Journal of International Money and Finance 142 (2024) 103026

Contents lists available at ScienceDirect

Journal of International Money and Finance

journal homepage: www.elsevier.com/locate/jimf

Exchange rate predictability: Fact or fiction?

Karen Jackson [a,*], Georgios Magkonis [b]

[a] Westminster Business School, University of Westminster, 35 Marylebone Road, London NW1 5LS, United Kingdom
[b] Faculty of Business and Law, University of Portsmouth, Portsmouth PO1 3DE, United Kingdom

ARTICLE INFO

JEL classification:
F31
F37
C83

Keywords:
Exchange rates
Forecasting performance
Meta-analysis

ABSTRACT

The present study investigates the factors that affect the forecasting performance of several models that have been used for exchange rate prediction. We provide a quantitative survey collecting 8,413 reported forecast errors and we investigate which forecasting characteristics tend to improve forecasting ability. According to our evidence, predictions can beat random walk when certain types of models and econometric methods are used. In particular, linear specifications based on PPP outperform random walk. Furthermore, higher data frequency and longer forecasting horizon also improve forecasting performance. In this way, we identify under which conditions it is feasible to solve the 'Meese-Rogoff' puzzle.

We should not expect the models to have much power to forecast changes in exchange rates. This might be disappointing news for forecasters working on Wall Street, but may be good news for open-economy macroeconomists. Engel et al. (2007)

1. Introduction

Exchange rates are one of the most central figures in forecasting. The need for accurate exchange rate forecasts is a primary concern for central banks, traders as well as academic economists. This need has been the incentive for investing in better modelling and the recent developments in the field of macroeconometrics. More elaborate models as well as advances in computer capabilities have provided more refined forecasting methodologies.

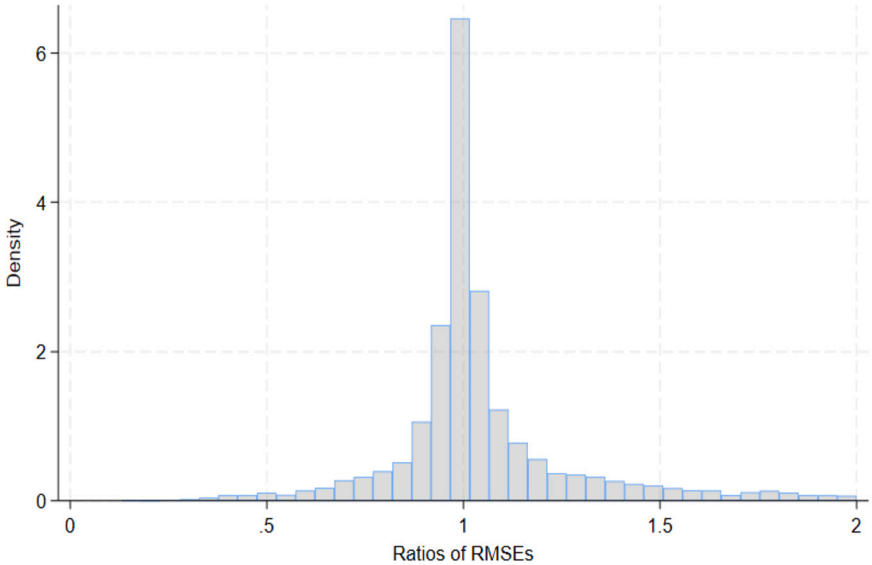
The empirical pattern of exchange rate forecasting, however, raises a lot of concerns. The so-called Meese-Rogoff puzzle is one of the most famous and influential concepts in international economics, according to which the random walk forecasts better than economic models. Since their seminal papers (Meese and Rogoff, 1983a,b) a large empirical literature has been developed exploring whether exchange rates can be predicted using different and more advanced economic models. In many cases, it has been proved that random walk can beat a range of structural models. As a reaction to this early evidence, the more recent literature has identified new predictors and new methods, claiming that forecasting the exchange rate is feasible.

The aim of this paper is a quantitative navigation of the existing literature over the last forty years. We explore whether there are factors that systematically provide better forecasts. And vice versa; to investigate whether there are factors that systematically provide worse forecasts. To achieve this aim, we distinguish among different models, econometric methods, forecasting horizons and periods as well as different data characteristics. Overall, the contribution is twofold. Firstly, we examine whether exchange rates are actually predictable, and identify the most useful factors to make these predictions. Secondly, instead of developing another horse race of forecasts, we collect data from the forecasting exercises conducted over the last forty years. We then use the most recently

* Corresponding author.
E-mail addresses: k.jackson@westminster.ac.uk (K. Jackson), georgios.magkonis@port.ac.uk (G. Magkonis).

https://doi.org/10.1016/j.jimonfin.2024.103026

Available online 15 February 2024



Note: The figure depicts the histogram of the relative forecasting root mean squared errors, ϕ reported by individual studies.

Fig. 1. Histogram of RMSEs.

developed techniques to perform a quantitative literature survey. Therefore, the present study can be read as a companion to the traditional surveys on exchange rate forecasts.

The remainder of the paper is structured as follows. Section 2 describes the data collection process and discusses the potential factors that affect the forecasting performance. Section 3 analyses the econometric methodology, while Section 4 presents the main results along with a series of robustness checks. Section 5 concludes.

2. Data collection

The most frequently used and reported metric in the forecasting literature is the relative root mean squared forecast error (RMSFE). Following Eickmeier and Ziegler (2008) and Filippidis et al. (2024), the main variable of interest in our study is the ratio of RMSFEs. This ratio compares the forecasting performance of a model relative to a benchmark one. The most common benchmark model across the literature is random walk; $\phi_{ij} = (RMSFE)/(RMSFE^{Bench})$, where the superscript *Bench* indicates the benchmark model. If the ratio is higher (lower) than 1, then random walk has a better (worse) predictive ability.[1] Fig. 1 shows the histogram of the collected data. In the remainder of this section, we analyse in detail all the steps of the data collection.

The data collection process follows the guidelines developed by Havránek et al. (2020). The starting point of our data collection process is the survey of Rossi (2013) from where we collect 59 papers. Then, we move to Google scholar using the key combinations ‘exchange rate predictability’, ‘exchange rate forecasting’, ‘exchange rate random walk’ and ‘exchange rate modelling’. In order to ensure that a key paper was not missed, we also checked the references of older important papers as well as other survey studies on exchange rate forecasting (Frankel and Rose, 1995; Sarno and Valente, 2009; Melvin et al., 2013; Wieland and Wolters, 2013; Kavtaradze and Mokhtari, 2018; Fang et al., 2023). Through this search, we are able to find 62 additional papers.

The next step is to impose a number of criteria according to which a study and a reported estimate can be used. Here, we impose three criteria. The first one requires a study to report at least one RMSFE. Almost all collected studies contained an empirical section. However, some of the papers report alternative evaluation metrics for the relative forecasting performance, such as the direction of change criterion, the consistency criterion and MSFE differences (Cheung and Chinn, 1998; Cheung et al., 2005; Giacomini and Rossi, 2010). Thus, papers that do not report any RMSFEs are excluded.

