92

CUADERNOS DE ECONOMÍA

ISSN 0121-4772

Facultad de Ciencias Económicas
Escuela de Economía
Sede Bogotá

UNIVERSIDAD NACIONAL DE COLOMBIA

Indexación, resúmenes o referencias en

SCOPUS
Thomson Reuters Web of Science
(antiguo ISI)-SciELO Citation Index
ESCI (Emerging Sources Citation Index) - Clarivate Analytics

EBSCO
Publindex - Categoría B - Colciencias
SciELO Social Sciences - Brasil
RePEc - Research Papers in Economics
SSRN - Social Sciences Research Network
EconLit - Journal of Economic Literature
IBSS - International Bibliography of the Social Sciences
PAIS International - CSA Public Affairs Information Service
CLASE - Citas Latinoamericanas en Ciencias Sociales y Humanidades
Latindex - Sistema regional de información en línea
HLAS - Handbook of Latin American Studies
DOAJ - Directory of Open Access Journals
CAPES - Portal Brasilero de Información Científica
CIBERA - Biblioteca Virtual Iberoamericana España / Portugal
DIALNET - Hemeroteca Virtual
Ulrich's Directory
DOTEC - Documentos Técnicos en Economía - Colombia
LatAm-Studies - Estudios Latinoamericanos
Redalyc

Universidad Nacional de Colombia
Carrera 30 No. 45-03, Edificio 310, primer piso
Correo electrónico: revcuaeco_bog@unal.edu.co
Página web: www.ceconomia.unal.edu.co
Teléfono: (571)3165000 ext. 12308, AA. 055051, Bogotá D. C., Colombia

Cuadernos de Economía Vol. 43 No. 92 - 2024

ARTÍCULO

COMBINATION OF THEORETICAL MODELS FOR EXCHANGE RATE FORECASTING

María Paula Bonel

Bonel, M. P. (2024). Combination of theoretical models for exchange rate forecasting. *Cuadernos de Economía*, *43*(92), 437-467.

This paper states that there are exchange rate forecasting gains when combining in-sample data from different models based on economic theory. Data combination is performed using Bayesian model averaging (BMA). Using pooled data by group of countries (developed and emerging economies) generates accuracy gains in an important amount of cases, with respect to forecasts that use country information. Gains are larger for currencies of developed economies, but accuracy decreases as the forecast horizon is extended. BMA models for developed countries tend to be more "sparse" than emerging countries models.

Keywords: Bayesian model averaging; exchange rate; forecasting; model uncertainty.
JEL: C11, C30, C53, F31.

M. P. Bonel
Facultad de Ciencias Económicas, Universidad Católica Argentina, Ciudad de Buenos Aires (Argentina). Universidad de Buenos Aires, Facultad de Ciencias Económicas. CONICET-Universidad de Buenos Aires, Instituto Interdisciplinario de Economía Política de Buenos Aires (IIEP-BAIRES), Ciudad de Buenos Aires (Argentina). Email: pbonel@cbc.uba.ar.

Suggested citation: Bonel, M. P. (2024). Combination of theoretical models for exchange rate forecasting. *Cuadernos de Economía*, *43*(92), 437-467. https://doi.org/10.15446/cuad.econ.v43n92.98393

This paper was received on September 13, 2021, revised on September 27, 2023, and finally accepted on September 28, 2023.

Bonel, M. P. (2024). Combinación de modelos teóricos para el pronóstico de tipo de cambio. *Cuadernos de Economía*, *43*(92), 437-437.

Este artículo propone la existencia ganancias en la predicción de tipo de cambio cuando se combinan datos *in-sample* de diferentes modelos basados en la teoría económica. La combinación se realiza mediante Bayesian Model Averaging. Entrenar el modelo con información de otras economías genera ganancias de precisión en una cantidad importante de casos, respecto a pronósticos que utilizan solo información del país. Mayores ganancias de precisión se encuentran para divisas de economías desarrolladas. Los modelos entrenados para países desarrollados tienden a ser más “escasos” que los modelos de países emergentes.

Palabras clave: promedio modelo bayesiano; tipo de cambio; pronósticos macroeconómicos; incertidumbre del modelo.
JEL: C11, C30, C53, F31.

INTRODUCTION

The forecast of macroeconomic variables plays a central role both in academic studies and in the analysis of monetary and financial policy. Li and Chen (2014) emphasise that having accurate forecasts helps to better understand the dynamics of the economy. Any economic decision, such as the management of portfolios and hedging strategies or monetary policies, necessarily implies establishing beliefs about the evolution of macroeconomic variables.

Exchange rate forecasting presents additional difficulties. It is characterised by presenting a *puzzle* related to a disconnection from its fundamentals. Meese and Rogoff (1983) found that the estimates from a random walk were more accurate than the forecasts from other models based on different economic theories. This resulted in different theoretical models developed to explain exchange rate changes over time (Fisher, 1896; Cassel, 1918; Dornbusch, 1976). Naturally, this problem is not foreign to the different conditions of the economy under study. Those models are a starting point for the choice of predictors when forecasting. Recent empirical work on exchange rate forecasting has found evidence of models that obtain more accurate forecasts than a random walk, although there is consensus that the predictive performance is sensitive to the choice of predictor, forecast horizon, sample period, model and also the chosen evaluation method (Rossi, 2013).

As an alternative for dealing with the lack of knowledge around the true model, the researcher can consider possible gains resulting from forecast combinations. Timmermann (2006) states that in several empirical studies, combinations of forecasts have been found to achieve more accurate results on average than other methods based on the selection of the best individual model ex ante. This happens even when working with simple combinations that ignore correlations between forecast errors. One of the possible causes of the forecast combination gains identified by the paper is the portfolio diversification argument. Due to the difficulties in detecting structural changes in real time, it is plausible that combinations of forecasts based on models with different degrees of adaptability outperform individual model forecasts. Finally, individual forecast models may be subject to unknown specification bias, causing forecast combinations to obtain results that are more robust to such misspecification.

When selecting and combining forecasts, the forecaster's uncertainty about the true model and the known instability of macroeconomic fundamentals, initially documented by Meese and Rogoff (1983), are explicitly addressed. In this paper we will seek to study the predictive ability of different models using the Bayesian model averaging (BMA) methodology, a particular case within forecast combination literature. This methodology is flexible enough to address both the problems of variable selection and the combination of variables.

The main objective of this paper is to study whether there are forecasting gains, with respect to a benchmark model, when combining fitted data from different empirical models based on economic theory. From this, another question arises:

which Bayesian models perform better for this task? The main results obtained are summarised below. BMA models provide more accurate forecasts than random walk, the benchmark model in the literature. This difference is statistically significant for an important number of country and horizon combinations (1, 2 and 4 quarters ahead). Training the model with information from other countries using pooled data generates accuracy gains in about 60% of the cases, with respect to forecasts generated using country information at the individual country level. This percentage is maintained throughout the different forecast horizons. Gains in accuracy decrease as the forecast horizon is extended. A regularity found for both country groups is that the BMA1 model tends to perform better. This model is the least restrictive in terms of both coefficient constraint and model size, indicating that the in-sample forecasts we generate under different theoretical models have relevant information for exchange rate estimation. By country group, the gain is larger for developed countries. The model trained for developed countries is "sparser" than for emerging ones. The trained model, in both country panels, becomes denser as the forecast horizon is extended. The analysis of average a posteriori inclusion probabilities helps us to understand which theoretical models tends to be selected for forecasting with the BMA methodology. Here, important differences emerge between country groups. In the developed countries model, the price growth differential model containing information for the previous 4 quarters stands out strongly, while for emerging countries the interest rate differential model is also incorporated with high probability in the different estimation horizons. Forecasting exercises on the direction of change (presented in an online annex) also seem to yield positive results in forecast performance.

