

Real Exchange Rate Forecasting and PPP: This Time the Random Walk Loses

Michele Ca' Zorzi¹ · Jakub Muck^{2,3} ·
Michal Rubaszek^{2,3}

Published online: 4 January 2016
© Springer Science+Business Media New York 2016

Abstract This paper brings four new insights into the Purchasing Power Parity (PPP) debate. First, we show that a half-life PPP (HL) model is able to forecast real exchange rates better than the random walk (RW) model at both short and long-term horizons. Second, we find that this result holds if the speed of adjustment to the sample mean is calibrated at reasonable values rather than estimated. Third, we find that it is preferable to calibrate, rather than to elicit as a prior, the parameter determining the speed of adjustment to PPP. Fourth, for most currencies in our sample, the HL model outperforms the RW also in terms of nominal exchange rate forecasting.

Keywords Exchange rate forecasting · Purchasing power parity · Half-life

We have benefited from helpful comments by G. Amisano, G. Benigno, L. Dedola, J. Nosal, L. Sarno and an anonymous referee. The paper has been presented at internal seminars of the ECB and NBP and at the MMF (London 2013), Macromodels (Warsaw 2013) and IWH-CIREQ Macroeconometric Workshop (Halle 2013) conferences. The views in this paper are those of the authors and do not necessarily reflect the views of the European Central Bank and the National Bank of Poland. All remaining errors are the sole responsibility of the authors.

✉ Jakub Muck
jakub.muck@nbp.pl

Michele Ca' Zorzi
michele.cazorzi@ecb.europa.eu

Michal Rubaszek
michal.rubaszek@nbp.pl

¹ European Central Bank, Frankfurt, Germany

² National Bank of Poland, Warsaw, Poland

³ Warsaw School of Economics, Warsaw, Poland

JEL Classification C32 · F31 · F37

1 Introduction

Exchange rates have long fascinated, challenged and puzzled researchers in international finance. Since the seminal papers by Meese and Rogoff (1983a, b), there has been wide agreement that macroeconomic models are not very helpful for exchange rate forecasting.¹ The exchange rate literature provides, however, at least two reasons for cautious optimism.

First, the dismal forecasting performance of exchange rate models can to some extent be explained by estimation and not only misspecification error (Engel et al. 2008). The significant role of estimation error is confirmed, among other things, by the relatively good forecasting performance of economic models estimated with a large panels of data (Mark and Sul 2001; Engel et al. 2008; Ince 2014) or long time series (Lothian and Taylor 1996). A number of tests have also been developed to take into account the sampling variance of estimated models, which would unduly give the RW a head start in a forecasting horse race (see Engel 2013, for a review).

The second reason for being cautiously optimistic about the usefulness of exchange rate models comes from the evidence in favor of the PPP model. According to Taylor and Taylor (2004), the exchange rate literature has turned full circle to the pre-1970s view that PPP holds in the long run. The mean reverting nature of real exchange rates has in particular found some support by panel unit root tests, which have higher power than the conventional univariate tests in the case of highly persistent or non-linear processes (Sarno and Taylor 2002; Holmes et al. 2012). Only a handful of studies have instead tested more directly whether the mean-reverting properties of the real exchange rate can be exploited in a forecasting setting. In the late 1980s, Meese and Rogoff (1988) extended their classic analysis to reach the conclusion that, like nominal exchange rates, real exchange rates are disconnected from economic fundamentals. Two studies in the mid-1990s argued instead that the RW can be beaten for larger datasets, for example in the case of long-data series (Lothian and Taylor 1996) or in multivariate frameworks (Jorion and Sweeney 1996). More recently, Rogoff (2009) suggested that is worth investigating whether PPP deviations and current account positions may help predict real exchange rate movements.

In this paper we aim to establish if Rogoff's insight on PPP is correct. We will show that it is not enough to assume that real exchange rates revert to their mean, but it is also necessary to assume that the reversion pace is rather slow. This was first established in a series of studies conducted between the mid-1980s and early 1990s, which employed more than a hundred years of annual data. From an informal meta-analysis of these studies, Rogoff (1996) inferred that it takes between 3 and 5 years to halve real exchange rate deviation from the mean. More recent studies

¹ In the mid-1990s Mark (1995) and Chinn and Meese (1995) suggested that the RW model could be beaten at longer horizons. This more optimistic perspective was however short-lived and failed to overturn the previous consensus (Faust et al. 2003, Cheung et al. 2005 or Rogoff 2009).

have been more skeptical about what is typically dubbed as “Rogoff consensus”. For example, Kilian and Zha (2002) proposed a prior probability distribution based on a survey of professional international economists and derived a posterior probability distribution of the half-life on the basis of a Bayesian autoregressive model. Their results provide very limited support for the view that the half-life is between three and five years. In a similar vein Murray and Papell (2002) stressed how univariate methods provide virtually no information on the size of the half-lives. Finally, a large cross-country heterogeneity in terms of point estimates and confidence intervals has also been found in the studies by Murray and Papell (2005) and Rossi (2005). There is however another strand of the literature which finds it instead plausible that at the aggregate level half-lives would be in the range between 3 and 5 years. They argue that aggregation bias both at the time and product dimension helps reconcile this high duration of the adjustment process with faster convergence at the product and sectoral level (Imbs et al. 2005; Crucini and Shintani 2008; Mayoral and Dolores Gadea 2011; Bergin et al. 2013).

In this paper we analyse whether mean-reversion of real exchange rates can be exploited for forecasting purposes. We test whether a calibrated, half-life PPP (henceforth, HL) model, which imposes a very gradual linear adjustment of the real exchange rate toward its mean, is able to forecast real exchange rates better than would be the case with (i) the same autoregressive (henceforth, AR) model, where the pace of mean-reversion is instead estimated rather than imposed and with (ii) a RW model. In our baseline the HL model is calibrated so that half of the adjustment of the real exchange rate toward its mean is completed within five years. We have set this initial value, rather conservatively, i.e. at the top of the range proposed by Rogoff (1996), so that in the short-run the model predictions resemble quite closely those of the RW model. As Rogoff’s consensus was built on the basis of the data from the pre-1990s and our forecast evaluation sample starts in 1990, this choice is not influenced by the data that are used to assess the quality of forecasts. We shall later extend our analysis with a thorough sensitivity analysis to show that, contrary to our initial expectations, all the key results hold true for the whole range of half-lives between 3 and 5 years proposed by Rogoff.

The key findings of our paper are as follows. We show that, exploiting simultaneously the evidence of (i) real exchange rate persistence and (ii) long-term convergence to PPP, leads to a considerable improvement in our ability to forecast real exchange rates even for short samples. To be more specific, we show that the HL model is able to forecast real exchange rates better than the RW for seven out of nine currencies. Particularly persuasive is that this simple approach beats the RW also at short-horizons.

Another remarkable result of our study is that the forecast accuracy of the estimated AR model is clearly considerably worse than that of the RW. We will explain this result both analytically and empirically, emphasizing that this is to be ascribed to the large impact of the estimation error, even if we have as many as 15 years of monthly data. Our empirical investigation is then taken a final step forward: we find that the mean reverting nature of the real exchange rate can also be exploited to outperform the RW model also for forecasting nominal effective exchange rates. This is because, for the majority of the currencies in our sample, real exchange rates revert

to the mean mainly via nominal exchange rate rather than relative price adjustments. Finally, we will show that the results also hold in bilateral and not just effective terms.

The reason for the success of the HL model can be explained by Diebold's shrinkage principle, which asserts that imposing restrictions on model parameters may be helpful for out of sample forecasts, even if the restrictions are false. There is also an interesting parallelism between our work and that of Faust and Wright (2013) on inflation forecasting. These authors show that a minimalistic two-steps forecasting strategy proves to be very effective out of sample. The first step requires finding a good long-term forecast, which could be underpinned by economic theory, survey data, etc. The second step consists of connecting the current and long-term value of inflation with a smooth path, giving up any attempt to explain a more complex dynamic of adjustment. Our proposed HL model is an exact application of this strategy, as we postulate a linear adjustment between the current level of the real exchange rate and a good long-term forecast, i.e. PPP.

The rest of the paper is structured as follows. Sections 2 and 3 outline the alternative models that we use in the real exchange rate forecasting horse race and present the outcome of the competition. Sections 4 and 5 provide an analytical investigation that sheds some light on our findings and contain a more general discussion on why the calibrated HL model is so competitive in terms of real exchange rate forecasting. Section 6 shows that the results are very robust to several alternative specifications. We also illustrate that our key results are valid for a broad range of half-lives (i.e. even wider than the range proposed by Rogoff). We conclude by showing that our improved ability to forecast real exchange rates is helpful also in the context of nominal exchange rate forecasting. The results are also valid for bilateral exchange rates of eight currencies vis-à-vis the US dollar. The last section concludes.

2 The Models

Let us define the log of the real exchange rate (RER) according to the convention that $y_t \equiv s_t + p_t - p_t^*$, where s_t is the log of the nominal exchange rate expressed as the foreign currency price of a unit of domestic currency, and p_t and p_t^* are the logs of home and foreign price levels, respectively.

Consider a simple autoregression model for the real exchange rate:

$$y_t - \mu = \rho(y_{t-1} - \mu) + \epsilon_t, \epsilon_t \sim \mathcal{N}(0, \sigma^2). \quad (1)$$

For a stationary AR process the parameter ρ measures the speed of reversion to μ , which we interpret as the level of PPP. The half-life of adjustment toward PPP is equal to:

$$hl = \log(0.5) / \log(\rho). \quad (2)$$

For $\rho = 1$ the RER is generated by the RW process.

In the forecasting contest we employ three alternatives of model (1).