Secondly, our dataset excludes all the RMSFEs that consider a benchmark model other than random walk. This is not a strictly restrictive criterion since all the collected papers are using random walk as the benchmark model. Therefore, we do not exclude any paper; we only omit a small number of reported estimates that are based on other benchmark models (such as autoregressive models).

Thirdly, we impose a quality threshold in order to ensure that our collected studies contain a solid empirical analysis, avoiding papers that simply report results from horse races of several forecasting exercises without a prevalent economic context. In this way, we exclude papers that are published in journals either classified as 1-star in the ABS list or not classified.[2] Fig. 2 summarises the process in a PRISMA chart. Additionally, Table 1 reports the studies that constitute our dataset along with some main characteristics, as reported in Table 2 in Rossi (2013). These characteristics are discussed in the next section.

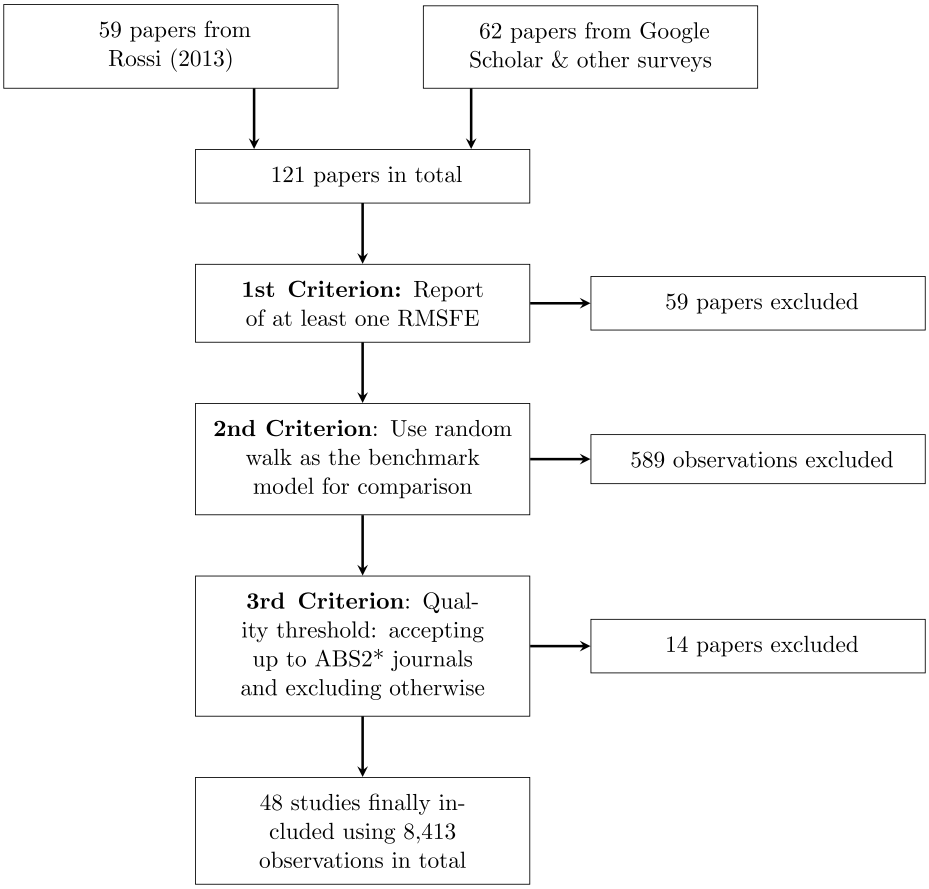
[1] We use the symbol ϕ and the term ‘relative forecasting performance’ interchangeably throughout the text.

[2] NBER papers are treated as published high quality papers and therefore are included.

Table 1
Collected papers and main characteristics.