THEORETICAL MODEL SELECTION

This section reviews a selection of well-known theoretical models for exchange rate forecasting. A brief literature review of each is given below. We also detail the empirical specifications used to estimate each of the models under the ordinary least squares (OLS) method recursively. This part of the work allows us to obtain in-sample adjusted data for each of the theoretical models, which are then combined using the for BMA methodology. Although these models are useful for anticipating exchange rate movements in both large and small economies, it is also true that the relevance of each of the models can vary with the specificity of each economy. In this sense, the BMA methodology used later will allow us to adapt the parameters of the model in order to obtain the best out-of-sample results.

Inflation differentials model

Relative purchasing power parity requires that the growth rates in the exchange rate offset the differential between the growth rates of domestic and international prices. Empirical evidence from out-of-sample estimation has mixed results.

In their recent paper, Cheung et al. (2019) find that, when the forecast horizon is one year or longer, there are some improvements in forecasting with price fundamentals.

The specification of the model is determined as follows:

$$s_{t+h} - s_t = \alpha_t + \beta_{1t\left(\pi_{t,t-k} - \pi^{*}_{t,t-k}\right)} + \beta_{2t\,(s_t - s_{t-k})} + u_{t+h,t} \tag{1}$$

where s_t is the logarithm of bilateral nominal exchange rate at time *t, h* is the forecast horizon, *h=1,2,4* quarters. We use four specifications for this model that includexchange rate changes and inflation differentials. $\pi_{t,t}-k$ and $\pi_{t,t}{}^{*}-k$ are the domestic and foreign inflation rates in the last *k quarters*, respectively. In each specification of the model, *k* takes the value 1, 2, 4 or 12 quarters. We will refer to each of these models as PPP3, PPP6, PPP12 and PPP36, respectively.

Interest rate parity model

In 1896, Fisher conducted an analysis of how interest rates can be related to expected changes in international currencies. This relationship is known as uncovered interest rate parity (UIRP).

Cheung et al. (2005) and Alquist and Chinn (2008) argue that, although for some countries rate parity obtains better predictions than the random walk model at long horizons, its performance is never significantly better. Chinn and Meredith (2004) and Molodtsova and Papell (2009) have reported slightly more positive results at short horizons for a group of developed countries. Cheung et al. (2019) find that interest rate parity rarely works well, but if it does, it does so over a longer horizon, such as one to 5 years.

Following Molodtsova and Papell (2009), the interest rate differential model is estimated using:

$$s_{t+h} - s_t = \alpha + \beta_{t\left(i_t - i^{*}_t\right)} + u_{t+t,t} \tag{2}$$

where i_t and i_{*_t} are the domestic and foreign nominal interest rates. Since we do not constrain $\beta = 1$, or even its sign, this equation can be consistent with UIRP, where a positive interest rate differential produces forecasts of exchange rate depreciation, and the risk premium *puzzle* literature. Hereafter, we will refer to this model as UIP.

Monetary model

The monetary model, also called the asset model, holds that the exchange rate varies in order to balance the international demand for assets stock rather than the flow of demand for goods, as in more traditional views. According to the monetary model of exchange rate determination, bilateral exchange rate fluctuations should reflect movements in relation to money, output, interest rates and prices between the two countries. The monetary model was introduced during the 1970s in various works such as Frenkel (1976), Dornbusch (1976), Frankel (1979) and, later, Mussa (1982). It is based on a simple small open economy model where real output is exogenous. The demand for real money is viewed as a function of income and the interest rate. There are two approaches within this model depending on the assumption made about price behaviour. We will focus on models that hold that prices are sticky, at least in the short term.

Mark (1995) and Mark and Sul (2001) find strong and statistically significant evidence in favour of the monetary model at long horizons (three to four years). Some positive out-of-sample prediction results using monetary fundamentals were found by Medel et al. (2015) for the UK and Euro area. Research has also found that positive results depend on the sample period used (Rossi, 2013).

Below we describe the monetary model specification (SMON):

$$s_t = \alpha + \beta_1(m_t - m^*_t) + \beta_2(y_t - y_t^*) + \beta_3(i_t - i^*_t) + \beta_4(\pi_t - \pi_t^*) + u_t \qquad (3)$$

where m is the logarithm of money supply, y the real GDP, i and π represent the interest and inflation rates respectively, and u is the error term. As in previous cases, * indicates that the variable belongs to the foreign country. In this case, no restriction is imposed on the coefficients since the theory does not provide clear guidance as to the value of these coefficients. After incorporating the one-period lags of fundamentals and differencing the equation above, we consider the following forecasting equation:

$$\begin{gathered} s_t + h - s_t = \alpha + \beta_1(\Delta m_t - \Delta m_{t*}) + \beta_2(\Delta y_t - \Delta y_{t*}) + \beta_3(\Delta i_t - \Delta i_{*t}) + \\ \beta_4(\Delta \pi_t - \Delta \pi_{t*}) + u_t + h, t \end{gathered} \qquad (4)$$

where $st + h - st$ is the forecast for the change in exchange rate in the next $h = 1,2,4$ quarters and $\Delta x_t = x_t - x_{t-1}$ indicates the change in the previous quarter for a variable x.

Taylor rule model

Taylor (1993) formalises the idea that the monetary authority sets the real interest rate according to how inflation differs from its target level (the higher the inflation, the more contractionary the monetary policy will be) and also according to the output gap. That is, if output is below the potential output, monetary policy will be more expansionary.

Molodtsova and Papell (2009) adapt this concept by taking the decisions in two countries into account in order to analyse their relationship with the exchange rate through the output gap and inflation differentials. Cheung et al. (2019) also uses a simple specification of this model. Results in Molodtsova and Papell (2009) show that the Taylor rule model predicts the out-of-sample exchange rate significantly better than the random model for several countries, although performance depends on the exact specification. Some of the specifications that performed better include heterogeneous coefficients across countries and interest rate smoothing. Cheung et al. (2019) find positive but not significant results except for the period 1983q1 2014q4.

Cheung et al. (2019) use the following regression to estimate the exchange rate following the Taylor rule (TR) model:

$$s_{t+h} - s_t = \beta_0 + \beta_1 \left(y_t - \widetilde{y_t} \right) + \beta_2 \left(\pi_t - \pi_t^* \right) + u_{t+h,t} \tag{5}$$

where $\widetilde{y_t}$ is the output gap. This equation uses inflation differentials as we assume that the international inflation rate works as a nominal anchor.