1. The first is a RW model, for which we calibrate $\rho = 1$ and $\mu = 0$. The h step ahead forecast is:

$$y_{T+h|T}^{RW} = y_T. \quad (3)$$

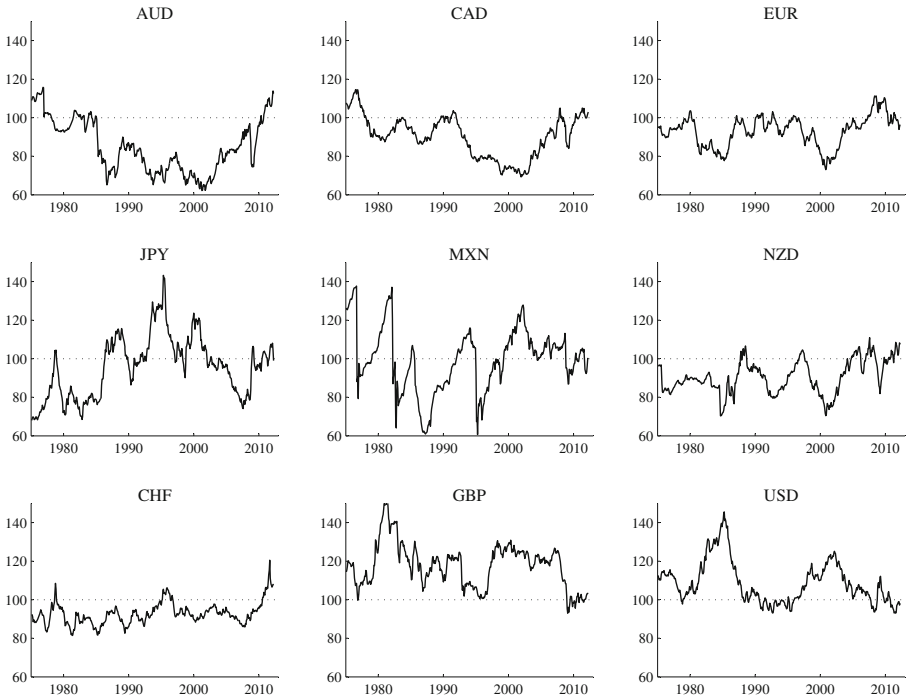


Fig. 1 Real exchange rates (2010 = 100)

2. The second is the HL model for which we assume that the real exchange rate gradually converges to its sample mean ($\bar{\mu}$). The pace of convergence ($\bar{\rho}$) is calibrated with Eq. 2 so that the half-life is equal to five years, i.e. at the top of the range proposed by Rogoff.² The h step ahead forecast is:

$$y_{T+h|T}^{HL} = \bar{\mu} + \bar{\rho}^h (y_T - \bar{\mu}). \tag{4}$$

3. The third is an autoregressive AR model, whose two parameters are estimated with OLS $\hat{\mu}$ and $\hat{\rho}$, so that the h step ahead forecast is:

$$y_{T+h|T}^{AR} = \hat{\mu} + \hat{\rho}^h (y_T - \hat{\mu}). \tag{5}$$

3 Empirical Evidence

To assess the predictability of real exchange rates we gather monthly data for nine major currencies of the following economies: Australia (AUD), Canada (CAD), euro area (EUR), Japan (JPY), Mexico (MXN), New Zealand (NZD), Switzerland (CHF), the United Kingdom (GBP) and the United States (USD) for the period between

²In the sensitivity analysis we will check the forecasting accuracy of different calibrations for ($\bar{\rho}$)

1975:1 and 2012:3. For all currencies we take the real effective exchange rates provided by the Bank for International Settlements (Klau and Fung 2006). The values of the analyzed series are presented in Fig. 1.

We assess the out-of-sample forecast performance for horizons ranging from one to sixty months ahead. In our baseline specification the models are estimated using rolling samples of 15 years ($R = 180$ months). The first set of forecasts is elaborated with the rolling sample 1975:1-1989:12 for the period 1990:1-1994:12. This procedure is repeated with rolling samples ending in each month from the period 1990:2-2012:2. Since the data available end in 2012:3, the 1-month-ahead forecasts are evaluated on the basis of 267 observations, 2-month-ahead forecasts on the basis of 266 observations, and 60-month-ahead forecasts on the basis of 208 observations.

We measure the forecasting performance of the three competing models with three statistics: the mean squared forecast errors (MSFEs), the correlation coefficient between forecast and realized real exchange rate changes and the frequency of correct predictions in terms of directional change. Table 1 and Fig. 2 present the values

Table 1 Mean squared forecast errors (15Y rolling window)

h	RW	HL	AR	RW	HL	AR	RW	HL	AR
	AUD			CAD			EUR		
1	0.05	1.01	1.02*	0.02	1.02	1.03*	0.02	1.00	1.04**
6	0.44	1.03	1.06*	0.23	1.03	1.11**	0.19	0.96	1.12*
12	0.82	1.06	1.09*	0.46	1.06	1.20**	0.43	0.92*	1.14*
24	1.53	1.10*	1.09	0.94	1.10	1.19*	0.89	0.83**	1.18**
36	2.10	1.12	1.06	1.58	1.06	1.20*	1.28	0.77**	1.13*
60	3.00	1.11	1.06	3.02	0.94	1.45**	2.06	0.66**	0.91
	JPY			MXN			NZD		
1	0.06	1.00	1.01	0.12	0.99	0.99	0.03	1.00	1.03*
6	0.59	0.98	1.04	0.83	0.96*	0.92	0.32	0.96	1.06
12	1.00	0.97	1.10*	1.55	0.92*	0.87*	0.72	0.92*	1.03
24	2.34	0.91	1.17**	3.01	0.84**	0.78*	1.59	0.83**	0.89*
36	3.53	0.86	1.19**	3.66	0.78**	0.74**	2.44	0.74**	0.74**
60	3.12	0.89	1.19*	3.56	0.74**	0.72*	3.01	0.64**	0.62**
	CHF			GBP			USD		
1	0.02	1.00	1.06	0.03	1.00	1.02	0.02	1.00	1.03*
6	0.12	0.98	1.22*	0.24	0.97	1.04	0.19	0.96	1.10*
12	0.25	0.97	1.16	0.47	0.95	1.03	0.31	0.93	1.21**
24	0.50	0.88*	0.99	1.06	0.87*	0.98	0.55	0.84*	1.21*
36	0.72	0.80**	0.78**	1.49	0.82**	0.93	0.69	0.72**	1.13
60	0.79	0.72**	0.69**	1.98	0.67**	0.70**	1.41	0.53**	0.91

For the RW model MSFEs are reported in levels (multiplied by 100), whereas for the remaining methods they appear as the ratios to the corresponding MSFE from the RW model. Asterisks ** and * denote the rejection of the null of the Diebold and Mariano (1995) test, stating that the MSFE from RW are not significantly different from the MSFE of a given model, at 1 %, 5 % significance level, respectively

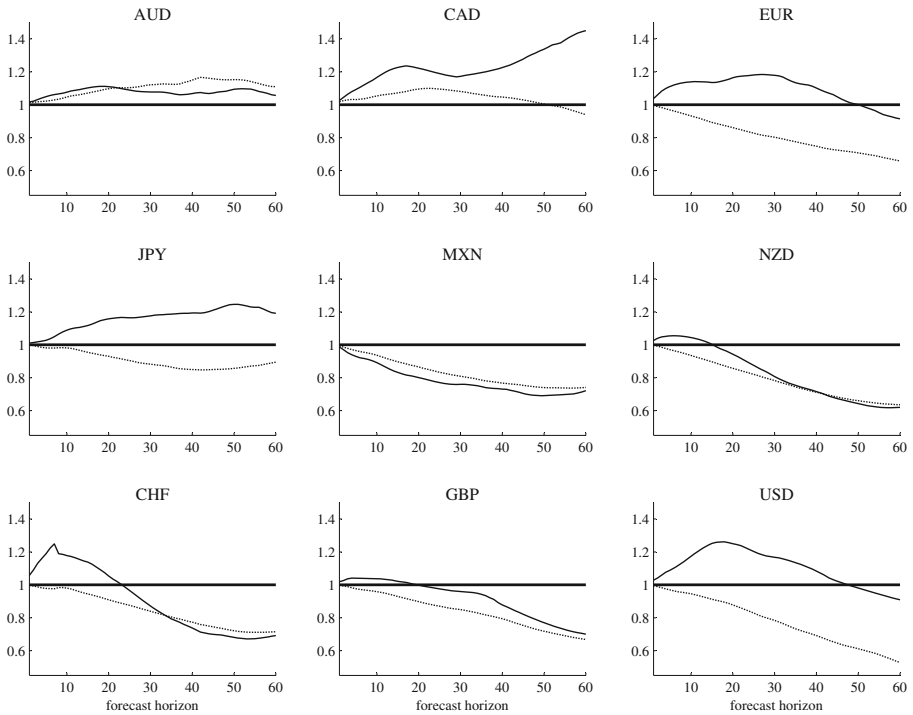


Fig. 2 Mean squared forecast errors. Notes: Each *line* represents the ratio of MSFE from a given method to MSFE from the random walk, where values below unity indicate better accuracy of point forecasts. The *straight* and *dotted lines* stand for AR and HL5, respectively. The forecast horizon is expressed in months

of MSFEs. For the RW we report the MSFEs in levels. For the HL and AR models, we report them divided by MSFEs of the RW, so that values below unity indicate that such model outperforms the RW. We also test the null of equal forecast accuracy with the two-sided Diebold and Mariano (1995) test (corrected for heteroscedasticity and autocorrelation) to assess whether there is significant evidence that the model under or over-performs compared to the RW.³

In terms of the MSFE criterion the HL model-based forecasts beat the RW for seven out of nine currencies (EUR, MXN, NZD, CHF, GBP, USD, JPY). The MSFEs of the HL model are on average 9 and 23 % lower than that of the RW model at two and five-year horizon, respectively. The HL model-based forecasts are also

³Diebold (2012) states that if the aim is to test the null of equal forecast accuracy between two models the best choice is to use the Diebold Mariano (DM) statistic, especially as for rolling forecasting schemes its distribution is asymptotically normal. We however also calculated bootstrapped DM test *p*-values using the algorithm of Kilian (1999) to evaluate whether it is possible to reject the null that RW is the true DGP. The results, available upon request, suggest that at five year horizon the null against the alternative of the HL model is rejected for six out of nine currencies at the 5 % significance level. The Clark and West test cannot be used to evaluate RW against HL because they are not-nested models.

considerably more precise than those based on the AR model for five currencies (CAD, EUR, JPY, GBP and USD) while are broadly comparable for the other four. At short-horizons the HL model is clearly the best whereas the AR the worst one. At the one-year horizon the MSFEs of the HL model are on average 3 and 12 % lower than those from the RW and AR models, respectively.