Paper	Models & Pedictors		Data		
Abhyankar et al. (2005)	(V)ECM	m,y	M	1977-2000	CA,JP,UK
Bergin (2003)	Structural	y,p,CA	Q	1973-1996	AUS,CA,UK
Berkowitz & Giorgianni (2001)	(V)ECM	m,y	Q	1973-1991	CA,GE,SWI,JP
Byrne et al. (2016)	Linear/Nonlinear	i,π,y(gap)	Q	1973-2013	AUS,CA,DE,UK,JP,KO,NO,SWE, SWI,AUT,BE,FR,GE,SP,IT,FI,NE
Byrne et al. (2017)	Model averaging	m,y,i,π,p,y(gap)	M	1977-2013	AUS,CA,UK,JP,NO,SWE,SWI,EU
Canova (1993)	Nonlinear	i	W	1979-1987	FR,SWI,GE,UK,JP
Carriero et al. (2009)	Structural	s	M	1995-2008	AUS,BR,CA,CO,CHI,DE,EU,FI,UK,JP, CZ,HU,IN,IR,IS,ME,NO,SWE,SWI,TH, MA,PA,PE,PH,PL,SI,SL,TU,UR
Ca'Zorzi et al. (2016)	Linear	p	M	1975-2012	AUS,CA,EU,JP,ME,NZ,SWI,UK
Ca'Zorzi et al. (2017)	Structural	m,y	Q	1975-2013	EUR,UK,CA,AUS
Ca'Zorzi and Rubaszek (2020)	Linear	p	M	1975-2017	AUS,CA,JP,NZ,SWI,UK,EUR,
Ca'Zorzi et al. (2022)	Linear	y,π, CA	Q	1975-2018	AUS,CAN,CHE EA,JPN,NOR,NZL,SWE
Cheung et al. (2005)	(V)ECM	m,y,i,π	Q	1973-2000	CA,UK,GE,JP,SWI
Cheung et al. (2019)	(V)ECM	m,y,i,π,p	M	1973-2014	CA,GE,JP,UK
Chinn (1991)	Linear/(V)ECM	m,y,i,π,w	Q	1974-1988	GE,JP
Chinn & Meese (1995)	Linear/(V)ECM	m,y,i,π,tb	M	1973-1990	GE,CA,UK,JP
Chinn & Moore (2011)	(V)ECM	m,y,i,p,of	M	1999-2001	EU,JP
Clarida & Taylor (1997)	(V)ECM	F	W	1977-1993	UK,GE,JP
Engel (1994)	Linear/Nonlinear	F,s	Q	1973-1991	FR,IT,JP,SWI,UK,GE
Eichenbaum et al. (2021)	Nonlinear	p	Q	1982-2010	AUS,CA,GE,NZ,SWE,UK
Engel and Hamilton (1990)	Nonlinear	s,i	Q	1973-1988	GE,FR,UK
Engel and Wu (2018)	Linear	i	M	1999-2017	AUS,CAN,EUR,JAP,NZL,NOR, SWE,SWI,UK
Engel et al. (2007)	(V)ECM	m,y	Q	1973-2005	AUS,AUT,BE,CA,DE,FI,FR,GE,GR UK,IT,JP,KO,NE,NO,SP,SWE,SWI
Engel et al. (2015)	Linear factor	m,y,i,π,p	Q	1973-2007	AUS,AUT,BE,CA,DE,FI,FR,GE, UK,IT,JP,KO,NE,NO,SP,SWE,SWI
Faust et al. (2003)	(V)ECM	m,y	Q	1977-1991	CA,GE,JP,SWI
Garratt & Mise (2014)	Model averaging	m,i,p,π,F,y(gap)	Q	1990-2008	EU,UK,JP,CA,DE,NO,SWE,SWI,AUS,KO
Groen (1999)	(V)ECM	m,y	M	1973-1994	GE,FR,NE,CA
Groen (2005)	(V)ECM	m,y	Q	1975-2000	EMU,GE
Ince (2014)	Linear	π,p,y(gap)	Q	1977-2009	AUS,CA,FR,GE,IT,JP,NE,SWE,UK
Lam et al. (2008)	Linear/moving averaging	p,π,i,m,y,y(gap)	Q	1973-2007	UK,JP
Lopez-Suarez & Rodríguez-Lopez (2011)	Nonlinear	p	Q	1973-2009	AUS,AUT,BE,CA,DN,FI,FR,GE,GR UK,IT,JP,KO,NE,NO,SP,SWE,SWI
MacDonald & Taylor (1993)	(V)ECM	m,y,i	M	1976-1990	GE
Mark (1995)	(V)ECM	m,y	Q	1973-1991	CA,GE,SWI,JP
Mark & Sul (2001)	(V)ECM	m,y,p	Q	1973-1997	AUS,AUT,BE,CA,DE,FI,FR,GE,GR, IT,JP,KO,NE,NO,SP,SWE,SWI,UK
Meese (1988)	Linear	i,π	M	1980-1986	GE,JP,UK
Meese and Rogoff (1983a)	Linear/Structural	m,y,i,π,F,tb	M	1973-1981	GE,UK,JP
Meese and Rogoff (1983b)	Linear/Structural	m,y,i,π,tb	M	1973-1981	GE,UK,JP
Meese & Rose (1991)	Nonlinear	m,y,i,π,tb	M	1974-1987	UK,CA,GE,JP
Mizrach (1992)	Nonlinear	s	D	1974-1988	FR,IT,GE
Moosa & Burns (2014)	Linear/(V)ECM/Nonlinear	m,y,i,s	M	1998-2013	CA,UK,JP
Qi & Wu (2003)	Nonlinear	m,y,i	M	1973-1997	JP,GE,UK,CA
Rapach & Wohar (2006)	Nonlinear	s	M	1980-2003	UK,GE,FR,JP
Rossi (2005)	(V)ECM/Structural	m,y,i,π	Q	1973-1998	GE,JP
Rocci (2006)	Linear/Nonlinear	m,y,i	M	1973-1998	CA,FR,GE,IT,JP
Rogoff & Stavrakeva (2008)	(V)ECM	i,π,y(gap)	Q	1973-2006	UK,DE,FR,GE,NE,CA,JP,AUS, IT,SWI,SWE,PO
Rogoff$_{0}$8/Molodtsova & Papell (2008)	(V)ECM	π,y(gap)	Q	1973-2005	UK,DE,FR,GE,NE,CA,JP,AUS, IT,SWI,SWE,PO
Schinasi & Swamy (1989)	Linear	m,y,i,tb	M	1973-1980	UK,JP,GE
Wolff (1987)	Nonlinear	m,y,i,π	M	1973-1984	UK,JP,GE
Wright (2008)	Model averaging	m,y,i,p,z,SP,D,CA	M/Q	1973-2005	CA,GE,JP,UK KO,NO,SWE

Notes: Regarding the econometric methods, (V)ECM represents for (vector) error correction models. For the predictors we use the following codes: 'm' for money aggregates, 'i' for interest rates, 'y' for output, 'y(gap)' for output gap, 'p' for price level, 'π' for inflation, 'CA' for current account, 'F' for forward discount, 'D' for dividend, 'z' for productivity, 's' for lagged exchange rate, 'of' for order flow, 'SP' for stock price, 'tb' for trade balance and 'w' for wealth. Regarding the frequency, M stands for monthly data, Q for quarterly, W for weekly and D for daily. The country codes are as follows: AUS = Australia, AUT = Austria, BE = Belgium, BR = Brazil, CA = Canada, CH = Chile, CZ = Czechia, DE = Denmark, EMU/Euro = Euro area, GE = Germany, GR = Greece, FI = Finland, FR = France, HU = Hungary, IN = India, IR = Ireland, IS = Israel, IT = Italy, JP = Japan, KO = South Korea, MA = Malta, ME = Mexico, NE = Netherlands, NO = Norway, NZ = New Zealand, PA = Pakistan, PE = Peru, PH = Philippines, PL = Poland, PO = Portugal, SL = Slovakia, SI = Singapore, SP = Spain, SWE = Sweden, SWI = Switzerland, TH = Thailand, TU = Turkey, UK = United Kingdom, UR = Uruguay.



Note: The figure summarises the process of data collection.

Fig. 2. PRISMA chart.

3. Heterogeneity of forecast errors

In order to examine the heterogeneity of the reported RMSFEs in detail, we look into five categories. For each category we identify factors that may systematically influence the reported RMSFEs. In what follows we describe this process.

Models. Following Cheung et al. (2005), Cheung et al. (2019), Rossi (2013), we distinguish seven broad families of models; i) monetary, ii) purchasing power parity (PPP), iii) Taylor rule (Taylor), iv) behavioural equilibrium exchange rate (BEER), v) uncovered interest rate parity (UIP), vi) real interest rate parity (real) and vii) models that combine several of the above mentioned models (mixed). Focusing on models, instead of specific variables, provides a better economic rational.

Econometric methods. Regarding the econometric techniques that have been used, we distinguish among the following choices. The first and most straightforward category is the one of linear models. This family of models (*linear*) is the reference category. In this category, we do not include error correction models (ECM). Since ECM have been extensively used in the forecasting literature, we treat them as a separate category. Therefore, we create a separate variable, *(V)ECM*, which takes 1 when an error or a vector error correction model is used and 0 otherwise. Next, we take into account the non-linear models, such as smooth transition models, threshold models, Markov switching etc. Furthermore, we take into account that many forecasting exercises use structural models. With the term ‘structural’ we mean both DSGE and VAR models. So, in a similar vein, we create the variable *structure* that takes 1 when the reported ϕ comes from these kinds of structural models. Finally, we treat as a separate category all the methodologies that are based on any kind of model averaging methodologies. The distinguishing characteristic of this family of techniques is that they deal with the concept of model uncertainty. In the next section, we discuss this issue in more technical detail. Another characteristic that we take into account is the distinction between panel data vs. individual time series. One of the key issues is related to estimation error (Ca’Zorzi et al., 2016). Therefore, studies that use panel data may tend to deliver higher accuracy than models using individual country series.

Data characteristics. Another important feature of exchange rate prediction is the data frequency. Throughout the exchange rate forecasting literature there is a trade-off between the frequency and the span of the dataset. The majority of the studies use monthly and quarterly data, with the later covering shorter samples. More recent studies, however, tend to use higher frequencies. Therefore, we distinguish between two categories; quarterly and weekly/daily frequencies, with monthly data frequency being the reference

Table 2
List of Moderator Variables.