External imbalances measures

There are several papers that highlight the importance of external imbalances vis-à-vis changes in the exchange rate. Gourinchas and Rey (2007) argue that not only the current account but the whole dynamic process of net exports, foreign asset holdings and net foreign asset portfolio returns are important predictors of exchange rates. When a country experiences a current account imbalance, the traditional intertemporal approach to the current account suggests that the country will need to run surpluses of funds to reduce this imbalance. Gourinchas and Rey (2007) argue, instead, that part of the adjustment can be made through a transfer of wealth between that country and the rest of the world which occurs through a depreciation in the value of its currency. They find empirical evidence in favour of external imbalance measures as predictors of exchange rate.

Della Corte et al. (2012) find that the net foreign asset model can predict significantly better out-of-sample (effective) exchange rates than the random walk at both long and short horizons. Alquist and Chinn (2008) find that in some subsamples the (bilateral) exchange rates predicted from the net foreign assets model are better than the random walk at short horizons for some countries. However, results are less favourable at longer horizons.

We decided to use the following specification, hereafter referred to as CA, as a simple way to address the relationship between external imbalances and the exchange rate:

$$st+h - st = \beta 0 + \beta 1cat + ut+h,t \tag{6}$$

where ca_t represents the current account as a percentage of gross domestic product.

Dense forecasting methods

The section described above makes it possible to count on different exchange rate forecasts supported by theoretical models. In several cases, accuracy gains were found related to random walk results, but we can also state that no unique model arises as the true one. There is empirical evidence that affirms that combinations of forecasts produce better estimates on average than methods based on the best (ex ante) forecast model at the individual level (Timmermann, 2006).

In the context of exchange rate prediction, several recent empirical studies claim to have found a relationship between exchange rates and macroeconomic fundamentals, although these relationships are often unstable or short-lived (Rossi, 2013). The empirical literature has also shown that combinations of exchange rate forecasts perform better than models based on individual fundamentals (see, for example, Della Corte et al. (2009)). In this context, it may be desirable to work with techniques that can deal with uncertainty on forecasting models and predictors' selection.

There are different methodologies that aim to deal with the researcher's lack of knowledge around the best predictors of the exchange rate. This paper will focus on "dense" prediction techniques. As Giannone et al. (2017) describe, it is recognised in these cases that all possible explanatory variables may be important for the prediction, although the impact of some of them may be small. Factor models, Bayesian averaging models or Ridge regressions are examples of dense models.

Bayesian model averaging methodology (BMA) proposes a methodological framework for generating combinations of forecasts by taking advantage of the information gains contained in the different variables and therefore dealing with uncertainty. BMA is essentially an application of Bayesian inference to model selection problems. The methodological section describes its methodology in detail.

Wright (2008) has used this methodology to study a set of developed country currencies and finds that, for most currency-horizon pairs, BMA forecasts using a sufficiently high degree of shrinkage obtain a slightly smaller out-of-sample mean square prediction error than the random walk benchmark. This paper, identified as the main reference for the exercise, differs from our current work in two major ways. Firstly, Wright (2008) studies the combination of different predictors and n ot in-sample forecasts obtained through empirical models. The second difference is that it only works with currencies of developed countries.

Bayesian model averaging

The lack of knowledge about the best model, coupled with parameter uncertainty, is an important problem in econometrics. The Bayesian model averaging technique emerges as an alternative for dealing with this type of uncertainty. The methodology we use is adapted from Wright (2008). In this paper, we evaluate the combination of eight predictors that are the in-sample forecasts based on theoretical models: PPP3, PP6, PPP12, PPP36, UIP, SMON, TR, and CA.

A set of n models is considered, $M_1 \dots M_n$. The i-th model is indexed by a vector of parameters θ_i. The researcher knows that one of these models is true, but she does not know which one. Also, the researcher has prior beliefs about the probability that the i-th model is correct. That probability is written as $P(M_i)$. Subsequently, the researcher observes the data, D, and updates her beliefs to compute the a posteriori probability that the i-th model is correct:

$$P\left(M_i|D\right)=\frac{P\left(D|M_i\right)P\left(M_i\right)}{\sum_{j=1}^{n}P\left(D|M_j\right)P\left(M_j\right)} \tag{7}$$

where

$$P(D\mid Mi)=\int P(D\mid \theta i, Mi)P(\theta i\mid Mi)d\theta i \tag{8}$$

is the marginal likelihood of the i-th model. $P(\theta_i|M_i)$ is the a priori density of the parameter vector in this model and $P(D|\theta_i,M_i)$ is the likelihood. Each model results in a forecast. In the presence of uncertainty, the final forecast weighs each of those forecasts by the posterior for that model. This gives the minimum mean square forecast error. The researcher only needs to establish the set of models, the model priors $P(M_i)$ and the parameter priors $P(\theta_i|M_i)$.

The *i-th* model is specified as:

$$y = \beta_i X_i + \epsilon \tag{9}$$

where y is a vector of observations T x 1 on the variable we wish to predict and X_i is a matrix of predictors $T \, x \, p_i$, β is a vector of parameters $p_i \, x \, 1$. ϵ is the vector of disturbances. The disturbances are i.i.d, meaning that they have zero mean and variance equal to σ^2, and $\theta_i = (\beta'_i, \sigma^2)$.

In this case, regressors are assumed to be strictly exogenous to obtain the closed form of the model. We know that this assumption is false for the problem we are dealing with. However, this assumption does not prevent these methods from performing well when forecasting (Wright, 2008; Stock & Watson, 2005).

The need to obtain a posteriori distributions requires that prior beliefs about the model parameters are specified. For the parameter priors we take the specification of the natural conjugate g-prior for β_i, such that the prior for β_i is conditional on σ^2 is $N(X'_i X'^{-1}_i)^1$. For σ^2 we assume the "improper" prior which is proportional to $1/\sigma^2$.

According to Zellner (1971), one can calculate the likelihood of the model as:

$$p\left(D|M_i\right)=\frac{1}{2}\frac{``\left(\frac{T}{2}\right)}{\pi^{\frac{T}{2}}}\left(1+\phi\right)^{\frac{p_i}{2}} S_i^{-T} \tag{10}$$

Where $S_i^2 = Y'Y - Y'X_i\left(X_i'X_i\right)^{-1}X_iY\frac{\phi}{1+\phi}$. Note that Γ is the gamma function and π is the number pi.

The a posteriori mean of β_i is:

$$\widetilde{\beta_i}=E(\beta_i\ |D,M_i)=\frac{\phi}{1+\phi}\left(X_i'X_i\right)^{-1}X_i'Y \tag{11}$$

The prior of β is centred around zero leading each model to the assumption of no predictive ability. The level of shrinkage is bound by ϕ. The hyperparameter ϕ represents the certainty that the coefficients are zero. A small ϕ implies smaller priors on the coefficient's variance and, therefore, implies that the researcher is fairly certain that the coefficients are actually zero. In contrast, a larger size means that the researcher is not sure that the coefficients are zero. That is, we are more willing to move from our prior beliefs in response to what we observe in the data. A popular default approach is to determine this parameter according to the unit information prior (UIPr) criterion, which states $\phi = n$. This can be considered as an a priori distribution containing the same amount of information as a single observation.