Further evidence that the HL model beats the other two models can be found using our second criterion, which consists in computing the correlation coefficient between the realized and forecast changes of real exchange rates:

$$r_{M,h} = cor \left(y_{T+h|T}^M - y_T, y_{T+h} - y_T \right), \tag{6}$$

where M stands for the model name. Note that Eq. 3 implies that $r_{RW,h}$ is zero: for that reason in Table 2 we report only the results for the HL and AR models. It shows that the correlation coefficients for the HL model are generally positive for all currencies at all horizons, except for the AUD. The average value of $r_{HL,h}$ also increases with the forecast horizon: from just 0.04 for one-month ahead forecasts to 0.53 for five-year ahead forecasts. The results do not provide support instead for the AR model: MXN is the only currency with a positive $r_{AR,h}$ throughout the forecast horizon. Moreover, the average value of $r_{AR,h}$ is positive only for horizons above two years. Finally, at all horizons $r_{AR,h}$ is visibly lower than $r_{HL,h}$.

We also compare how well investigated models predict the sign of change in real exchange rates. Table 3 reports the frequency of correct direction of change in the real exchange rate predicted by the HL and AR models. The RW by construction is

Table 2 Correlation of forecast and realized changes of real exchange rates

h	AUD	CAD	EUR	JPY	MXN	NZD	CHF	GBP	USD	Mean
HL MODEL										
1	-0.04	-0.03	0.06	0.04	0.11	0.06	0.08	0.06	0.07	0.04
6	-0.01	0.02	0.21	0.15	0.30	0.23	0.21	0.19	0.22	0.17
12	-0.05	0.01	0.32	0.18	0.42	0.36	0.25	0.26	0.29	0.23
24	-0.10	0.02	0.46	0.31	0.60	0.56	0.50	0.43	0.43	0.36
36	-0.12	0.11	0.55	0.38	0.66	0.73	0.66	0.53	0.56	0.45
60	-0.20	0.26	0.71	0.31	0.61	0.81	0.74	0.71	0.78	0.53
AR MODEL										
1	-0.04	-0.09	-0.19	-0.04	0.13	-0.07	-0.13	-0.03	-0.09	-0.06
6	-0.07	-0.22	-0.29	-0.10	0.33	-0.05	-0.21	0.02	-0.17	-0.08
12	-0.06	-0.31	-0.24	-0.22	0.44	0.05	-0.04	0.08	-0.24	-0.06
24	0.00	-0.14	-0.22	-0.33	0.60	0.35	0.22	0.19	-0.13	0.06
36	0.07	-0.09	-0.06	-0.32	0.64	0.61	0.52	0.26	0.05	0.19
60	0.10	-0.43	0.32	0.04	0.63	0.78	0.64	0.58	0.34	0.33

Table 3 Frequency of correct sign forecast for the real exchange rates

<i>h</i>	AUD	CAD	EUR	JPY	MXN	NZD	CHF	GBP	USD	Mean
HL MODEL										
1	0.48	0.47	0.52	0.48	0.49	0.50	0.50*	0.53	0.49	0.50
6	0.40**	0.44*	0.49	0.47	0.52**	0.52*	0.48*	0.57	0.55	0.49
12	0.37**	0.42**	0.55*	0.45*	0.61**	0.55*	0.55*	0.61**	0.59**	0.52
24	0.37**	0.43*	0.57*	0.50*	0.61**	0.66**	0.66**	0.64**	0.67**	0.57
36	0.37**	0.43*	0.56*	0.53*	0.67**	0.75**	0.73**	0.69**	0.63**	0.59
60	0.45**	0.50*	0.67**	0.61**	0.75**	0.80**	0.71**	0.80**	0.78**	0.67
AR MODEL										
1	0.49	0.49	0.45	0.44	0.48	0.48	0.50*	0.49	0.45	0.48
6	0.42**	0.44**	0.39**	0.37**	0.52	0.47*	0.50*	0.48	0.42**	0.45
12	0.44*	0.39**	0.39**	0.34**	0.62**	0.51*	0.55*	0.54	0.42	0.47
24	0.48*	0.50*	0.39**	0.29**	0.64**	0.62**	0.63**	0.54	0.48	0.51
36	0.47*	0.52*	0.42**	0.34**	0.69**	0.70**	0.70**	0.59	0.46	0.54
60	0.57*	0.40**	0.57**	0.37*	0.69**	0.77**	0.75**	0.76**	0.57	0.61

The figures represents the fraction of forecasts that correctly predict the sign of the change in the real exchange rates. Asterisks ** and * denote the rejection of the null of the Pesaran and Timmermann (1992) χ^2 test, stating that the sign of forecasts and realizations are independent, at 1 %, 5 % significance level, respectively

agnostic about directional change. The results reveal that at short horizons both the AR and HL models are unable to capture directional change systematically better than with a simple toss of the coin.⁴ At medium term horizons, however, the HL model anticipates the direction of change in the real exchange rate between 60 and 80 % of times for seven out of nine currencies, which is somewhat better than in the case of the AR model. Using the Pesaran and Timmermann (1992) Chi-squared test we also establish that at five-year horizon the null of independence between the projected real exchange rate changes (on the basis of the HL model) and the realized changes is rejected at the 1 % significance level for all currencies except the Canadian dollar.

To sum up, the evidence suggests that real exchange rates of major currencies tend to be mean reverting and forecastable, as shown by the good performance of the calibrated HL model. At short horizons the HL model performs well compared to the AR model and slightly better than the RW. Over medium term forecasting horizons, however, the HL model strongly outperforms the RW exactly because it captures the tendency of the real exchange rate to be mean reverting. In the next section we provide an analytical explanation of why the estimated AR model instead performs so poorly.

⁴Pippenger and Goering (1998) present a similar application of this criterion to exchange rate forecasts generated by non-linear models.

4 Analytical Interpretation of the Results

In what follows we show analytically why the finite sample determines a sizable estimation error, which distorts the results in favor of the RW model even when the rolling estimation window covers several years of monthly data. Let us assume that the data generating process (DGP) for y_t is given by Eq. 1 so that the unbiased and efficient forecast is:

$$y_{T+h|T} = \mu + \rho^h(y_T - \mu), \tag{7}$$

and the variance of the forecast error:

$$E \left\{ (y_{T+h} - y_{T+h|T})^2 \right\} = \sigma^2 \frac{1 - \rho^{2h}}{1 - \rho^2}. \tag{8}$$

If the DGP is known the only source of forecast errors comes from the random term. The variance of forecast errors generated by our three competing models is however higher than that in Eq. 8 because the coefficients μ and ρ are unknown and have to be estimated or calibrated.

Let us decompose the variance of the forecast error from a generic model $M \in \{RW, HL, AR\}$ into three components:

$$E \left\{ (y_{T+h} - y_{T+h|T}^M)^2 \right\} = E \left\{ (y_{T+h} - y_{T+h|T})^2 \right\} + E \left\{ (y_{T+h|T} - y_{T+h|T}^M)^2 \right\} + 2E \left\{ (y_{T+h} - y_{T+h|T}) (y_{T+h|T} - y_{T+h|T}^M) \right\}. \tag{9}$$

The first component, which is given by Eq. 8, represents the random error that is common to all models. The second component captures the error due to model misspecification. The third component is equal to zero as future shocks are not forecastable. The key component that determines the different performance of our three competing models is therefore the second one. It is particularly advantageous that we can derive such component analytically for all three models.