Variable Name	Description	Mean	SD
ϕ	Ratio of root mean square forecast errors as in equation (1)	1.042	0.231
Models			
Monetary	1 if monetary models (base)	0.203	0.403
PPP	1 if PPP model is used	0.136	0.343
Taylor	1 if Taylor-rule model is used	0.170	0.376
BEER	1 if behavioural model is used	0.010	0.097
UIP	1 if uncovered parity model is used	0.028	0.167
RIRP	1 if real parity model is used	0.011	0.106
Mixed	1 if a mixed model is used	0.143	0.350
Econometric Methods			
Linear	Linear models (base)	0.207	0.405
Non-linear	1 if non-linear models	0.172	0.377
Structural	1 if a structural model	0.254	0.435
(V)ECM	1 if an ECM or VECM model	0.192	0.394
Model Averaging	1 if a model averaging model	0.171	0.377
Panel	1 if panel data is used	0.547	0.497
Data Characteristics			
Monthly	monthly frequency (base)	0.406	0.472
Daily/Weekly	1 if daily/weekly frequency	0.024	0.154
Quarterly	1 if quarterly frequency	0.479	0.499
Emerging	1 if emerging economy	0.163	0.369
Forecasting Characteristics			
Forecasting horizon	number of months	13.750	15.681
Forecasting period	average period of the forecast	20.548	7.712

Note: The table shows the moderator variables included in this study. SD stands for standard deviation.

category.[3] Finally, an interesting aspect is the heterogeneity of reported forecast errors according to which economies are examined. Due to the large number of countries, we separate between developed and emerging economies.

Forecasting characteristics. We consider two features of the forecasting exercise. The first is the forecasting horizon. As analysed above, different studies use different data frequencies. Therefore, the forecasting horizon is expressed in various ways. To overcome this problem, we convert the horizons into months in order to obtain a homogeneous measure across all studies. The second characteristic that we take into account is the average forecasting period used for each forecasting exercise. The earliest starting year is 1974 in order to avoid the Bretton-Woods period of fixed exchange rates. In this way, we examine whether there is a trend in the reported results. A detailed description and summary statistics for each variable used in our analysis is reported in Table 2. Overall, this design provides 16 different explanatory variables (see Table A.1 for multicollinearity test). In the robustness section, we examine whether there is any publication bias by taking into account the quality of the collected studies. Furthermore, we examine whether papers published after Rossi (2013) have improved their forecasting performance.

4. Econometric framework and results

In order to identify the factors that systematically affect the forecasting performance, we adopt a simple regression model estimated by a model averaging. The estimated equation is written as:

$$ln(\phi_{ij}) = c^{\theta} + \sum_{S=1}^{K} \beta_S^{\theta} X_{S,ij} + \epsilon_{ij} \tag{1}$$

where $ln(\phi_{ij})$ is natural logarithm of the relative forecasting performance i from study j, matrix X contains the moderator variables, β_S are the coefficients of each moderator, K is the number of regressors, while $\epsilon \sim N(0,\sigma)$. i is an index for a regression estimate from the jth study. The superscript θ indicates that the above equation is valid under model M_{θ}. With a maximum of 2^K possible models, the model space consists of $M_1,...,M_{\theta}$ models, with $\theta \in [1,...,2^K]$. We use the logarithm of the dependent variable in order to avoid unintended nonlinearity into the relationship between the regressors and RMSFE. The rationale behind the decision to use model averaging is to tackle model uncertainty. Even with a moderate number of explanatory variables, the total number of models could be quite large. The traditional 'general-to-specific' approach, which is based on removing the most insignificant variable one by one, may lead to erroneous results. The benefit of model averaging is that it identifies the most frequently statistically significant variables by examining a large model space. In this way, the analysis and the results are more robust. Among the several model averaging

[3] A related idea that captures other aspects of data characteristics would be the distinction of data transformations such as seasonal adjustment, detrending, filtering as well as real vs. final vintage data and rolling vs. expanding window. However, the majority of the collected papers do not report these details.

Table 3
Baseline results.

Variable	BMA$_1$ PIP	BMA$_1$ Coefficient	BMA$_1$ SD	BMA$_2$ PIP	BMA$_2$ Coefficient	BMA$_2$ SD
Models						
PPP	**0.999**[a]	**-0.069**	**0.009**	**0.999**[a]	**-0.068**	**0.008**
Taylor	0.108	0.002	0.006	0.115	0.002	0.007
BEER	0.073	0.003	0.015	0.079	0.004	0.016
UIP	0.166	-0.006	0.015	0.164	-0.006	0.016
RIRP	**0.930**[c]	**0.082**	**0.032**	**0.917**[c]	**0.081**	**0.033**
Mixed	0.052	0.000	0.003	0.051	0.000	0.004
Econometric Methods						
Nonlinear	**0.999**[a]	**0.053**	**0.007**	**0.999**[a]	**0.053**	**0.007**
Structural	**0.999**[a]	**0.058**	**0.009**	**0.999**[a]	**0.058**	**0.009**
(V)ECM	0.012	0.000	0.000	0.013	0.000	0.000
Model Averaging	0.030	0.000	0.002	0.031	0.000	0.003
Panel	**0.999**[a]	**-0.033**	**0.006**	**0.999**[a]	**-0.033**	**0.006**
Data Characteristics						
Daily/Weekly	**0.998**[b]	**-0.089**	**0.019**	**0.998**[b]	**-0.090**	**0.019**
Quarterly	0.210	-0.003	0.007	0.214	-0.003	0.007
Emerging	0.018	0.000	0.001	0.019	0.000	0.001
Forecasting Characteristics						
Forecasting Horizon	**0.999**[a]	**-0.002**	**0.000**	**0.999**[a]	**-0.002**	**0.000**
Forecasting period	0.017	0.000	0.000	0.018	0.000	0.000

Notes: PIP stands for posterior inclusion probabilities. The BMA$_1$ shows the results using uniform and UIP as model and parameters priors, respectively. BMA$_2$ shows the results using random and BRIC as model and parameters priors, respectively. a/b/c denotes decisive/strong/positive evidence of a regressor having an effect, respectively, according to Kass and Raftery (1995).

schemes, we adopt the Bayesian version, as it has become the dominant approach. In appendix B, we provide more technical details about the estimation process as well as the different prior distributions that we use in order to ensure the robustness of our results. More precisely, we analyse the different combinations of parameters and models' priors.

Models. Table 3 shows the first round of results. To increase their readability the variables with high PIPs are shown in bold. Following Kass and Raftery (1995), the effect of a variable is considered as weak, positive, strong, and decisive if its PIP lies between 0.5-0.75, 0.75-0.95, 0.95-0.99 and 0.99-1, respectively. Beginning from the choice of models, our results suggest that the use of PPP model tends to reduce the relative forecast errors. This result is in accordance with the evidence provided by Rogoff (1996) according to which PPP exchange rates tend to move towards the PPP in the long run. According to this outcome, whenever there is a predictive failure of PPP, it is due to short run deviations. The potential explanations for these deviations have been extensively discussed in the literature (Cheung and Lai, 2000; Murray and Papell, 2002; Imbs et al., 2005). On the other hand, models that are based on real interest rate parity tend to increase the relative forecast errors. This empirical result is in accordance with evidence presented in Meese and Rogoff (1988) and, more recently in Alquist and Chinn (2008), where the predictive ability of interest rates does not seem to beat random walk.