Fernandez et al. (2001) argue that a comparatively large likelihood minimises the impact of prior beliefs on the results, keeps the results close to the OLS coefficients, and represents the absolute lack of prior knowledge. On the other hand, Ciccone and Jarociński (2010) show that a large ϕ may not be robust against noise and generates risk of overfitting, particularly if the noise component plays an important role in the data. Wright (2008) indicates that better results are usually obtained in exchange rate prediction exercises when the priors are informative. In order to take different prior beliefs into account in the model construction, estimations will be carried out under two different parameter cases. The BMA1 and BMA2 models use the default approach *UIPr* in which the same information is attributed to the priors as is contained in an observation. The BMA3 and BMA4 cases use a hyperparameter $\phi = 0.5$ which favours prior beliefs that the coefficients are zero. The timely choice of that parameter value is based on how well it performed in Wright's (2008) work.

The i-th model that forecasts the exchange rate has the following form:

$$St + h - St = \beta_i' Xi.t + \epsilon_t \tag{12}$$

where $X_{i,t}$ is the vector of regressors in period t for model i and ϵ_t is the error term. Each model obtains a forecast $\widetilde{\beta_i}'X$, where $\widetilde{\beta_i}'$ indicates the a posteriori mean of

β'. Then, the forecast weights each of these models by their a posteriori probabilities, obtaining a forecast that is equal to:

$$\sum_{i=j}^{n} P(M_i \mid D)\tilde{\beta}_i{}' X_{i,t} \tag{13}$$

BMA models consist of all possible permutations of λ potential predictors (in our case, in-sample forecasts), from the model that includes all predictors and none of them, thus obtaining 2^{λ} candidate models. Each of these models includes a constant, except for the model that has no regressors, implying that this model is the random walk. Following Wright (2008), the a priori probability of each model with k predictors (excluding the intercept) is determined as:

$$P(M_i) = \rho^{k}(1 - \rho)^{\lambda-k} \tag{14}$$

If $\rho = 0.5$ then all models have the same weight. Assigning equal a priori probability to all models means that models with fewer predictors may receive little a priori weighting. A smaller value of ρ favours models that are sparser or have fewer predictors. The probability that the correct model is the one that does not include any predictors is equal to $(1 - \rho)^{\lambda}$.

Forecasts will be estimated under two particular cases of the parameter ρ in order to address the lack of knowledge about the size of the model. It follows that a first approach to the problem would be to establish beliefs about the model size that are uniform across all possible models, i.e., $\rho = 0.5$. This will be reflected in the BMA1 and BMA3 specifications. Thus, since there are eight predictors (or in-sample forecasts) with a probability of inclusion of 0.5 each, the expected size of the a priori model is determined to be 4 (M prior). This implies that, given that there are more possible models of size 3 than, for example, of size 1 or 8, uniform beliefs attribute higher probability to models of intermediate size. This is not necessarily an assumption that applies a priori to the problem under study. We therefore determine an alternative specification used in BMA2 and BMA4 where $\rho = 0.5$ for models with 0 and 1 variables, while for the rest of the models with more variables the probability is close to zero. This results in the expected size of the a priori model (M prior) being around 0.9, favouring sparser models and variable and selection.

The methodology described above results in four BMA model specifications under different combinations of parameters ϕ and ρ. The choice of these four specifications is aimed at considering different assumptions that a researcher may have.

Results evaluation

The performance of the different models will be evaluated relative to a benchmark model. From the literature listed above, it arises that the best predictor of the exchange rate is the random walk model without constant term or intercept (Meese & Rogoff, 1983):

$$E_t(s_{t+h} - s_t) = 0 \tag{15}$$

In other words, the benchmark model implies that the best forecast of tomorrow's exchange rate is today's exchange rate.

Subsequently, to evaluate the performance of the out-of-sample models, the total sample is divided into two parts: the in-sample part, which consists of the window of observations used to train the model, and the out-of-sample part on which the performance of the models is actually tested. In this case, a recursive forecasting scheme is used, where the model parameters are initially estimated with a window of 60 observations and, subsequently, the model parameters are re-estimated using all the previous observations, adding the observation of the new period. This is a standard way of simulating the available data at the moment of making a forecast. At this stage, we generate the in-sample adjusted data from all the theoretical models. We then combine that information using Bayesian model averaging. Estimates are made where h corresponds to one, two or four steps forward, corresponding to estimation horizons of 1, 2 and 4 quarters in advance.

The forecast evaluation process requires two main decisions: the choice of loss function to evaluate the forecast and the choice of statistical test to evaluate the significance of performance differences. The root mean square prediction error (RMSPE) was chosen as the loss function, as in Meese and Rogoff (1983). This indicator gives equal weight to forecasts that underestimate or overestimate the exchange rate. To simplify the visualisation of the data, the ratio between the RMSPE of the model to be tested and the RMSPE of the random walk model will be shown in the results section. A number less than unity indicates that the chosen BMA model obtains better results than the random walk.

To test the significance of the differences on the errors in the models estimated under time series, we used the Diebold and Mariano (1995) test designed specifically for forecast comparison, taking the forecast horizon to be analysed into account in all cases ,and correcting for error correlation. For the analysis of panel data, we used Pesaran et al.'s (2009) extension of this test. This allowed us to statistically test the differences between forecasts while considering the panel structure of the data. One problem that arises is that the forecasts overlap with h-staggered forecasts (in the 2- and 4-quarter forecast horizon cases). Therefore, it is possible that forecast errors separated by less than h periods are correlated. This problem is avoided by using only each h-th observation to compute statistical inference. A positive sign of the t-statistic indicates that the Bayesian forecasts perform better (in a differential sense of squared error loss) than the forecasts corresponding to the reference model.

DATA

We work with quarterly data from Q11986-Q42018, with some differences in the sample extension depending on data availability. Panel A of developed countries, includes Australia(AUD), Canada(CAD), Japan(JPY), South Korea(KRW), New Zealand(NZD), Singapore (SGD), Sweden (SEK), Switzerland (CHF), and Great Britain (GBP). In the group of emerging countries (panel B) we worked with Argentina (ARG), Brazil (BRA), Colombia (COP), Philippines (PHP), Indonesia (IDR), Malaysia (MYR), and Mexico (MXN).

The main source of data used to construct the macroeconomic fundamentals is the IMF's international financial statistics database (IFS). In cases where the necessary information was not available, we used information provided by central banks and official institutions. The information on exchange rates is obtained from IFS and corresponds to bilateral exchange rate data with the US dollar at the end of the period.

The short-term interest rate corresponds to 3-month "money market" interest rates. M1 is used to measure the money supply. The price level in the economy is measured by the consumer price index and used to calculate inflation rates. Current account balance as a percentage of GDP was obtained from the databases provided by the OECD and the IMF.

Output indices are used as a proxy for the level of activity due to their availability at the quarterly level in all cases. The output gap depends on the measure of potential output. Since there is no assumption about the definition of potential output used by central banks in their interest rate reaction functions, we consider the percentage deviations of output from a trend defined by Hodrick and Prescott (1997). To mimic the information available to central banks at time t, when decisions were made, as closely as possible, only data up to time t-1 are used to construct the trend. Therefore, in each period the regression is re-estimated by adding an additional observation to the sample.