In the case of the RW model the forecast error equals:

$$y_{T+h|T} - y_{T+h|T}^{RW} = (\rho^h - 1)(y_T - \mu) \tag{10}$$

and thus:

$$E \left\{ (y_{T+h|T} - y_{T+h|T}^{RW})^2 \right\} = (\rho^h - 1)^2 \times E \left\{ (y_T - \mu)^2 \right\}, \tag{11}$$

where:

$$E \left\{ (y_T - \mu)^2 \right\} = \frac{\sigma^2}{1 - \rho^2}.$$

For the HL model, such error is equal instead to:

$$y_{T+h|T} - y_{T+h|T}^{HL} = (\rho^h - \bar{\rho}^h)(y_T - \mu) - (1 - \bar{\rho}^h)(\bar{\mu} - \mu). \tag{12}$$

The first term describes the forecast error caused by the wrong calibration of parameter ρ and the second one is the error related to the estimation of μ . The

resulting variance is:

$$\begin{aligned}
 E \left\{ \left(y_{T+h|T} - y_{T+h|T}^{HL} \right)^2 \right\} &= (\rho^h - \bar{\rho}^h)^2 \times E \{ (y_T - \mu)^2 \} + (1 - \bar{\rho}^h)^2 \\
 &\quad \times E \{ (\bar{\mu} - \mu)^2 \} - 2 (\rho^h - \bar{\rho}^h) (1 - \bar{\rho}^h) \\
 &\quad \times E \{ (y_T - \mu) (\bar{\mu} - \mu) \},
 \end{aligned} \tag{13}$$

where:

$$\begin{aligned}
 E \{ (\bar{\mu} - \mu)^2 \} &= \frac{\sigma^2}{1 - \rho^2} \times \frac{1}{R^2} \times \left(R + 2 \sum_{j=1}^{R-1} (R - j) \rho^j \right) \\
 E \{ (y_T - \mu) (\bar{\mu} - \mu) \} &= \frac{\sigma^2}{1 - \rho^2} \times \frac{1}{R} \times \frac{1 - \rho^R}{1 - \rho}.
 \end{aligned}$$

Finally, as derived in Fuller and Hasza (1980), for the AR model the second component is approximately equal to:

$$E \left\{ \left(y_{T+h|T} - y_{T+h|T}^{AR} \right)^2 \right\} \simeq \sigma^2 \times \frac{1}{R} \times \left[h^2 \rho^{2(h-1)} + \left(\frac{1 - \rho^h}{1 - \rho} \right)^2 \right] \tag{14}$$

and is entirely caused by the estimation error.

Given Eqs. 8–14, the assumptions for the DGP coefficients (μ , ρ and σ) and the sample size (R), one can calculate the theoretical value of MSFE for all competing models (RW, HL and AR) at different forecast horizons ($h = 1, 2, \dots, H$). The theoretical MSFEs of all models do not depend on the value of μ and are proportional to the value of σ^2 . The relative MSFEs depend hence only on the convergence coefficient ρ , the sample size R and the forecast horizon h .

Let us now consider values of ρ corresponding to DGPs where the underlying half-life parameter varies from one to ten years. We also postulate the same sample size and forecast horizons as in the empirical application from Section 3. The results are presented in Fig. 3, where the theoretical MSFEs of a given model are shown as a ratio with respect to the RW model.

The analytical results depend on the half-lives of the underlying DGP process. For half-lives above one year, the HL model beats the AR model; for values below 10 years it also beats the RW. This means that for a very wide range of half-lives, between 1 and 10 years, the HL model beats its competitors. For values higher than three years the AR model loses also with the RW model, as estimation error of estimating an autoregressive process is more severe than the model misspecification error of assuming a RW.⁵

The bottom line is that in most univariate applications, unless the sample is very long, the AR model is likely to produce very imprecise forecasts. It is hence much preferable to simply employ a reasonably calibrated HL model, which assumes a gradual mean reversion to the sample mean.

⁵All the results have also been cross-checked with Monte Carlo simulations.

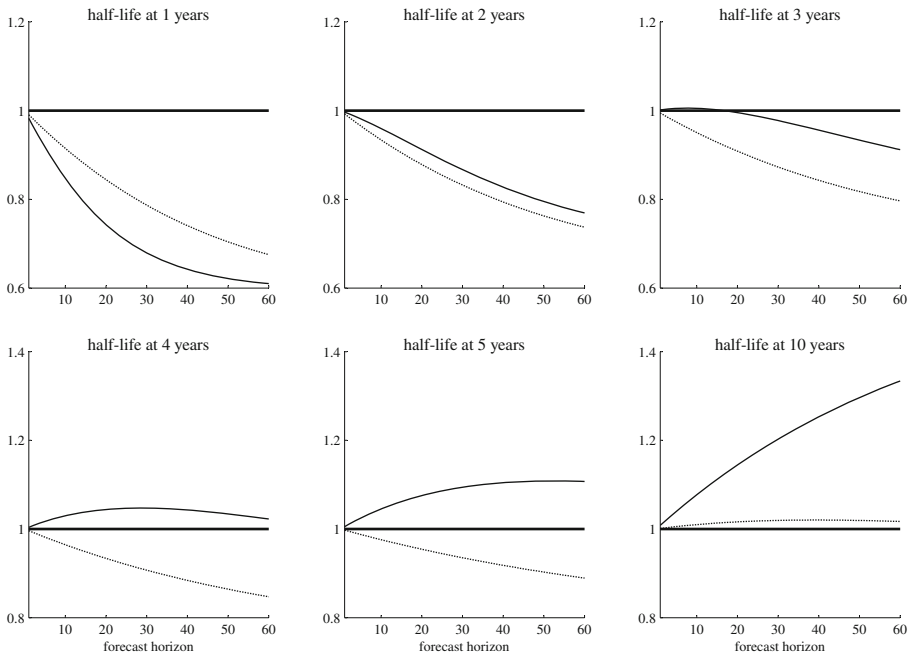


Fig. 3 Theoretical mean squared forecast errors. Notes: Each *line* represents the ratio of MSFE from a given method to MSFE from the random walk, where values below unity indicate better accuracy of point forecasts. The *straight* and *dotted lines* stand for AR and HL, respectively. The forecast horizon is expressed in months

5 Additional Remarks on the HL Model

While not particularly appealing from the theoretical perspective, apart from being consistent with long-run PPP, the HL model performs extremely well in the forecasting competition. What can explain this success? The shrinkage principle tells us that applying restrictions to economic models may be helpful out of sample even if they deteriorate their in-sample fit. Moreover, studies based on Monte Carlo techniques (e.g. Gilbert 1995), as well as our analytical derivations discussed in Section 4, show that the knowledge of the true structure of the DGP is often not particularly useful for predictive purposes if the parameter values are unknown. This is essentially the reasons why the estimated AR model is beaten by the calibrated HL framework. A helpful way of thinking about this is to recall George Box (1979) famous remark that “essentially, all models are wrong, but some are useful”.

From the results of our estimated autoregressive model it is straightforward to understand why a most likely false restriction (that the half-life adjustment is constant and equal to five years) has improved our forecasting performance. Table 4 reports rolling descriptive statistics for the estimated AR model. The first row presents the share of estimated half-lives in Rogoff’s range of 3 and 5 years, which is in general small. The second row broadens the range to between 1 and 10 years. The third

Table 4 Rolling descriptive statistics; half-lives expressed in monthly terms

	AUD	CAD	EUR	JPY	MXN	NZD	CHF	GBP	USD
15Y									
N_1	29.85	10.07	6.72	22.76	61.19	40.67	49.25	23.13	7.09
N_2	69.78	30.97	60.45	69.40	100.00	76.49	96.27	88.81	41.42
N_3	2.61	20.15	2.99	2.24	0.00	0.00	1.49	1.12	0.00
Mean	13.81	90.67	12.21	13.98	3.89	7.98	4.22	7.10	14.60
std	43.40	324.96	11.27	26.57	1.18	5.78	2.17	5.48	10.87
Median	5.65	16.98	9.17	6.32	3.60	6.31	3.83	6.29	10.76
$Q(.05)$	3.16	3.70	4.77	3.41	2.24	3.18	2.50	3.62	4.29
$Q(.25)$	4.50	7.48	7.24	4.99	3.13	3.69	2.95	4.98	7.67
$Q(.75)$	10.80	71.67	12.02	11.09	4.62	9.73	4.70	7.16	19.09
$Q(.95)$	28.69	317.50	32.07	47.45	5.87	19.22	7.52	14.57	34.16
10Y									
N_1	34.33	13.06	14.18	41.42	41.04	30.60	42.16	14.93	12.69
N_2	87.31	40.67	62.69	83.96	88.81	65.67	91.42	82.09	62.31
N_3	1.49	26.12	8.96	1.87	4.48	1.49	3.73	3.73	6.34
Mean	9.59	52.95	29.42	20.34	13.18	11.70	3.93	7.04	17.66
std	38.78	167.66	254.40	209.90	89.37	26.25	2.89	9.51	44.76
Median	4.61	9.37	7.24	5.36	3.14	8.30	3.38	5.29	7.10
$Q(.05)$	2.24	3.10	1.86	3.13	1.84	2.42	1.62	1.37	1.69
$Q(.25)$	3.36	5.33	3.63	3.97	2.57	3.62	2.56	2.42	4.46
$Q(.75)$	7.41	14.84	10.64	7.72	3.86	11.84	4.20	8.50	14.93
$Q(.95)$	14.24	263.53	50.41	17.50	26.03	29.55	8.85	14.76	61.83
20Y									
N_1	13.06	0.00	0.00	7.46	73.51	42.91	70.90	25.75	4.48
N_2	55.22	8.96	70.15	52.24	100.00	90.30	98.13	90.30	39.93
N_3	1.12	13.06	0.00	0.00	0.00	0.00	1.12	0.00	0.00
Mean	12.08	1435.28	9.77	13.30	4.22	6.15	4.21	6.78	12.43
std	14.79	20805.36	5.23	16.38	1.51	2.60	2.15	3.55	4.94
Median	9.63	17.03	8.14	9.75	3.63	5.35	4.15	5.99	12.08
$Q(.05)$	3.60	8.74	6.22	4.31	2.85	3.07	2.50	3.98	5.55
$Q(.25)$	6.10	11.92	7.07	7.85	3.32	4.10	3.41	4.92	7.91
$Q(.75)$	12.18	26.99	10.51	12.57	4.09	8.37	4.61	7.10	15.69
$Q(.95)$	25.97	234.78	17.00	31.06	7.51	10.63	5.86	11.22	20.38

$$N_1 = \sum \mathcal{I}(\hat{HL} \in (3, 5))/N; N_2 = \sum \mathcal{I}(\hat{HL} \in (1, 10))/N \text{ and } N_3 = \sum \mathcal{I}(|\hat{\rho}| \notin (-1, 1))/N$$

row provides the fraction of cases for which the AR model produces explosive forecasts. The remaining rows report the mean, standard deviation, median and different quantiles of the estimated half-lives. These statistics illustrate perfectly well the high degree of estimation error and why calibration turns out to be the winning strategy out of sample.