Econometric methods. As mentioned in Section 3, we use the linear model as our benchmark category. In this category, we include all models that use a single equation of the form $E_t(s_{t+h} - s_t) = \beta_0 + \beta_1 p^*$ with the predictor p^* receiving either realized contemporaneous or lagged values. Our results indicate that the use of nonlinear models tends to produce higher relative forecast errors. This outcome is consistent with the earlier evidence that nonlinear models prove to be the least successful (Teräsvirta, 2006). More specifically, Sarantis (1999), Clements and Smith (2001) and Boero and Marrocu (2002) report evidence according to which threshold and smooth-transition models cannot increase the predictive ability against random walk. These two broad range of models are included in the *nonlinear* moderator variable described above. However, this variable takes into account additional categories of nonlinear modelling. For instance, part of the reported RMSFEs comes from other model families, such as time-varying coefficients and markov-switching. Prior evidence is also in accordance with our results; Rossi (2006) reports very limited predictive ability for time-varying models. Bacchetta et al. (2009) argue that parameters' instability cannot explain the Meese-Rogoff puzzle. Finally, inconclusive results are also reported in the forecasting exercise of Ferrara et al. (2015), where linear models beat several nonlinear specifications. Overall, these outcomes regarding modeling and econometric estimations suggest that simple linear regressions assuming PPP works best. This supports the results of Ca'Zorzi and Rubaszek (2020), Eichenbaum et al. (2021) and Engel and Wu (2023). From this perspective, it is not surprising that structural econometric modelling does not seem to offer better forecasts. Finally, our initial hypothesis that panel data should produce higher accuracy than individual series is confirmed.

Data characteristics. Higher (daily or weekly) frequency improves the forecasting ability. Usually, the studies exploiting higher frequency datasets are more able to reduce the forecast error of each model (Ferraro et al., 2015). In our study, we did not take into account RMSFEs coming from annual data, which are the least frequently used in the literature. The only paper that passed the three inclusion criteria discussed earlier is the study of Rapach and Wohar (2002). Due to the fact that they report only seven RMSFEs, we eventually do not include it in the final sample. As far as the country data coverage is concerned, our results suggest that there is no difference in the forecasting performance between developed and emerging economies.

Table 4
Robustness checks.

Variable	BMA_3 Mean	BMA_3 SD	BMA_4 Mean	BMA_4 SD	OLS Mean	OLS SD
Models						
PPP	**-0.061**[a]	**0.010**	**-0.059**[a]	**0.010**	-0.057**	0.026
Taylor	0.017	0.012	0.022	0.011	0.027	0.023
BEER	0.036	0.035	0.054	0.032	0.067	0.084
UIP	-0.020	0.021	-0.026	0.019	-0.029	0.030
RIRP	**0.088**[a]	**0.024**	**0.087**[a]	**0.023**	0.091***	0.025
Mixed	-0.001	0.006	0.001	0.009	0.007	0.025
Econometric Methods						
Nonlinear	**0.050**[a]	**0.008**	**0.050**[a]	**0.008**	0.052**	0.020
Structural	**0.056**[a]	**0.010**	**0.056**[a]	**0.011**	0.058**	0.027
(V)ECM	0.001	0.004	0.003	0.007	0.007	0.026
Model Averaging	-0.002	0.006	-0.005	0.009	-0.009	0.027
Panel	**-0.036**[a]	**0.007**	**-0.037**[a]	**0.007**	-0.038*	0.024
Data Characteristics						
Daily/Weekly	**-0.085**	**0.019**	**-0.084**	**0.019**	-0.087	0.067
Quarterly	**-0.016**[c]	**0.009**	**-0.019**[b]	**0.008**	-0.022	0.025
Emerging	0.001	0.005	0.003	0.007	0.008	0.010
Forecasting Characteristics						
Forecasting Horizon	**-0.002**[a]	**0.000**	**-0.002**[a]	**0.000**	-0.002***	0.001
Forecasting period	0.000	0.000	0.000	0.000	0.000	0.001

Notes: The BMA_3 shows the results using uniform and EBL as model and parameters priors, respectively. BMA_4 shows the results using random and EBL as model and parameters priors, respectively. a/b/c denotes decisive/strong/positive evidence of a regressor having an effect, respectively, according to Kass and Raftery (1995). For the OLS, clustered standard errors at study level are reported. ***, ** and * indicate statistical significance at 1%, 5% and 10%, respectively.

Forecasting characteristics. The distinction between short-run and long-run prediction is one of the factors that explains the reported variation. Specifically, as we move from shorter to longer horizons, the predictive ability is improved. Mark (1995) and Cheung et al. (2005) report evidence in favour of the long-horizon predictability of the monetary fundamentals. On the contrary, the forecasting period does not systematically influence the reported RMSFEs. Therefore, there is no difference in the forecasting performance between the primary studies that use the first decade of the post-Bretton Woods system and later studies that use the late 2000s-early 2010s as the forecast period.

5. Robustness and further evidence

As a first robustness check we use different combinations of parameters' and models' priors. Additionally, we also apply (the pure frequentist exercise of) least squares. Table 4 shows the results. Both types of estimation lead to results that are quantitatively and qualitatively the same with the benchmark results. The only difference is that lower (quarterly) frequency data tends to improve the forecasting ability. This can be explained by the fact that lower frequencies contain more information about potential long-run trends as they exploit longer spans of data (Rossi, 2013). Overall, the moderator variables that were found to be robust drivers of the observed heterogeneity in the literature remain the same.[4]

In order to make our analysis more inclusive, we extend equation (1) by adding two more variables. The first variable accounts for potential publication bias. Specifically, we add the dummy variable 'ABS4*' that takes 1 when the study is published in a 4-star journal according to the ABS list and 0 otherwise. Secondly, in order to test whether forecasters have learnt something important from the influential study of Rossi (2013), we create a new dummy variable ('afterRossi') which takes 1 when the study has been published after the publication of Rossi. As an extra robustness check, we estimate the extended version with BMA, OLS as well as with least absolute shrinkage and selection operator (LASSO), an alternative method to tackle model uncertainty. A technical description is provided in Appendix B. The new outcomes are very similar to the previous set of results. Interestingly, we find evidence of a publication bias, supported by both BMA and LASSO; studies published in top journals tend to report better forecasts. However, published before or after Rossi (2013) does not play a role. (See Table 5.)

The final exercise is to focus only on the subset of observations that use the US dollar as the peer currency. In this way, we test whether our results are driven by the mean-reversion of the USD rate against other currencies (Engel and Wu, 2023). Our findings in Table 6 show that this is not the case, as the new estimates are not qualitatively or quantitatively different from the previous sets of results.