RESULTS

At the beginning of this section, we present the results of the RMSPE ratio when the models are estimated using pooled data for all countries within panel A (developed countries) or panel B (emerging countries), respectively. In all cases estimates are made for 1, 2 and 4 quarters ahead. Estimates are direct (non-iterated) forecasts for all horizons. In addition, the statistic and p-value corresponding to the significance test of the differences are detailed. In this case, observations corresponding to times when these countries were under a fixed exchange rate regime are not considered during the estimation (see table 8 in Annex). The objective of training the models using information from the whole group of developed countries in one case, and the whole group of emerging countries in the other, is to test whether incorporating information related to other countries improves the out-of-sample forecast results in each of the panels.

Table 1 includes the average performance of the models across all countries included within each panel. This is expressed through the simple average of the RMSPE ratios. In addition to this, the results of the significance tests of differences at the global level within each panel are described using Pesaran et al.'s (2009) t-statistic to take the panel structure of the data into account. The results look strongly significant for the 1 and 2 quarter horizons in both cases, while at 4 quarters ahead the forecast gains are not significant. In line with works like Wright (2008), we find that the results at the forecast accuracy level deteriorate with longer estimation horizons. The positive results for the group of developed countries are stable across the different Bayesian model specifications, while in the case of emerging countries we clearly identify a deterioration in BMA3 and BMA4 cases.

Tables 2 and 3 present these same results but disaggregated for each of the countries in each panel. Regarding the RMSPE ratios, we find that for developed countries the BMA models obtain lower RMSPEs than the benchmark model. Although the differences are of small magnitude, the results are found to be stable across countries, model specifications and estimation horizons. This is summarised in the simple average results, which are less than 1 in all panel A cases. In general terms, the BMA1 model seems to obtain the most accurate results. This model is the least restrictive in terms of both coefficient constraint and model size, indicating that the in-sample forecasts we generate under different theoretical models have relevant information for exchange rate estimation. The improvements in forecasting turn out to be significant for several cases under the BMA1 and BMA2 specifications. Within Table 3 for panel B (emerging countries), results are less stable across the different BMA model specifications. However, the magnitudes of differences in the ratio of the Bayesian model to the random model are larger. Again, the results seem to indicate that on average the BMA1 specification obtains the most accurate results. On the contrary, the BMA4 specification that includes more restrictive priors is the one that obtains the worst forecasting results. In this case, the results are only significant in some countries.

It is relevant to note that we see an improvement in both country panels when results are estimated at the pool level with respect to individual level data (described in tables 4 and 5). This indicates that the information in the exchange rate evolution of other countries contains relevant information for the estimation of the exchange rates under study. This is true for about 60% of the cases, something that is stable over the different horizons but more important for developed countries.

Table 1.
RMSPE ratio and significance test–Pooled data results (avg.)

Panel A:	Developed countries (Period: 1986-2018)				
Horizon		**BMA-1**	**BMA-2**	**BMA-3**	**BMA-4**
1q	Ratio RMSPE (avg.)	0.97	0.973	0.988	0.993
	t-stat	-7.81	-8.09	-4.78	-3.46
	p-value	**0.00**	**0.00**	**0.00**	**0.00**
2q	Ratio RMSPE (avg.)	0.979	0.975	0.988	0.989
	t-stat	-2.82	-4.42	-2.11	-2.24
	p-value	**0.00**	**0.00**	**0.02**	**0.01**
4q	Ratio RMSPE (avg.)	0.99	0.987	0.98	0.985
	t-stat	-1.22	-1.00	-1.79	-1.06
	p-value	0.11	0.16	**0.04**	0.14
Panel B:	Emerging countries (Period: 1990-2018)				
Horizon		**BMA-1**	**BMA-2**	**BMA-3**	**BMA-4**
1q	Ratio RMSPE (avg.)	0.756	0.756	0.996	1.033
	t-stat	-4.92	-4.92	-1.37	0.14
	p-value	**0.00**	**0.00**	**0.09**	0.55
2q	Ratio RMSPE (avg.)	0.882	0.882	1.132	1.123
	t-stat	-1.83	-1.84	0.93	1.10
	p-value	**0.03**	**0.03**	0.82	0.86
4q	Ratio RMSPE (avg.)	0.957	0.957	1.199	1.216
	t-stat	0.79	0.79	1.15	1.35
	p-value	0.79	0.78	0.87	0.91

RMSPE ratio is calculated as the simple average across all countries in each panel, respectively. We are considering forecasts from models trained using pool data by group of countries. The t-statistic and p-value data were computed using the Pesaran et al. (2009) test that extends the Diebold Mariano test for forecasting panels. More details are described in section Results evaluation. *ModelBMA1:* ϕ*: UIPr*, $\rho = 0.5$ uniform at all model sizes. *BMA2:* ϕ*: UIPr*, $\rho = 0.5$ for model sizes *0* and *1*. *BMA3:* ϕ*: 0.5*, $\rho = 0.5$ uniform to all model sizes. *BMA4:* ϕ*: 0.5*, $\rho = 0.5$ for model sizes *0* and *1*.

Table 2.
Panel A: RMSPE ratio and significance test–Pooled data results
Panel A: Developed countries

Ratio RMSPE–Period 1986-2018													
Country	**Horizon**	**BMA-1**	**t-stat**	**p-value**	**BMA-2**	**t-stat**	**p-value**	**BMA-3**	**t-stat**	**p-value**	**BMA-4**	**t-stat**	**p-value**
AUD	1q	0.971	3.549	**0.000**	0.973	3.052	**0.002**	0.991	2.142	**0.018**	0.994	1.277	0.103
	2q	0.979	1.491	**0.070**	0.973	2.039	**0.023**	0.992	0.596	0.276	0.988	1.123	0.133
	4q	1.002	-0.253	0.599	0.987	1.037	0.152	0.991	0.750	0.228	0.990	1.078	0.142
CAD	1q	0.967	2.967	**0.002**	0.972	2.641	**0.005**	0.989	1.852	**0.034**	0.993	1.061	0.146
	2q	0.962	2.233	**0.014**	0.975	2.128	**0.018**	0.989	1.470	**0.073**	0.993	0.925	0.179
	4q	0.976	1.044	0.150	0.984	1.110	0.135	0.985	1.334	**0.093**	0.991	0.897	0.186
JPY	1q	0.978	1.558	**0.062**	0.982	1.726	**0.044**	0.993	0.840	0.202	0.996	0.725	0.235
	2q	1.005	-0.420	0.662	0.994	0.242	0.405	1.000	-0.250	0.599	0.994	0.391	0.349
	4q	0.975	0.210	0.417	0.987	-0.350	0.636	0.971	0.706	0.241	0.974	0.679	0.250
KRW	1q	0.973	2.523	**0.007**	0.974	2.817	**0.003**	0.990	1.341	**0.092**	0.995	0.876	0.192
	2q	0.965	1.674	**0.049**	0.970	2.491	**0.008**	0.990	0.927	0.179	0.993	0.838	0.202
	4q	0.979	0.665	0.254	0.984	1.244	0.109	0.986	1.048	0.149	0.991	0.858	0.197
NZD	1q	0.959	4.125	**0.000**	0.967	3.740	**0.000**	0.986	2.910	**0.002**	0.992	1.600	**0.057**
	2q	0.960	2.487	**0.008**	0.965	3.868	**0.000**	0.988	1.970	**0.026**	0.991	1.635	**0.053**
	4q	0.979	1.137	0.130	0.984	2.113	**0.019**	0.986	2.172	**0.017**	0.991	1.494	**0.070**