This strategy is remarkably close to that proposed by Faust and Wright (2013) on inflation forecasting. These authors have shown that assuming a smooth path of adjustment, between the last inflation reading and a good long term forecast, results in a projection that is hard to beat by more sophisticated econometric approaches. This confirms Box (1979) intuition that a cunningly well-chosen parsimonious model tends to beat overwhelmingly any method giving too much weight to in-sample fluctuations.

6 Sensitivity Analysis

In this section we will show that the HL model turns out to be the most competitive model also when we change the forecast setting in our baseline. We shall then exploit its good performance and extend the analysis to nominal exchange rate forecasting.

6.1 Rolling Window Length

We begin by analyzing whether a change in the length of the rolling window has an impact on our findings. A longer rolling window should, in theory, increase the accuracy of the HL and AR models, as implied by Eqs. 13 and 14. In the case of the HL model, a longer rolling window helps the modeler to determine with more precision the PPP level. In the case of the AR model, a longer window also helps one to better determine the degree of real exchange rate persistence. A longer rolling window may, on the other hand, be potentially counterproductive, if we relax the assumption that we had in the analytical section that the equilibrium value of the real exchange rate is time-invariant (see Rossi 2006 for a discussion on the importance of parameter instability). As shown by Tables 5 and 6, for most currencies in our sample this latter effect seems to play a lesser role, considering that both the HL and AR models tend to become more competitive for longer rolling windows. For a 20 year rolling window as well as in the case of recursive estimation, the HL models outperform the RW model for 8 out of 9 currencies at almost all horizons.⁶ For shorter samples, as in the case of a 10 year rolling window, the HL model continues to generally beat the RW (but this is no longer the case for the US dollar). The AR model instead generates, as expected, inaccurate forecasts, which confirms that the estimation error is the main source of the weak performance of the AR model. To sum up, for the currencies in our sample a rolling window of at least 15–20 years represents a good choice.

6.2 Prior on the Half-Life Parameter

A Bayesian autoregressive process may potentially outperform the HL models. To establish this, we set the mean-reversion parameter ρ as prior information rather than

⁶Results for the case of recursive estimation are available upon request and would not change the overall assessment of this paper.

Table 5 Mean squared forecast errors (10Y rolling window)

<i>h</i>	RW	HL	AR	RW	HL	AR	RW	HL	AR
	AUD			CAD			EUR		
1	0.05	1.00	1.05	0.02	1.03	1.02	0.02	1.01	1.04*
6	0.44	0.99	1.15	0.23	1.07	1.10*	0.19	0.99	1.14**
12	0.82	0.98	1.10	0.46	1.14*	1.21	0.43	0.96	1.24**
24	1.53	0.96*	1.09	0.94	1.23**	1.24**	0.89	0.92	1.55**
36	2.10	0.95	1.20	1.58	1.21**	1.32*	1.28	0.88	1.96**
60	3.00	1.02	1.63	3.02	1.09	1.26*	2.06	0.77*	2.34*
	JPY			MXN			NZD		
1	0.06	1.00	1.02	0.12	0.99	1.02	0.03	1.00	1.00
6	0.59	0.98	1.03	0.83	0.95	1.06	0.32	0.96	0.99
12	1.00	0.99	1.08*	1.55	0.90**	1.15	0.72	0.92**	0.97
24	2.34	0.92	1.10	3.01	0.81**	1.26*	1.59	0.84**	0.92
36	3.53	0.87	1.08	3.66	0.73**	1.10	2.44	0.75**	0.96
60	3.12	0.95	1.14	3.56	0.68**	0.89	3.01	0.64**	1.26
	CHF			GBP			USD		
1	0.02	1.00	1.05	0.03	1.00	1.05	0.02	1.01	1.04**
6	0.12	0.99	1.24	0.24	0.98	1.25	0.19	1.01	1.17**
12	0.25	0.99	1.09	0.47	0.98	1.52	0.31	1.04	1.38**
24	0.50	0.92	1.02	1.06	0.93	3.35	0.55	1.05	1.55**
36	0.72	0.84**	0.88**	1.49	0.91	13.45	0.69	1.03	1.78**
60	0.79	0.76**	0.86	1.98	0.78**	0.82*	1.41	0.75*	1.87**

As in Table 1

simply just impose it as we had done in the calibrated version of the model. To assess the implication of this choice let us consider a Bayesian autoregressive model (BAR), along the line suggested by Kilian and Zha (2002). We use the standard Minnesota setting for vector autoregressions to elicit our prior on the degree of PPP persistence. In particular, we write down the model (1) in the standard AR form:

$$y_t = \delta + \rho y_{t-1} + \epsilon_t, \tag{15}$$

where $\delta = (1 - \rho)\mu$. The prior for $\alpha = [\delta \ \rho]'$ is assumed to be $\mathcal{N}(\underline{\alpha}, \underline{V})$ with $\underline{\alpha} = [(1 - \bar{\rho})\bar{\mu} \ \bar{\rho}]'$ and $\underline{V} = \text{diag}(\lambda\sigma^2, \lambda)$, where σ is the residual standard error from the AR model, $\bar{\rho}$ is the mean-reversion parameter calibrated so that the half-life is five years and λ is the overall tightness hyperparameter. The expected value of the posterior is:

$$\bar{\alpha} = \left(\underline{V}^{-1} \underline{\alpha} + \sigma^{-2} X' X \hat{\alpha} \right),$$

where $\hat{\alpha}$ is the LS estimate of α , X is the observation matrix and $\bar{V} = (\underline{V}^{-1} + \sigma^{-2} X' X)$. The parameter λ has a very simple intuitive explanation for it allows us to

Table 6 Mean squared forecast errors (20Y rolling window)

<i>h</i>	RW	HL	AR	RW	HL	AR	RW	HL	AR
	AUD			CAD			EUR		
1	0.05	1.00	1.05	0.02	1.02	1.01	0.02	1.00	1.00
6	0.44	0.99	1.15	0.23	1.05	1.05*	0.19	0.95	0.98
12	0.82	0.98*	1.10*	0.46	1.09	1.11**	0.43	0.90*	0.94
24	1.53	0.96*	1.09**	0.94	1.15*	1.17**	0.89	0.82**	0.90*
36	2.10	0.95*	1.20*	1.58	1.11	1.18**	1.28	0.77**	0.86**
60	3.00	1.02	1.63	3.02	0.98	1.35**	2.06	0.67**	0.75**
	JPY			MXN			NZD		
1	0.06	1.00	1.01	0.12	0.99	1.00	0.03	1.00	1.01
6	0.59	0.99	1.05*	0.83	0.94 *	0.94	0.32	0.97	0.99
12	1.00	1.00	1.12**	1.55	0.89**	0.87	0.72	0.92*	0.94
24	2.34	0.94	1.17**	3.01	0.79**	0.72**	1.59	0.82**	0.82**
36	3.53	0.89	1.16**	3.66	0.72**	0.67**	2.44	0.72**	0.69**
60	3.12	0.90	1.13	3.56	0.68 **	0.69*	3.01	0.59**	0.55**
	CHF			GBP			USD		
1	0.02	1.00	1.02	0.03	0.99	1.00	0.02	1.00	0.99
6	0.12	0.98	1.06	0.24	0.96	0.96	0.19	0.96	0.96*
12	0.25	0.97	1.02	0.47	0.93	0.91	0.31	0.94	0.92**
24	0.50	0.89*	0.87	1.06	0.84**	0.85**	0.55	0.86*	0.87**
36	0.72	0.82**	0.74**	1.49	0.79**	0.77**	0.69	0.78*	0.86**
60	0.79	0.78*	0.66**	1.98	0.65**	0.62**	1.41	0.59**	0.82**

As in Table 1

choose an intermediate solution between the calibrated and autoregressive solution. The two corner solutions are the calibrated solution (for $\lambda = 0$ $\bar{\alpha}$ collapses to $\underline{\alpha}$) and the estimated solution (for $\lambda \rightarrow \infty$ $\bar{\alpha}$ equals to $\hat{\alpha}$).

We report in Table 7 the ratios between the MSFEs from the Bayesian autoregressive model (reported as BAR in the table) and the MSFEs from the RW model for λ equal to 0, .1 and ∞ . For the intermediate case $\lambda = 0.1$ such ratios are typically higher than those corresponding to the HL model and lower than those corresponding to the AR models. In other words in general the relative MSFEs tend to increase monotonically with the rising of λ . The best solution is therefore to set $\lambda = 0$, i.e. the calibrated solution.