[4] The results are also very similar when we use a frequentist model averaging, instead of Bayesian. Therefore, the results remain robust against different averaging methods.

Table 5
Further evidence.

	BMA		OLS		LASSO	
Variable	Mean	SD	Mean	SD	Mean	SD
Models						
PPP	**-0.064**[a]	**0.011**	-0.051*	0.025	**-0.049**	**0.004**
Taylor	0.012	0.017	0.038	0.033	0.000	0.000
BEER	0.008	0.024	0.074	0.079	0.000	0.000
UIP	-0.005	0.014	-0.023	0.029	0.000	0.000
RIRP	**0.086**[b]	**0.032**	0.104***	0.024	**0.097**	**0.011**
Mixed	0.000	0.001	-0.001	0.024	0.000	0.000
Econometric Methods						
Nonlinear	**0.047**[a]	**0.008**	0.045**	0.020	**0.046**	**0.003**
Structural	**0.054**[a]	**0.009**	0.053*	0.027	**0.061**	**0.004**
(V)ECM	-0.006	0.035	-0.006	0.030	0.000	0.000
Model Averaging	0.000	0.002	0.001	0.024	0.000	0.000
Panel	**-0.047**[a]	**0.007**	-0.049**	0.023	**-0.049**	**0.003**
Data Characteristics						
Daily/Weekly	**-0.096**[a]	**0.018**	-0.090	0.062	**-0.079**	**0.008**
Quarterly	-0.029	0.067	-0.014	0.026	0.000	0.000
Emerging	0.000	0.001	0.000	0.011	0.000	0.000
Forecasting Characteristics						
Forecasting Horizon	**-0.002**[a]	**0.000**	-0.002*	0.001	**0.002**	**0.000**
Forecasting period	0.017	0.000	0.000	0.018	0.000	0.000
Publication Characteristics						
ABS4*	**-0.030**	**0.007**	-0.032	0.021	**-0.027**	**0.004**
After Rossi	-0.012	0.014	-0.031	0.020	0.000	0.000

Notes: For the BMA uniform and UIP are used as model and parameters priors, respectively. The same notation is used as explained in Tables 3 and 4. For LASSO estimates, the variables that are found significant are shown in bold.

Table 6
Further evidence using only USD.

	BMA_1		BMA_2		OLS	
Variable	Mean	SD	Mean	SD	Mean	SD
Models						
PPP	**-0.067**[a]	**0.009**	**-0.066**[a]	**0.009**	-0.049*	0.026
Taylor	0.003	0.010	0.004	0.011	0.039	0.038
BEER	0.004	0.017	0.005	0.019	0.081	0.082
UIP	-0.007	0.016	-0.006	0.016	-0.021	0.029
RIRP	**0.078**[c]	**0.034**	**0.077**[c]	**0.036**	0.102***	0.024
Mixed	0.000	0.001	0.000	0.001	0.007	0.025
Econometric Methods						
Nonlinear	**0.050**[a]	**0.007**	**0.050**[a]	**0.008**	0.046**	0.020
Structural	**0.062**[a]	**0.008**	**0.062**[a]	**0.008**	0.061**	0.030
(V)ECM	0.000	0.002	0.000	0.003	-0.006	0.031
Model Averaging	0.000	0.001	0.000	0.002	0.004	0.024
Panel	**-0.050**[a]	**0.007**	**-0.050**[a]	**0.007**	-0.052**	0.023
Data Characteristics						
Daily/Weekly	**-0.093**[a]	**0.018**	**-0.093**[a]	**0.018**	-0.089	0.062
Quarterly	-0.007	0.003	-0.008	0.003	-0.013	0.025
Emerging	0.000	0.001	0.000	0.001	-0.001	0.012
Forecasting Characteristics						
Forecasting Horizon	**-0.002**[a]	**0.001**	**-0.002**[a]	**0.001**	-0.002**	0.001
Forecasting period	0.000	0.000	0.000	0.000	0.000	0.001
Publication Characteristics						
ABS4*	**-0.027**[b]	**0.008**	**-0.027**[b]	**0.008**	-0.030	0.022
After Rossi	-0.002	0.007	-0.003	0.008	-0.026	0.020

Notes: BMA_1 and BMA_2 are explained in Table 3. For the OLS, clustered standard errors at study level are reported.

6. Conclusions

Our meta analysis of predicting models contributes to the practice of exchange rate forecasting. Using a dataset that covers roughly forty years of different models, econometric techniques, datasets, horizons as well as several economies, we identify the most important drivers of forecasting accuracy. Based on the most frequently used metric of forecasting performance, the relative RMSFEs, we conclude that the answer to our paper's subtitle is the following: 'it depends'. Firstly, the choice of the model plays an important role; PPP models provide systematically better forecasts; the opposite is true for models based on real interest rate parity. Secondly, linear beat nonlinear specifications. Also, structural econometric techniques are not found to be better than random walk. This can be viewed as a reflection of the developments reported in the studies of Rogoff (1996) and Taylor and Taylor (2004). Thirdly, high-frequency data also tends to improve the forecasting performance. Fourthly, the forecasting horizon is also one of the driving forces that explains the variation of the reported RMSFEs, with more long-run horizons improving forecasting ability. Finally, we find evidence that studies published in 4* ABS journals tend to report better forecasts.

The aim of this study is not to provide a golden rule of exchange rate forecasting, such as Armstrong et al. (2015). On the contrary, our target is that this quantitative survey can be read in parallel with traditional surveys; Frankel and Rose (1995), Cheung et al. (2005), Engel et al. (2007) and Rossi (2013). For almost forty years, a vast number of forecasting exercises have been considered as successful, or not, based on their performance relative to random walk. However, this is only one aspect that may underestimate the contribution of newly developed models. An extension to the current study would be to consider alternative forecasting metrics. Also, the distinction between random walk and alternative models used as benchmark modelling would provide further insights into forecasting practice. Finally, looking into in-sample forecasts is another avenue for further examination. We leave these issues for future research.

Declaration of competing interest

None.

Data availability

Data will be made available on request.

Appendix A. Multicollinearity test

Table A.1
Testing for multicollinearity-VIFs.

Variable	VIF
Models	
PPP	1.73
Taylor	2.28
BEER	1.13
UIP	1.28
RIRP	1.05
Mixed	3.70
Econometric Methods	
Nonlinear	1.66
Structural	4.06
(V)ECM	2.00
Model Averaging	4.10
Panel	2.11
Data Characteristics	
Daily/Weekly	1.38
Quarterly	2.43
Developed	1.74
Forecasting Characteristics	
Forecasting Horizon	1.28
Forecasting period	1.95
Mean VIF	2.12

Notes: The table shows the individual and the average variance inflation factors.