Ratio RMSPE–Period 1986-2018													
Country	**Horizon**	**BMA-1**	**t-stat**	**p-value**	**BMA-2**	**t-stat**	**p-value**	**BMA-3**	**t-stat**	**p-value**	**BMA-4**	**t-stat**	**p-value**
SGD	1q	0.973	0.927	0.179	0.970	1.652	**0.051**	0.974	1.159	0.125	0.980	1.693	**0.047**
	2q	1.037	-0.532	0.702	0.972	1.027	0.154	0.974	0.766	0.223	0.970	1.795	**0.038**
	4q	1.115	-1.135	0.870	0.970	1.372	**0.087**	0.954	0.974	0.167	0.955	2.184	**0.016**
SEK	1q	0.972	2.585	**0.006**	0.973	3.384	**0.001**	0.991	1.187	0.120	0.995	0.971	0.167
	2q	0.981	1.462	**0.074**	0.977	2.740	**0.004**	0.996	0.461	0.323	0.995	0.824	0.206
	4q	0.972	1.290	0.101	0.980	1.855	**0.034**	0.986	1.405	**0.082**	0.994	0.911	0.183
CHF	1q	0.966	3.472	**0.00**	0.974	3.226	**0.001**	0.981	2.473	**0.008**	0.990	1.731	**0.044**
	2q	0.957	1.915	**0.030**	0.973	2.058	**0.022**	0.974	1.719	**0.045**	0.980	1.574	**0.060**
	4q	0.939	2.103	**0.020**	0.999	-0.472	0.681	0.964	2.214	**0.015**	0.981	1.342	**0.092**
GBP	1q	0.971	2.288	**0.013**	0.975	2.153	0.017	0.994	1.017	0.156	0.998	0.299	0.383
	2q	0.967	1.430	**0.078**	0.979	1.755	0.042	0.991	0.611	0.272	0.995	-0.125	0.549
	4q	0.976	0.729	0.234	1.004	-0.708	0.759	0.991	0.613	0.271	1.001	-0.681	0.751
Avg.	1q	0.970	-	-	0.973	-	-	0.988	-	-	0.993	-	-
	2q	0.979	-	-	0.975	-	-	0.988	-	-	0.989	-	-
	4q	0.990	-	-	0.987	-	-	0.980	-	-	0.985	-	-

BMA1: ϕ*: UIPr*, ρ *=0.5* uniform to all model sizes. *BMA2:* ϕ*: UIPr*, ρ *=0.5* for model sizes *0* and *1*. *BMA3:* ϕ*: 0.5*, ρ *=0.5* uniform to all model sizes. *BMA4:* ϕ*: 0.5*, ρ *=0.5* for model sizes *0* and *1*. The t-statistics and p-values were estimated using the Diebold-Mariano test described in section Results evaluation.

Table 3.
Panel B: RMSPE ratio and significance test–Pooled data results
Panel B: Emerging countries

Ratio RMSPE–Period 1990-2018													
Country	**Horizon**	**BMA-1**	**t-stat**	**p-value**	**BMA-2**	**t-stat**	**p-value**	**BMA-3**	**t-stat**	**p-value**	**BMA-4**	**t-stat**	**p-value**
ARG	1q	0.758	1.039	0.152	0.759	1.034	0.153	0.762	1.805	**0.038**	0.796	1.991	**0.026**
	2q	0.786	1.238	0.111	0.788	1.226	0.113	0.769	2.152	**0.018**	0.777	2.295	**0.013**
	4q	0.670	2.329	**0.012**	0.670	2.327	**0.012**	0.731	3.092	**0.002**	0.758	2.857	**0.003**
BRA	1q	0.692	3.253	**0.001**	0.692	3.250	**0.001**	0.949	0.804	0.213	1.004	-0.074	0.529
	2q	0.839	2.539	**0.007**	0.840	2.532	**0.007**	1.058	-0.848	0.800	1.064	-0.937	0.824
	4q	0.902	1.194	0.119	0.902	1.193	0.119	1.102	-1.272	0.896	1.116	-1.348	0.908
COP	1q	0.702	3.942	**0.000**	0.702	3.939	**0.000**	0.966	0.575	0.284	1.006	-0.125	0.550
	2q	0.857	1.889	**0.032**	0.857	1.885	**0.032**	1.105	-1.413	0.918	1.086	-1.268	0.895
	4q	0.974	0.331	0.371	0.974	0.331	0.371	1.230	-2.495	0.992	1.196	-2.192	0.984
PHP	1q	0.890	0.967	0.169	0.890	0.967	0.169	1.342	0.967	0.169	1.310	0.967	0.169
	2q	0.966	0.308	0.380	0.966	0.313	0.378	1.661	-5.346	1.000	1.553	-4.617	1.000
	4q	0.992	0.056	0.478	0.992	0.057	0.477	1.740	-5.400	1.000	1.781	-5.510	1.000

Ratio RMSPE–Period 1990-2018													
Country	**Horizon**	**BMA-1**	**t-stat**	**p-value**	**BMA-2**	**t-stat**	**p-value**	**BMA-3**	**t-stat**	**p-value**	**BMA-4**	**t-stat**	**p-value**
IDR	1q	0.808	1.788	**0.040**	0.808	1.791	**0.040**	0.975	0.285	0.389	1.062	-0.703	0.757
	2q	0.989	0.114	0.455	0.987	0.135	0.447	1.088	-0.822	0.793	1.161	-1.495	0.929
	4q	1.252	-1.878	0.967	1.252	-1.879	0.967	1.210	-1.652	0.948	1.253	-1.932	0.971
MYR	1q	0.941	0.440	0.331	0.937	0.440	0.331	1.064	0.440	0.331	1.141	0.440	0.331
	2q	1.304	-2.296	0.987	1.295	-2.283	0.987	1.235	-2.174	0.983	1.299	-2.645	0.995
	4q	2.364	-4.881	1.000	2.362	-4.891	1.000	1.211	-1.847	0.965	1.500	-3.848	1.000
MXN	1q	0.684	2.290	**0.013**	0.684	2.290	**0.013**	0.978	0.212	0.417	1.020	-0.251	0.598
	2q	0.855	1.213	0.115	0.855	1.215	0.115	1.111	-0.936	0.823	1.100	-0.98	0.834
	4q	0.951	0.609	0.273	0.951	0.609	0.273	1.185	-1.471	0.926	1.193	-1.550	0.936
Avg.	1q	0.756	-	-	0.756	-	-	0.996	-	-	1.033	-	-
	2q	0.882	-	-	0.882	-	-	1.132	-	-	1.123	-	-
	4q	0.957	-	-	0.957	-	-	1.199	-	-	1.216	-	-

BMA1: ϕ: UIPr, *ρ =0.5* uniform to all model sizes. *BMA2: ϕ: UIPr*, *ρ =0.5* for model sizes *0* and *1*. *BMA3: ϕ: 0.5*, *ρ =0.5* uniform to all model sizes. *BMA4: ϕ: 0.5*, *ρ =0.5* for model sizes *0* and *1*. The t-statistics and p-values were estimated using the Diebold-Mariano test described in section Results evaluation.