For the one-month horizon we also provide a graphical illustration of what we have just said for values of λ ranging on a continuous scale between zero and ∞ (see Fig. 4). On the vertical axis the values of MSFE are normalized so that MSFE is equal to 100 for $\lambda = 0$, which corresponds to the case of the calibrated HL model. For six currencies (EUR, JPY, NZD, CHF, GBP, USD) the relationship is indeed increasing and monotonic, i.e. the more weight one gives to estimation error the worse is the forecasting performance of the Bayesian autoregressive model. For one

Table 7 Mean squared forecast errors (MSFEs) – BAR model

<i>h</i>	HL	BAR	AR	HL	BAR	AR	HL	BAR	AR
λ	0	.1	∞	0	.1	∞	0	.1	∞
	AUD			CAD			EUR		
1	1.01	1.02*	1.02*	1.02	1.03*	1.03*	1.00	1.04**	1.04**
6	1.03	1.05*	1.06*	1.03	1.11**	1.11**	0.96**	1.12*	1.12*
12	1.06	1.09*	1.09*	1.06	1.19**	1.20**	0.92*	1.14*	1.14*
24	1.10*	1.09	1.09	1.10	1.19*	1.19*	0.83**	1.17*	1.18*
36	1.12	1.06	1.06	1.06	1.20*	1.20*	0.77**	1.12*	1.13*
60	1.11	1.06	1.06	0.94	1.44**	1.45**	0.66**	0.91	0.91
	JPY			MXN			NZD		
1	1.00	1.01	1.01	0.99	0.99	0.99	1.00	1.03*	1.03*
6	0.98	1.04	1.04	0.96*	0.92	0.92	0.96	1.05	1.06
12	0.97	1.10*	1.10*	0.92*	0.87*	0.87*	0.92*	1.03	1.03
24	0.91	1.16**	1.17**	0.84**	0.78*	0.78*	0.83**	0.89**	0.89**
36	0.86	1.19**	1.19**	0.78**	0.74**	0.74**	0.74**	0.75**	0.74**
60	0.89	1.19*	1.19*	0.74**	0.72*	0.72*	0.64**	0.62**	0.62**
	CHF			GBP			USD		
1	1.00	1.05	1.06	1.00	1.02	1.02	1.00	1.03*	1.03*
6	0.98	1.21*	1.22*	0.97	1.04	1.04	0.96	1.10*	1.10*
12	0.97	1.15	1.16	0.95	1.03	1.03	0.93	1.21*	1.21*
24	0.88*	0.99	0.99	0.87*	0.98	0.98	0.84*	1.21*	1.21*
36	0.80**	0.79**	0.78**	0.82**	0.93	0.93	0.72**	1.12	1.13
60	0.72**	0.70**	0.69**	0.67**	0.70**	0.70**	0.53**	0.91	0.91

As in Table 1

currency (MXN), the estimated model performs the best. Only for two currencies (AUD and CAD) and very specific ranges of λ we find additional gains from using a Bayesian autoregressive model.

6.3 Other Currencies

As a third robustness check we evaluate whether the results are applicable to other currencies as well. We thus consider the full set of real effective exchange rates indices available in the Bank for International Settlements database. The additional sample consists of eighteen currencies for the following countries: Austria (ATS), Belgium (BEF), Taiwan (TWD), Denmark (DKK), Finland (FIM), France (FRF), Germany (DEM), Greece (GRD), Hong Kong (HKG), Ireland (IEP), Italy (ITL), the South Korea (KRW), the Netherlands (NLG), Norway (NOK), Portugal (PTE), Singapore (SGD), Spain (ESP) and Sweden (SEK). The results are reported in Table 8 and lead to similar conclusions to those reached earlier. The forecasts based on the HL model are better than those based on the RW for 9 of the 18 currencies, comparable

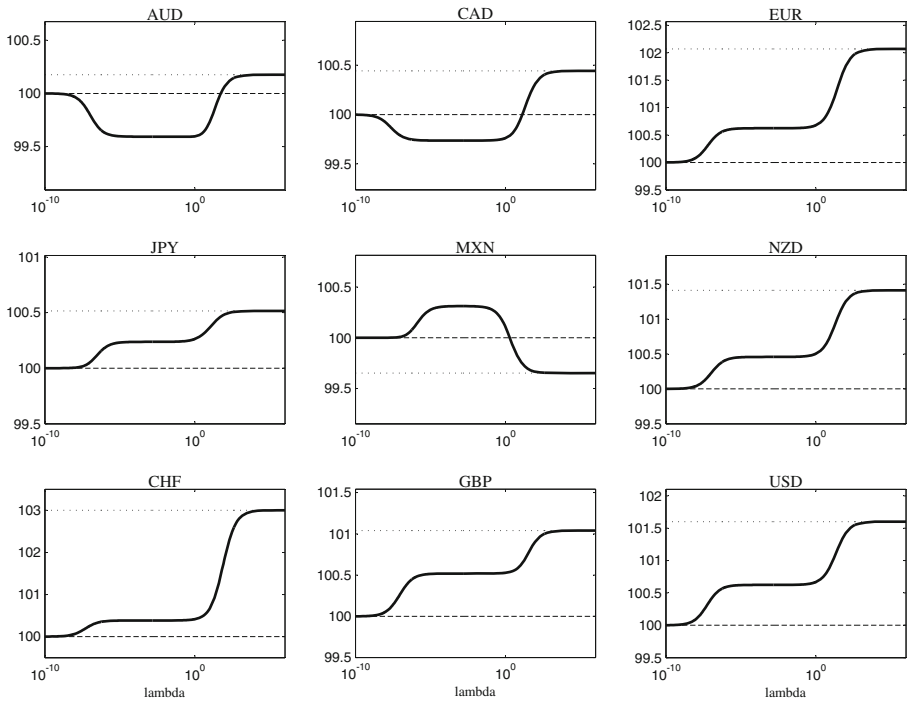


Fig. 4 Sensitivity analysis of MSFE on the λ (forecast horizon: 1 month). Notes: Each *line* represents the ratio of MSFE from a given method to MSFE from the HL5 multiplied by 100, where values 100 unity indicate better accuracy of point forecasts. The *straight, dashed and dotted lines* stand for BAR, HL and AR, respectively. The value of λ parameter is expressed using the logarithmic scale

for 6 and less accurate for 3. The HL model also delivers more precise forecasts than the AR model for most currencies.

6.4 Sensitivity to the HL Parameter

Next, we evaluate if the performance of the HL model is sensitive to the duration of the adjustment process. Table 9 reports the relative performance of the HL model compared to the RW assuming that half of the adjustment is completed in 1, 3 and 10 years respectively. In the large majority of cases the HL model outperforms the RW regardless of this choice: the HL model beats the RW at the lower bound proposed by Rogoff (HL3) but is also quite competitive for half-lives in the broad range of 1 to 10 years. Opting for fast convergence to PPP, such as in the case of the HL1 model, the calibrated half-life model continues to perform satisfactorily for forecast horizons above two years. Opting for a much lower pace of convergence, such as in the HL10 model, the HL beats the RW at all horizons. However, at longer horizons the performance of the HL10 model is not as good as the HL3 or HL5 model, suggesting that it is still preferable to select a faster pace of convergence to PPP.

Table 8 Mean squared forecast errors for other currencies

<i>h</i>	RW	HL	AR	RW	HL	AR	RW	HL	AR
	ATS			BEF			TWD		
1	0.05	1.00	1.05	0.00	1.00	1.04	0.02	1.01	1.01
6	0.44	0.99	1.15	0.03	0.96	1.09	0.13	1.07	1.06
12	0.82	0.98	1.10	0.07	0.92	1.07	0.25	1.15*	1.13
24	1.53	0.96	1.09	0.14	0.83*	1.00	0.43	1.39**	1.34**
36	2.10	0.95	1.20	0.21	0.74**	0.94	0.56	1.69**	1.49**
60	3.00	1.02*	1.63	0.38	0.60**	0.64**	1.05	1.83**	1.57**
	DKK			FIM			FRF		
1	0.00	1.00	1.02	0.01	1.01	1.01	0.00	1.00	1.02
6	0.04	0.97	1.07	0.15	0.99	1.03	0.03	0.97	1.06
12	0.08	0.94	1.10	0.40	0.94	1.06**	0.08	0.92	1.05
24	0.12	0.94	1.19*	1.07	0.85*	1.12**	0.15	0.83**	1.10
36	0.14	1.06	1.26*	1.34	0.83	1.19**	0.21	0.76**	1.11
60	0.18	1.23*	1.43**	0.89	1.02	1.60**	0.33	0.67**	0.89
	DEM			GRD			HKD		
1	0.01	0.99	1.03	0.02	1.01	1.03*	0.02	1.04	1.05*
6	0.06	0.95	1.12	0.05	1.17**	1.35**	0.23	1.11	1.23**
12	0.13	0.89*	1.11	0.08	1.41**	1.76**	0.48	1.20	1.50**
24	0.28	0.76**	1.03	0.23	1.52**	1.80**	1.23	1.18	1.84**
36	0.41	0.65**	0.91	0.41	1.55**	1.66**	2.06	1.14	2.13**
60	0.64	0.47**	0.75**	0.77	1.65**	1.60**	4.56	0.92	2.94**
	IEP			ITL			KRW		
1	0.01	1.01	1.03	0.02	1.00	1.04**	0.10	0.99	1.01
6	0.12	1.00	1.06	0.13	0.95	1.15**	0.82	0.94*	1.02
12	0.26	0.99	1.13	0.30	0.88	1.26**	1.48	0.89**	0.99
24	0.56	0.95	1.33**	0.63	0.74**	1.35**	2.65	0.81**	0.99
36	0.92	0.90	1.32	1.08	0.61**	1.38**	3.00	0.77**	1.17
60	1.33	0.91	0.94	1.09	0.46**	1.65**	3.11	0.83	3.76
	NLG			NOK			PTE		
1	0.01	1.00	1.02	0.02	0.99	1.01	0.01	1.08*	1.20**
6	0.05	0.97	1.05	0.12	0.96	1.00	0.05	1.29**	1.75**
12	0.10	0.93	0.98	0.22	0.92*	0.97	0.12	1.39**	1.94**
24	0.20	0.83**	0.86	0.25	0.93	1.07	0.23	1.72**	1.94**
36	0.26	0.79*	0.83	0.29	0.97	1.12	0.23	2.44**	2.14**
60	0.40	0.74**	0.72*	0.28	1.08	1.28*	0.38	3.17**	2.19**