Appendix B. Methodology

The rationale behind every averaging scheme is the assignment of a probability to each model. Under the Bayesian set-up, this probability is the prior model probability that is updated based on data. The distinguishing feature is that the inference is based on the weighted average across all models, rather than on an individual one. Using Bayes rule, the posterior density of β is written as:

$$p(\beta|\phi, X, M_i) \propto p(\phi|\beta, X, M_i)p(\beta|M_i) \tag{2}$$

where $p(\phi|\beta, X, M_i)$ is the likelihood function and $p(\beta|M_i)$ is the prior density. Considering M_i as an additional parameter whose posterior has to be estimated using data and a prior distribution, Bayes rule can be written as:

$$p(M_i|\phi, X) \propto p(\phi|M_i, X)p(M_i) \tag{3}$$

The left-hand side term is the posterior model probability (PMP) and the right-hand side term is the marginal likelihood function times the prior probability of model M_i. To determine this posterior density, the calculation of the marginal likelihood is needed. If we combine equations (3) and (4), the likelihood function can be written as:

$$p(\phi|M_i, X) = \int_{\beta} p(\phi|\beta, X, M_i)p(\beta|M_i, X)d\beta \tag{4}$$

Therefore, the posterior distribution of β is given by

$$p(\beta|\phi, X) = \sum_{j=1}^{2^K} p(\beta|\phi, X, M_j)p(M_j|\phi, X) \tag{5}$$

where $p(\beta|\phi, X, M_j)$ is the posterior distribution under model M_j and $p(M_j|\phi, X)$ is the posterior model probability, which is used as weight. Specifically, the posterior density of β for each model M_i is weighted by the posterior model probability of each model M_i. For benchmark results, $K = 16$, while for the robustness checks we increase K to 18 variables.

The frequency which with a regressor appears in model M_j determines whether the regressor can be considered a robust driver. This is the notion of posterior inclusion probability (PIP), which is the sum of posterior model probabilities of all the models that include the specific regressor;

$$PIP_i = \sum_{j=1} p(M_j|\phi, X) \tag{6}$$

with each regressor having a specific prior inclusion probability for $i \in [1, 15]$. The higher the PIP of a variable, the greater its explanatory power. This means that the variable with the highest estimated PIP is found to be present in almost all the alternative models and, therefore, constitutes a robust driver of the relative forecasting performance.

B.1. Parameter priors

Regarding the parameters, we have to specify the prior distributions for c, β and σ. We follow the standard convention of assuming non-informative priors for the intercept and the variance; that is, $p(c) \propto 1$ and $p(\sigma) \propto \sigma^{-1}$. As far as the parameters (β) are concerned, we follow Zellner (1986) by assuming that they are centered at zero and that the variance is proportional to $\sigma^2(g(X_i X_i)^{-1})$, where g is the so-called Zellner's g hyperparameter which indicates the level of our uncertainty. The smaller the g, the smaller the prior coefficient variance and, therefore, the lower our uncertainty. So, the coefficients' distribution depends on g:

$$\beta_i|g \sim N(0, \sigma^2(g(X_i X_i)^{-1}) \tag{7}$$

Here, we employ two different choices regarding g:

- Firstly, we set $g = n$, which leads to the most trivial case of unit information prior (UIP), where n is the sample size.
- Secondly, we employ the hyper-g prior as suggested by Liang et al. (2008). Specifically, $\frac{g}{1+g} \sim Beta(1, \frac{a}{2} - 1)$, where $\alpha \in (2,4]$ with a beta distribution mean equal to $\frac{2}{\alpha}$.
- Thirdly, we employ the 'empirical Bayes' (EBL) prior suggested by George and Foster (2000) and Hansen and Yu (2001) by using information contained in the data to elicit g via maximum likelihood.

B.2. Model priors

For the model prior setting we assume the binomial model prior according to which the model probability is given by:

$$p(M_j) = \gamma^{\lambda_j}(1-\gamma)^{\Lambda-\lambda_j} \tag{8}$$

where Λ is the maximum number of regressors, λ_j is the number of regressors included in the model M_j and γ is a hyperparameter that expresses the probability of each regressor. Based on this assumption, we discern between two different cases:

- Setting $\gamma = \frac{1}{2}$ assigns equal probability to all models under consideration. The expected model size is equal to $m = \frac{\Lambda}{2}$ and, therefore, favours models of medium size. In our case, $m = \frac{15}{2} = 7.5$. This leads to the uniform model prior, where each model has the same probability $p(M_i) = 2^{-\Lambda}$.
- Secondly, we use an alternative model prior that is less restrictive as far as the model size is concerned, assuming a hyperprior beta-binomial with a prior model size of $\Lambda/2^{15}$ from which the value of γ is drawn.

B.3. MCMC sampler algorithm

Using $15(=\Lambda)$ explanatory variables results in $32,768$ $(=2^{15})$ alternative models. Therefore, the posterior model distribution cannot be analytically calculated. On the contrary, a random walk chain Metropolis-Hastings sampler algorithm has to be used. More precisely, in step 1 the sampler begins with the current model M_1 that has a posterior model probability of a certain value, $p(M_1|\phi, X)$. In step 2, a new model, M_2, is proposed to replace M_1. The algorithm accepts the new model according to the following rate:

$$p_{1,2} = min\left(1, \frac{p(M_2|\phi, X)}{p(M_1|\phi, X)}\right) \tag{9}$$

If model M_2 is rejected, the next model M_3 is proposed and compared with the current M_1, using the same algorithm. If model M_2 is accepted, it becomes the current model and the process continues with a new model which may replace M_2. With a large number of iterations (here 200,000 with 100,000 used as burn-ins), the posterior model probabilities are empirically approached.

B.4. LASSO framework

Assuming the following distributions for the parameters of interest:

$$\gamma_i|h, \tau_i^2 \sim N(0, h^{-1}\tau_i^2) \tag{10}$$

$$p(\tau_i^2|\lambda) \propto exp(-\lambda\tau_i^2/2) \tag{11}$$

$$p(h) \propto 1/h \tag{12}$$

where h and τ are the crucial hyperparameters that determine the variance of the parameters of interest (γ), the inference is based on simulations on four posterior conditional distributions: $p(\gamma|\psi, h, \tau^2, \lambda^2)$, $p(h|\psi, \gamma, \tau^2, \lambda^2)$, $p(\tau^2|\psi, \gamma, h, \lambda^2)$ and $p(\lambda^2|\psi, \gamma, h, \tau^2)$. Following the standard practice, we assume an independent normal-gamma prior for the first two conditionals. For h we assume a non-informative prior as indicated by equation (12), while for τ (and for calculative purposes for $1/\tau$ we assume an inverse gamma) and for λ we assume a gamma distribution. With this prior structure, the simulations are reduced to a Gibbs sampler. In the main part of the paper, we report the posterior mean and the standard deviation for γ and the posterior mean for τ that shows which variables can remain in the model and which ones can be deleted. Coefficients that their estimated τ is zero are deleted from the model as their estimated mean is zero as equation (10) indicates. On the other hand, the meta-regression model keeps only the determinants whose estimated γ_i have a variance that is larger than zero (i.e., $\tau_i > 0$).

References

Alquist, R., Chinn, M.D., 2008. Conventional and unconventional approaches to exchange rate modelling and assessment. Int. J. Financ. Econ. 13 (1), 2–13.

Armstrong, J.S., Green, K.C., Graefe, A., 2015. Golden rule of forecasting: be conservative. J. Bus. Res. 68 (8), 1717–1731.

