian, J. R. and Taylor, M. P. (1996) Real exchange rate behavior: the recent float from the perspective of the last two centuries, *Journal of Political Economy*, **104**, 488–510.

Lothian, J. R. and Taylor, M. P. (2000) Purchasing Power Parity over two centuries: strengthening the case for real exchange rate stability: a reply to Cuddington and Liang, *Journal of International Money and Finance*, **19**, 759–64.

Michael, P., Nobay, A. R. and Peel, D. A. (1997) Transaction costs and nonlinear adjustment in real exchange rates: an empirical investigation, *Journal of Political Economy*, **105**, 862–79.

Murray, C. J. and Papell, D. H. (2002) The Purchasing Power Parity persistence paradigm, *Journal of International Economics*, **56**, 1–19.

Murray, C. J. and Papell, D. H. (2004) Do panels help solve the Purchasing Power Parity puzzle? Working Paper, University of Houston.

Murray, C. J. and Papell, D. H. (2005) The Purchasing Power Parity puzzle is worse than you think, *Empirical Economics*, **30**, 783–90.

Ng, S. and Perron, P. (2001) Lag length selection and the construction of unit root tests with good size and power, *Econometrica*, **69**, 1519–54.

Obstfeld, M. (1993) Model trending real exchange rates, Working Paper No. C93-011, Center for International and Development Economics Research.

Pierse, R. G. and Snell, A. J. (1995) Temporal aggregation and the power of tests for a unit root, *Journal of Econometrics*, **65**, 333–45.

Perron, P. (1989) Testing for a random walk: a simulation experiment of power when the sampling interval is varied, in *Advances in Econometrics* (Ed.) B. Raj, Kluwer Academic Publishers, Dordrecht, pp. 47–68.

Rapach, D. E. and Wohar, M. E. (2002) Testing the monetary model of exchange rate determination: new evidence from a century of data, *Journal of International Economics*, **58**, 359–85.

Rogoff, K. (1996) The Purchasing Power Parity puzzle, *Journal of Economic Literature*, **34**, 647–68.

Rossi, B. (2005) Confidence intervals for half-life deviations from Purchasing Power Parity, *Journal of Business and Economic Statistics*, **23**, 432–42.

Sarno, L. and Taylor, M. P. (2003) *Economics of Exchange Rates*, Cambridge University Press, Cambridge.

Shiller, R. J. and Perron, P. (1985) Testing the random walk hypothesis: power versus frequency of observation, *Economics Letters*, **18**, 381–6.

Taylor, A. M. (2001) Potential pitfalls for the Purchasing Power Parity puzzle? Sampling and specification biases in mean-reversion tests of the law of one price, *Econometrica*, **69**, 473–98.

Taylor, A. M. (2002) A century of Purchasing Power Parity, *The Review of Economics and Statistics*, **84**, 139–50.

Taylor, M. P., Peel, D. A. and Sarno, L. (2001) Nonlinear mean-reversion in real exchange rates: towards a solution to the Purchasing Power Parity puzzles, *International Economic Review*, **42**, 1015–42.