6.5 An Extension to Nominal Exchange Rate Forecasting

The next step in our analysis consists in testing whether the mean reverting nature of the real exchange rate helps us to forecast nominal exchange rates. A simple

Table 8 (continued)

<i>h</i>	RW	HL	AR	RW	HL	AR	RW	HL	AR
	SGD			ESP			SEK		
1	0.01	1.00	1.02*	0.01	1.01	1.04**	0.02	1.00	1.01
6	0.08	1.00	1.08*	0.06	0.99	1.17**	0.24	0.99	0.99
12	0.19	0.96	1.09	0.16	0.91	1.20**	0.52	0.97	0.96
24	0.52	0.84**	1.04	0.37	0.76*	1.18*	0.95	0.97	0.96
36	0.85	0.73**	0.97	0.55	0.61**	1.03	1.09	1.10	0.96
60	1.53	0.57**	0.83**	0.84	0.52**	0.82*	1.18	1.48**	1.03

As in Table 1

approach is to assume that the adjustment of the real exchange rate predicted by model *M* is entirely achieved via changes in nominal exchange rates, while the relative price channel is absent. The predicted change of log nominal exchange

Table 9 Mean squared forecast errors for other HL duration

<i>h</i>	HL1	HL3	HL10	HL1	HL3	HL10	HL1	HL3	HL10
	AUD			CAD			EUR		
1	1.18**	1.03	1.00	1.30**	1.04	1.01	1.09	1.00	1.00
6	1.47**	1.07	1.01	1.74**	1.10	1.01	1.17	0.96	0.97
12	1.71**	1.13	1.02	2.02**	1.17	1.01	1.11	0.92*	0.95*
24	1.79**	1.22**	1.04	2.05**	1.24*	1.02	0.92	0.83**	0.89**
36	1.72**	1.25*	1.04	1.68**	1.18	1.00	0.76	0.77**	0.85**
60	1.46**	1.22	1.03	1.13	0.99	0.93	0.52**	0.66**	0.78**
	JPY			MXN			NZD		
1	1.10*	1.01	1.00	1.01	0.99	1.00	1.07	1.00	1.00
6	1.22	0.99	0.98	0.97	0.94	0.98*	1.06	0.96	0.98
12	1.41*	1.00	0.97	0.92	0.89*	0.95*	0.96	0.90*	0.95*
24	1.19	0.92	0.93	0.76	0.79**	0.90**	0.72*	0.77**	0.90**
36	1.04	0.86	0.90*	0.71*	0.72**	0.86**	0.54**	0.65**	0.84**
60	1.20	0.97	0.88*	0.79	0.73**	0.81**	0.45**	0.55**	0.77**
	CHF			GBP			USD		
1	1.04	1.00	1.00	1.08*	1.00	1.00	1.10*	1.00	1.00
6	1.10	0.98	0.98	1.16	0.97	0.98	1.19	0.96	0.98
12	1.09	0.97	0.98	1.19	0.95	0.97*	1.33	0.94	0.95*
24	0.84	0.84*	0.93**	0.94	0.85*	0.92**	1.18	0.84	0.89**
36	0.66**	0.73**	0.88**	0.78	0.77**	0.88**	0.97	0.71*	0.81**
60	0.61**	0.65**	0.81**	0.53**	0.59**	0.78**	0.42**	0.45**	0.68**

As in Table 1

Table 10 Mean squared forecast errors for NEERs

<i>h</i>	RW	HL	AR	RW	HL	AR	RW	HL	AR
	AUD			CAD			EUR		
1	0.05	1.01	1.02*	0.02	1.02	1.03*	0.02	1.00	1.04*
6	0.44	1.02	1.06*	0.24	1.03	1.11**	0.19	0.97	1.12*
12	0.78	1.06	1.11**	0.47	1.06	1.22**	0.40	0.93	1.15*
24	1.30	1.12 *	1.13*	0.86	1.10	1.26**	0.81	0.85*	1.20**
36	1.70	1.16 *	1.10	1.42	1.05	1.29**	1.13	0.78**	1.16**
60	2.23	1.21 *	1.09	2.72	0.91	1.56**	1.78	0.66**	0.93
	JPY			MXN			NZD		
1	0.06	1.00	1.01	0.12	1.00	1.00	0.03	1.00	1.03*
6	0.57	0.99	1.03	1.23	0.98	0.95	0.32	0.96	1.06
12	1.03	0.99	1.08*	2.92	0.95*	0.91	0.69	0.91*	1.03
24	2.55	0.93	1.11*	7.19	0.90**	0.84*	1.41	0.81**	0.89*
36	3.98	0.88	1.11*	11.86	0.87**	0.81**	2.17	0.70**	0.74**
60	3.51	0.99	0.95	25.53	0.88**	0.82**	2.69	0.60**	0.59**
	CHF			GBP			USD		
1	0.02	1.00	1.06 *	0.03	1.00	1.02*	0.03	1.00	1.03*
6	0.13	1.00	1.22 *	0.23	0.99	1.06	0.25	0.97	1.09 *
12	0.29	0.99	1.19	0.49	0.98	1.06	0.45	0.95	1.17**
24	0.62	0.93*	1.01	1.12	0.92	1.02	0.90	0.88*	1.15**
36	0.90	0.87**	0.85*	1.69	0.88*	0.98	1.26	0.81*	1.05
60	1.12	0.87**	0.83*	2.42	0.74**	0.79**	2.52	0.67*	0.87**

As in Table 1

rate (*s*) at horizon *h* is thus simply equal to the predicted real exchange rate (*y*) adjustment:

$$s_{T+h|T}^M - s_T = y_{T+h|T}^M - y_T. \tag{16}$$

The results presented in Table 10 are based on the same settings that we had earlier in our baseline for real exchange rate forecasting. The calibrated HL model performs visibly better than the RW for exactly the same seven currencies for which it was superior in the context of real exchange rate forecasting. The forecasts generated by the HL model are also generally much more precise than those generated by the AR model. Comparing the numbers in Tables 1 and 10 clearly shows that for our set of currencies the ability to forecast real and nominal exchange rates is not very different. For the same seven currencies for which the real exchange rate are forecastable, the nominal exchange rate has contributed, amid high volatility, to the gradual mean reversion process of the real exchange rate rather than just followed the RW.⁷ The

⁷Zumaquero and Urrea (2002) similarly find, using cointegration techniques, that the adjustment to PPP takes place mainly via the exchange rate rather the relative price channel.

role of relative prices may be hypothetically more important in the case of countries that have a fixed exchange rate regime.

6.6 An Extension to Real and Nominal Bilateral Exchange Rates

In our study we use ex-post real effective exchange rates that are subject to revision of weights. Consequently, these data are not available in real time. Even though weights revisions are generally small, they might still influence the relative performance of

Table 11 Mean squared forecast errors for bilateral ERs against USD

<i>h</i>	RW	HL	AR	RW	HL	AR	RW	HL	AR	RW	HL	AR
REAL EXCHANGE RATE												
	AUD			CAD			EUR			JPY		
1	0.07	1.01	1.02**	0.03	1.01	1.02*	0.06	1.00	1.04*	0.08	1.00	1.02
6	0.72	1.00	1.06**	0.28	1.01	1.10**	0.52	0.96	1.15*	0.62	0.99	1.09*
12	1.52	1.00	1.08**	0.58	1.02	1.19**	0.98	0.91*	1.23**	1.02	0.99	1.20**
24	3.04	0.96	1.07	1.15	1.02	1.22*	2.00	0.81**	1.19*	2.44	0.92	1.22**
36	4.13	0.93	1.05	1.86	0.98	1.27*	2.79	0.72**	1.02	3.50	0.87	1.30**
60	6.87	0.83*	1.04	3.72	0.84	1.58**	4.77	0.59**	0.78**	3.51	0.80	1.50**
	MEX			NZD			CHF			GBP		
1	0.11	0.99	0.98	0.07	1.00	1.04**	0.08	1.00	1.04*	0.06	0.99	1.02
6	0.80	0.96	0.91	0.78	0.98	1.12**	0.59	0.97	1.14**	0.52	0.95*	1.01
12	1.42	0.92*	0.87	1.81	0.94	1.18**	1.04	0.94	1.22**	0.84	0.88**	0.97
24	2.62	0.84**	0.78*	3.96	0.85**	1.19**	1.99	0.85*	1.13	1.54	0.77**	0.76**
36	3.05	0.79**	0.76*	5.97	0.76**	1.11*	2.69	0.77**	0.95	1.73	0.64**	0.60**
60	2.72	0.75**	0.72**	8.72	0.65**	1.02	4.36	0.62**	0.78**	1.88	0.53**	0.40**
NOMINAL EXCHANGE RATE												
	AUD			CAD			EUR			JPY		
1	0.07	1.01	1.02**	0.03	1.01	1.03*	0.06	1.00	1.04*	0.07	1.00	1.02
6	0.76	1.00	1.07**	0.29	1.01	1.11**	0.56	0.96	1.14*	0.61	0.99	1.07
12	1.53	1.00	1.10**	0.59	1.02	1.21**	1.00	0.91*	1.22**	1.06	1.00	1.14*
24	2.85	0.96	1.10*	1.10	1.02	1.28**	1.96	0.81**	1.19*	2.62	0.93	1.11
36	3.81	0.93	1.09	1.78	0.97	1.35**	2.75	0.72**	1.03	3.83	0.88	1.12
60	6.35	0.81**	1.05	3.62	0.82*	1.68**	4.76	0.59**	0.79**	3.87	0.87	1.07
	MEX			NZD			CHF			GBP		
1	0.12	0.99	0.98	0.07	1.00	1.04**	0.08	1.00	1.03*	0.06	0.99	1.02
6	1.23	0.96*	0.91*	0.81	0.98	1.12**	0.63	0.98	1.13*	0.56	0.95*	1.00
12	2.85	0.93*	0.85*	1.83	0.94	1.18**	1.11	0.95	1.20**	0.89	0.88**	0.96
24	7.37	0.88**	0.78**	3.76	0.84**	1.20**	2.13	0.88*	1.10	1.64	0.77**	0.75
36	12.91	0.85**	0.77**	5.59	0.76**	1.12*	2.96	0.81*	0.92	2.04	0.67**	0.60**
60	29.66	0.86**	0.79**	8.31	0.65**	1.03	5.03	0.70**	0.76**	2.34	0.58**	0.44**