Bacchetta, P., Van Wincoop, E., Beutler, T., 2009. Can parameter instability explain the Meese-Rogoff puzzle? In: NBER International Seminar on Macroeconomics, pp. 125–173.

Boero, G., Marrocu, E., 2002. The performance of non-linear exchange rate models: a forecasting comparison. J. Forecast. 21 (7), 513–542.

Ca'Zorzi, M., Muck, J., Rubaszek, M., 2016. Real exchange rate forecasting and ppp: this time the random walk loses. Open Econ. Rev. 27 (3), 585–609.

Ca'Zorzi, M., Rubaszek, M., 2020. Exchange rate forecasting on a napkin. J. Int. Money Financ., 102168.

Cheung, Y.-W., Chinn, M.D., 1998. Integration, cointegration and the forecast consistency of structural exchange rate models. J. Int. Money Financ. 17 (5), 813–830.

Cheung, Y.-W., Chinn, M.D., Pascual, A.G., 2005. Empirical exchange rate models of the nineties: are any fit to survive? J. Int. Money Financ. 24 (7), 1150–1175.

Cheung, Y.-W., Chinn, M.D., Pascual, A.G., Zhang, Y., 2019. Exchange rate prediction redux: new models, new data, new currencies. J. Int. Money Financ. 95, 332–362.

Cheung, Y.-W., Lai, K.S., 2000. On the purchasing power parity puzzle. J. Int. Econ. 52 (2), 321–330.

Clements, M.P., Smith, J., 2001. Evaluating forecasts from setar models of exchange rates. J. Int. Money Financ. 20 (1), 133–148.

Eichenbaum, M.S., Johannsen, B.K., Rebelo, S.T., 2021. Monetary policy and the predictability of nominal exchange rates. Rev. Econ. Stud. 88 (1), 192–228.

Eickmeier, S., Ziegler, C., 2008. How successful are dynamic factor models at forecasting output and inflation? A meta-analytic approach. J. Forecast. 27 (3), 237–265.

Engel, C., Mark, N.C., West, K.D., Rogoff, K., Rossi, B., 2007. Exchange rate models are not as bad as you think [with comments and discussion]. NBER Macroecon. Annu. 22, 381–473.

Engel, C., Wu, S.P.Y., 2023. Forecasting the us dollar in the 21st century. J. Int. Econ. 141, 103715.

Fang, S., Wei, Y., Wang, S., 2023. 30 years of exchange rate analysis and forecasting: a bibliometric review. J. Econ. Surv.

Ferrara, L., Marcellino, M., Mogliani, M., 2015. Macroeconomic forecasting during the great recession: the return of non-linearity? Int. J. Forecast. 31 (3), 664–679.

Ferraro, D., Rogoff, K., Rossi, B., 2015. Can oil prices forecast exchange rates? An empirical analysis of the relationship between commodity prices and exchange rates. J. Int. Money Financ. 54, 116–141.

Filippidis, M., Filis, G., Magkonis, G., 2024. Evaluating oil price forecasts: a meta-analysis. Energy J. 45 (2).

Frankel, J.A., Rose, A.K., 1995. Empirical research on nominal exchange rates. In: Handbook of International Economics, vol. 3, pp. 1689–1729.

George, E., Foster, D.P., 2000. Calibration and empirical Bayes variable selection. Biometrika 87 (4), 731–747.

Giacomini, R., Rossi, B., 2010. Forecast comparisons in unstable environments. J. Appl. Econom. 25 (4), 595–620.

Hansen, M.H., Yu, B., 2001. Model selection and the principle of minimum description length. J. Am. Stat. Assoc. 96 (454), 746–774.

Havránek, T., Stanley, T., Doucouliagos, H., Bom, P., Geyer-Klingeberg, J., Iwasaki, I., Reed, W.R., Rost, K., Van Aert, R., 2020. Reporting guidelines for meta-analysis in economics. J. Econ. Surv. 34 (3), 469–475.

Imbs, J., Mumtaz, H., Ravn, M.O., Rey, H., 2005. Ppp strikes back: aggregation and the real exchange rate. Q. J. Econ. 120 (1), 1–43.

Kass, R.E., Raftery, A.E., 1995. Bayes factors. J. Am. Stat. Assoc. 90 (430), 773–795.

Kavtaradze, L., Mokhtari, M., 2018. Factor models and time-varying parameter framework for forecasting exchange rates and inflation: a survey. J. Econ. Surv. 32 (2), 302–334.

Liang, F., Paulo, R., Molina, G., Clyde, M.A., Berger, J.O., 2008. Mixtures of g priors for Bayesian variable selection. J. Am. Stat. Assoc. 103 (481), 410–423.
Mark, N.C., 1995. Exchange rates and fundamentals: evidence on long-horizon predictability. Am. Econ. Rev., 201–218.
Meese, R., Rogoff, K., 1983a. The out-of-sample failure of empirical exchange rate models: sampling error or misspecification? In: Exchange Rates and International Macroeconomics. University of Chicago Press, pp. 67–112.
Meese, R., Rogoff, K., 1988. Was it real? The exchange rate-interest differential relation over the modern floating-rate period. J. Finance 43 (4), 933–948.
Meese, R.A., Rogoff, K., 1983b. Empirical exchange rate models of the seventies: do they fit out of sample? J. Int. Econ. 14 (1–2), 3–24.
Melvin, M., Prins, J., Shand, D., 2013. Forecasting Exchange Rates: An Investor Perspective. Handbook of Economic Forecasting, vol. 2. Elsevier, pp. 721–750.
Murray, C.J., Papell, D.H., 2002. The purchasing power parity persistence paradigm. J. Int. Econ. 56 (1), 1–19.
Rapach, D.E., Wohar, M.E., 2002. Testing the monetary model of exchange rate determination: new evidence from a century of data. J. Int. Econ. 58 (2), 359–385.
Rogoff, K., 1996. The purchasing power parity puzzle. J. Econ. Lit. 34 (2), 647–668.
Rossi, B., 2006. Are exchange rates really random walks? Some evidence robust to parameter instability. Macroecon. Dyn. 10 (1), 20.
Rossi, B., 2013. Exchange rate predictability. J. Econ. Lit. 51 (4), 1063–1119.
Sarantis, N., 1999. Modeling non-linearities in real effective exchange rates. J. Int. Money Financ. 18 (1), 27–45.
Sarno, L., Valente, G., 2009. Exchange rates and fundamentals: footloose or evolving relationship? J. Eur. Econ. Assoc. 7 (4), 786–830.
Taylor, A.M., Taylor, M.P., 2004. The purchasing power parity debate. J. Econ. Perspect. 18 (4), 135–158.
Teräsvirta, T., 2006. Forecasting economic variables with nonlinear models. In: Handbook of Economic Forecasting, vol. 1, pp. 413–457.
Wieland, V., Wolters, M., 2013. Forecasting and Policy Making. Handbook of Economic Forecasting, vol. 2. Elsevier, pp. 239–325.
Zellner, A., 1986. On assessing prior distributions and Bayesian regression analysis with g-prior distributions. In: Bayesian Inference and Decision Techniques.