The results of the estimations at the individual level seem to indicate that the choice of parameters for obtaining the lowest error ratio depends on the country and forecast horizon. As for inference tests, the results obtained are modest, finding some cases where the difference in favour of the Bayesian model is significant, particularly in Canada, New Zealand, Singapore and Sweden when the forecast horizon is 1 quarter. Some positive results in the panel of emerging countries are also found in Argentina, Brazil and Malaysia.[1]

As a robustness test, we conducted a direction of change test presented in an online annex. In most cases the BMA model anticipates the correct sign of the change by more than 50%, which is a positive indicator of our results.

Trained models

In this section, we analyse the trained BMA models that generated the out-of-sample forecasts described above. Based on their good relative performance, we will work with models trained using pooled data (not at the individual level). Table 6 reports the prior beliefs of the different models' specifications with respect to the expected model size (M prior), depending on the different specifications of the model. As mentioned above, models 1 and 3 assume a prior on the model size of 4 variables while the other two cases assume smaller models of a size less than 1. *Mpost* is the average a posteriori model size over the whole testing period. We notice that a posteriori beliefs vary remarkably, and one of the points we will focus on next is the differences between developed and emerging countries. For emerging countries it stands out that, regardless of the horizon and specification, there is a high probability of a posteriori inclusion for more inputs. This could indicate that the lack of knowledge about the model is greater in these cases or that information gains are more evenly distributed across the different theoretical models.

Focusing on how the a posteriori model size varies over time, we look into the BMA1 specification results. We chose this specification since it is the one that obtains the best results for both developed and emerging countries. In table 2 we note that the BMA2 specification also obtains good forecasting results in the case of emerging countries. The results with respect to model size are similar to the BMA1 specification. We chose to focus on this specification only to facilitate clarity when presenting results. Figure 1 refers to the size of the model trained a posteriori at each point in time. The graphs located in the first line correspond to developed countries and the remaining ones to emerging countries. The trained models are systematically denser in developed countries for all horizons and at all time points in the sample.

[1] To test result robustness, complementary tests were performed where the adjusted data obtained at the individual level were generated using the vector error correction methodology (VEC) instead of OLS. Similarly, forecast results obtained were less accurate than the results using pooled data.

Figure 1.
Heat map of inclusion probabilities for the BMA1 specification

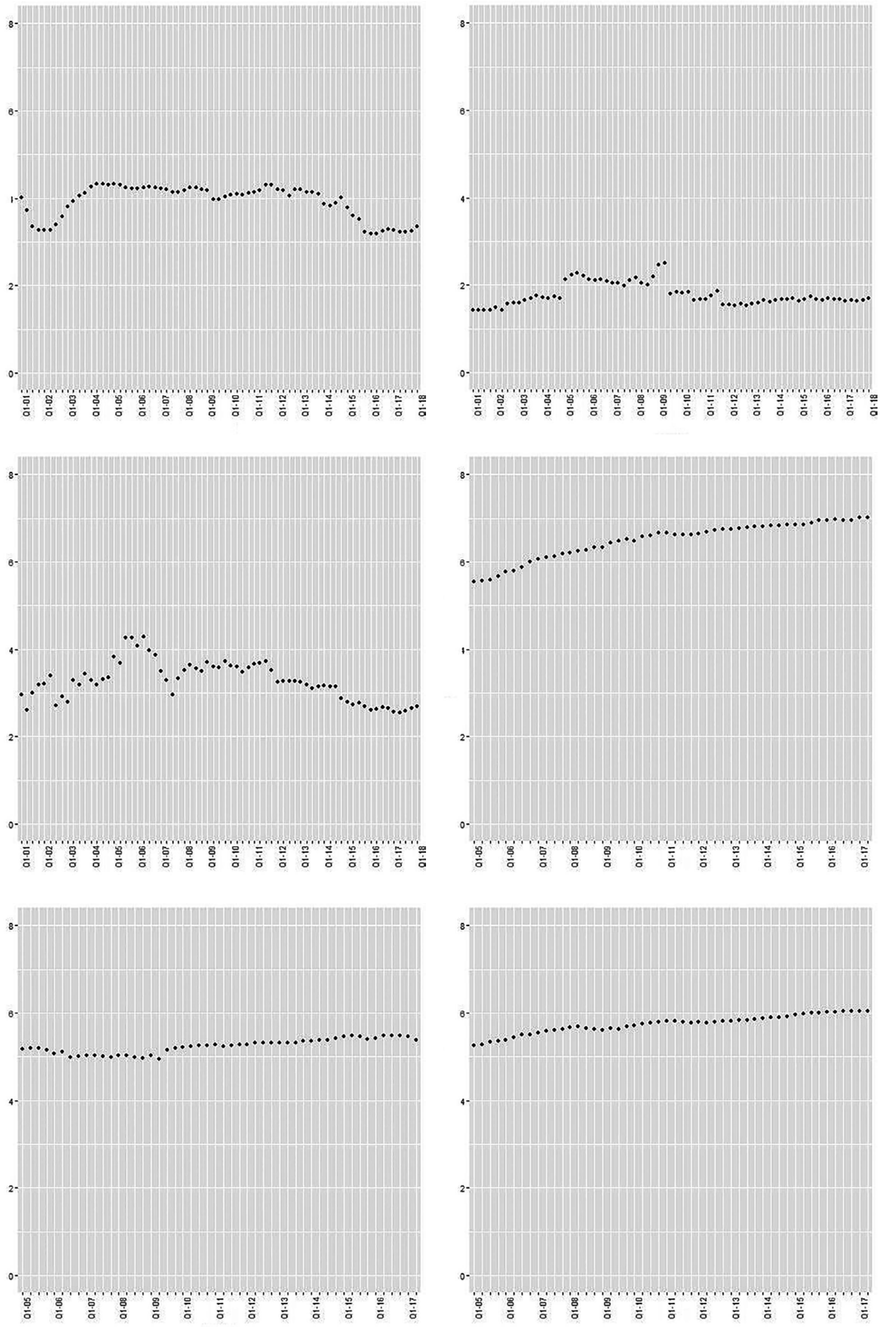


Table 4.
Panel A: RMSPE ratio and significance test–Individual data results
Panel A: Developed countries