As in Table 1

our competing models. It is hence important to show that the mean reverting properties of the real exchange rate are not dependent on the revision of the weights by the BIS nor are just a mere artefact of how different currency pairs are aggregated. The best way to prove this is to show that the HL model can forecast in a very precise way also bilateral exchange rates. In Table 11 we report our forecasting competition for the real and nominal exchange rate of the US dollar vis-à-vis the eight remaining currencies that we had analyzed before. As a reference currency we have chosen the US dollar, and not the euro, to avoid dependence of our results to the way we have defined the synthetic euro and the price index for the euro area before EMU creation. The results are exceptionally good. As it turns out the HL model clearly outperforms all other models for all currency pairs at medium term horizons. For most currencies the HL model performs comparably or often even better than the RW also at short-run horizons.

7 Conclusions

Notwithstanding the recent important progress made in the field of exchange rate economics, we still know very little of what drives the fluctuations of major currencies. Moreover, as evidenced by numerous studies, exchange rate forecasts tend to be inaccurate both in absolute sense and relative to a naïve RW. Solving the “exchange rate puzzle” has been an endeavor for many economists over the past three decades. The vast exchange rate literature provides, however, at least two reasons for being cautiously optimistic. First, a number of papers have emphasized that the dismal forecasting performance of exchange rate models is partly due to estimation error, which explains why the RW is less competitive for larger datasets. Second, the literature on PPP half-life has shown that there is evidence of mean reversion in real exchange rates.

In this paper we have illustrated how these two findings can be exploited in real exchange rate forecasting. In particular, we have proposed a simple model that assumes a gradual return of the real exchange rate to its sample mean. From the theoretical perspective this alternative is more appealing than the RW for it takes into account that PPP holds in the long-term horizon. It is also appealing from the empirical perspective for it is consistent with the evidence that real exchange rates are highly persistent but mean reverting.

The key finding of our analysis is that the proposed HL model is able to overwhelmingly beat the naïve RW in terms of real effective and bilateral exchange rate forecasting for seven out of nine major world currencies. Moreover, a model of gradual adjustment toward the PPP rate outperforms the RW also at short-horizons: the naïve model is beaten already at the 6 month horizon for both the US dollar and the euro. We believe that our results are quite intuitive: our preferred forecasting model for real exchange rates resembles quite closely the RW in the short-run while it gradually approaches and reaches PPP over medium to long term horizons.

A second key finding of our analysis is that if the speed of mean reversion pace is estimated then the model performs significantly worse than the RW. We explain this result analytically by showing that estimation forecast error plays an important role

even for horizons of 15 years of monthly data. We have carried out a comprehensive sensitivity analysis to check the robustness of our results to different rolling windows and the choice of analyzed currencies. We have found that the results are valid for a wide range of half-lives, as long as they are calibrated at reasonable values, rather than estimated, irrespective of whether we use Bayesian techniques. Finally, we have demonstrated that the mean reverting nature of real exchange rates can be exploited to outperform the RW also in terms of nominal exchange rate forecasting. For most currencies in our sample we find that the nominal exchange rate has contributed to the mean reversion process of the real exchange rate rather than just followed a RW.

References

- Bergin PR, Glick R, Wu J-L (2013) The micro-macro disconnect of purchasing power parity. *Rev Econ Stat* 95(3):798–812
- Box G (1979) Robustness in the strategy of scientific model building. *Robustness in Statistics* 4:201–236
- Cheung Y-W, Chinn MD, Pascual AG (2005) Empirical exchange rate models of the nineties: are any fit to survive? *J Int Money Financ* 24(7):1150–1175
- Chinn MD, Meese RA (1995) Banking on currency forecasts: how predictable is change in money? *J Int Econ* 38(1–2):161–178
- Crucini MJ, Shintani M (2008) Persistence in law of one price deviations: evidence from micro-data. *J Monet Econ* 55(3):629–644
- Diebold FX (2012) Comparing predictive accuracy, twenty years later: a personal perspective on the use and abuse of Diebold-Mariano tests. *NBER Working Papers*, p 18391
- Diebold FX, Mariano RS (1995) Comparing predictive accuracy. *J Bus Econ Stat* 13(3):253–63
- Engel C (2013) Exchange rates and interest parity. *NBER Working Papers*, p 19336
- Engel C, Mark NC, West KD (2008) Exchange rate models are not as bad as you think. In: Acemoglu D, Rogoff K, Woodford M (eds) *NBER Macroeconomics Annual 2007*, vol 22, pp 381–441
- Faust J, Rogers JH, Wright JH (2003) Exchange rate forecasting: the errors we've really made. *J Int Econ* 60(1):35–59
- Faust J, Wright JH (2013) Forecasting inflation. In: Elliott G., Timmermann A (eds) *Handbook of economic forecasting*, vol 2, Elsevier, Ch 1, p 2–56
- Fuller WA, Hasza DP (1980) Predictors for the first-order autoregressive process. *J Econ* 13(2):139–157
- Gilbert PD (1995) Combining VAR estimation and state space model reduction for simple good predictions. *J Forecast* 14:229–250
- Holmes M, Otero J, Panagiotidis T (2012) PPP in OECD countries: an analysis of real exchange rate stationarity, cross-sectional dependency and structural breaks. *Open Econ Rev* 23(5):767–783
- Imbs J, Mumtaz H, Ravn M, Rey H (2005) PPP strikes back: aggregation and the real exchange rate. *Q J Econ* 120(1):1–43
- Ince O (2014) Forecasting exchange rates out-of-sample with panel methods and real-time data. *J Int Money Financ* 43(C):1–18
- Jorion P, Sweeney RJ (1996) Mean reversion in real exchange rates: evidence and implications for forecasting. *J Int Money Financ* 15(4):535–550
- Kilian L (1999) Exchange rates and monetary fundamentals: what do we learn from long-horizon regressions *J Appl Econ* 14(5):491–510
- Kilian L, Zha T (2002) Quantifying the uncertainty about the half-life of deviations from PPP. *J Appl Econ* 17(2):107–125
- Klau M, Fung SS (2006) The new BIS effective exchange rate indices. *BIS Quarterly Review*
- Lothian JR, Taylor MP (1996) Real exchange rate behavior: the recent float from the perspective of the past two centuries. *J Polit Econ* 104(3):488–509
- Mark NC (1995) Exchange rates and fundamentals: evidence on long-horizon predictability. *Am Econ Rev* 85(1):201–18
- Mark NC, Sul D (2001) Nominal exchange rates and monetary fundamentals: evidence from a small post-Bretton Woods panel. *J Int Econ* 53(1):29–52

- Mayoral L, Dolores Gadea M (2011) Aggregate real exchange rate persistence through the lens of sectoral data. *J Monet Econ* 58(3):290–304
- Meese R, Rogoff K (1983a) The out-of-sample failure of empirical exchange rate models: Sampling error or misspecification? In: Frenkel JA (ed) *Exchange rates and international macroeconomics*. NBER Chapters, p 67–112
- Meese RA, Rogoff K (1983b) Empirical exchange rate models of the seventies: do they fit out of sample? *J Int Econ* 14(1–2):3–24
- Meese RA, Rogoff K (1988) Was it real? The exchange rate-interest differential relation over the modern floating-rate period. *J Financ* 43(4):933–48
- Murray C, Papell D (2005) The purchasing power parity puzzle is worse than you think. *Empir Econ* 30(3):783–790
- Murray CJ, Papell DH (2002) The purchasing power parity persistence paradigm. *J Int Econ* 56(1):1–19
- Pesaran MH, Timmermann A (1992) A simple nonparametric test of predictive performance. *J Bus Econ Stat* 10(4):561–565
- Pippenger M, Goering G (1998) Exchange rate forecasting: results from a threshold autoregressive model. *Open Econ Rev* 9(2):157–170
- Rogoff K (1996) The purchasing power parity puzzle. *J Econ Lit* 34(2):647–668
- Rogoff K (2009) Exchange rates in the modern floating era: What do we really know *Review of World Economics (Weltwirtschaftliches Archiv)* 145(1):1–12
- Rossi B (2005) Confidence intervals for half-life deviations from purchasing power parity. *J Bus Econ Stat* 23:432–442
- Rossi B (2006) Are exchange rates really random walks? Some evidence robust to parameter instability. *Macroecon Dyn* 10(01):20–38
- Sarno L, Taylor MP (2002) *The economics of exchange rates*. Cambridge University Press, Cambridge
- Taylor AM, Taylor MP (2004) The purchasing power parity debate. *J Econ Perspect* 18(4):135–158
- Zumaquero A, Urrea R (2002) Purchasing power parity: error correction models and structural breaks. *Open Econ Rev* 13(1):5–26