Ratio RMSPE–Period 1986-2018													
Country	**Horizon**	**BMA-1**	**t-stat**	**p-value**	**BMA-2**	**t-stat**	**p-value**	**BMA-3**	**t-stat**	**p-value**	**BMA-4**	**t-stat**	**p-value**
AUD	1q	1.017	-0.365	0.642	1.017	-0.399	0.655	0.996	0.272	0.393	0.996	0.448	0.328
	2q	0.987	0.655	0.258	0.982	0.838	0.204	0.991	1.045	0.151	0.990	1.146	0.130
	4q	1.022	-0.517	0.697	1.001	-0.026	0.510	1.000	-0.002	0.501	0.996	0.207	0.418
CAD	1q	0.944	3.041	**0.002**	0.963	2.807	**0.003**	0.972	3.565	0.000	0.988	2.631	**0.005**
	2q	0.949	1.313	**0.099**	0.980	0.682	0.250	0.979	1.355	0.092	0.990	0.960	0.172
	4q	0.928	1.619	**0.055**	0.962	1.484	0.071	0.972	1.879	0.032	0.984	0.875	0.192
JPY	1q	1.102	-2.018	0.976	0.983	0.990	0.163	1.009	-0.501	0.691	0.991	0.814	0.209
	2q	1.206	-1.783	0.958	0.969	0.573	0.285	1.017	-0.394	0.652	0.994	0.083	0.467
	4q	0.981	0.256	0.399	0.958	0.642	0.262	0.934	1.608	**0.056**	0.973	0.761	0.224
COR	1q	1.124	-4.076	1.000	1.020	-1.833	0.965	1.040	-2.786	0.997	1.001	-0.822	0.793
	2q	1.043	-1.009	0.840	1.009	-0.237	0.593	0.997	-0.232	0.591	0.992	0.139	0.445
	4q	1.152	-1.326	0.906	1.044	-0.386	0.650	1.001	-0.014	0.506	0.983	0.312	0.378
NZD	1q	0.848	4.105	0.000	0.892	4.451	0.000	0.935	3.751	0.000	0.969	3.097	0.001
	2q	0.813	2.556	0.007	0.869	3.596	0.000	0.924	2.425	0.010	0.967	2.164	0.019
	4q	0.848	1.318	0.096	0.906	2.041	0.022	0.919	2.084	0.020	0.974	1.097	0.138

Ratio RMSPE–Period 1986-2018													
Country	**Horizon**	**BMA-1**	**t-stat**	**p-value**	**BMA-2**	**t-stat**	**p-value**	**BMA-3**	**t-stat**	**p-value**	**BMA-4**	**t-stat**	**p-value**
SGD	1q	0.977	0.668	0.253	0.970	1.399	0.083	0.987	0.379	0.353	0.981	1.086	0.140
	2q	1.064	-0.920	0.818	0.960	0.900	0.187	0.981	0.312	0.379	0.981	0.798	0.215
	4q	1.137	-0.896	0.813	0.994	0.071	0.472	0.955	0.707	0.241	0.964	0.755	0.226
SEK	1q	0.950	1.830	**0.036**	0.973	1.522	**0.066**	0.962	2.593	**0.006**	0.983	1.604	**0.057**
	2q	0.995	-0.054	0.521	0.992	1.088	0.142	0.981	1.765	**0.043**	0.997	0.543	0.295
	4q	1.027	-0.347	0.635	0.995	0.112	0.456	0.979	0.508	0.307	0.997	0.103	0.459
CHF	1q	1.093	-2.344	0.989	1.042	-1.382	0.914	1.036	-1.722	0.955	0.999	0.053	0.479
	2q	1.068	-0.359	0.639	1.011	0.697	0.245	1.009	0.382	0.352	0.988	1.013	0.159
	4q	1.010	-0.140	0.555	1.458	-0.101	0.540	0.970	0.659	0.256	0.973	0.842	0.201
GBP	1q	1.063	-2.067	0.979	1.036	-2.500	0.993	1.012	-0.947	0.826	1.002	-0.308	0.620
	2q	1.115	-2.158	0.981	1.044	-1.343	0.906	1.031	-0.919	0.818	1.007	-0.109	0.543
	4q	1.000	0.006	0.498	0.995	0.146	0.442	0.990	0.557	0.290	0.989	0.855	0.198
Avg.	1q	1.013	-	-	0.989	-	-	0.994	-	-	0.990	-	-
	2q	1.022	-	-	0.980	-	-	0.990	-	-	0.990	-	-
	4q	1.012	-	-	1.035	-	-	0.969	-	-	0.981	-	-

BMA1: ϕ: UIPr, *ρ =0.5* uniform to all model sizes. *BMA2: ϕ: UIPr*, *ρ =0.5* for model sizes *0* and *1*. *BMA3: ϕ: 0.5*, *ρ =0.5* uniform to all model sizes. *BMA4: ϕ: 0.5*, *ρ =0.5* for model sizes *0* and *1*. The t-statistics and p-values were estimated using the Diebold-Mariano test described in section Results evaluation.

Table 5.
Panel A: RMSPE ratio and significance test–Individual data results
Panel B: Emerging countries

Ratio RMSPE–Period 1990-2018													
Country	**Horizon**	**BMA-1**	**t-stat**	**p-value**	**BMA-2**	**t-stat**	**p-value**	**BMA-3**	**t-stat**	**p-value**	**BMA-4**	**t-stat**	**p-value**
ARS	1q	0.771	1.048	0.150	0.748	1.549	**0.064**	0.743	2.433	**0.009**	0.793	2.735	**0.004**
	2q	0.782	0.988	0.164	0.759	1.087	0.141	0.709	2.226	**0.015**	0.788	2.484	**0.008**
	4q	0.775	2.067	**0.022**	0.797	2.126	**0.019**	0.764	2.072	**0.021**	0.768	2.097	**0.000**
BRL	1q	0.572	2.858	**0.003**	0.838	1.151	0.127	1.331	-2.453	0.991	1.386	-2.776	0.996
	2q	1.163	-0.769	0.777	1.152	-1.151	0.873	1.705	-2.980	0.998	1.717	-2.993	0.998
	4q	1.374	-1.650	0.948	1.475	-1.657	0.948	2.356	-3.058	0.998	2.407	-3.160	0.999
COP	1q	0.967	1.443	**0.077**	0.963	1.705	**0.047**	0.988	0.595	0.277	0.999	0.032	0.487
	2q	0.944	2.022	**0.024**	0.980	0.558	0.290	0.982	0.453	0.326	1.014	-0.287	0.612
	4q	0.959	0.896	0.187	0.987	0.365	0.358	1.036	-0.376	0.646	1.063	-0.599	0.724
PHP	1q	0.981	0.299	0.383	1.034	-0.776	0.779	1.018	-0.484	0.685	1.055	-1.320	0.904
	2q	0.956	0.426	0.336	0.954	0.491	0.313	1.014	-0.313	0.622	1.100	-1.295	0.899
	4q	1.140	-1.031	0.847	1.096	-0.579	0.718	1.111	-0.894	0.812	1.146	-1.042	0.849

Ratio RMSPE–Period 1990-2018													
Country	**Horizon**	**BMA-1**	**t-stat**	**p-value**	**BMA-2**	**t-stat**	**p-value**	**BMA-3**	**t-stat**	**p-value**	**BMA-4**	**t-stat**	**p-value**
IDR	1q	1.027	-0.322	0.626	1.048	-0.649	0.740	0.994	0.099	0.461	1.001	-0.012	0.505
	2q	1.487	-2.170	0.983	1.128	-0.896	0.813	1.047	-0.368	0.643	1.053	-0.459	0.676
	4q	1.442	-1.042	0.849	1.312	-1.125	0.867	1.138	-0.548	0.707	1.150	-0.649	0.740
MYR	1q	0.929	1.727	**0.045**	0.967	1.370	**0.088**	0.950	1.852	**0.035**	0.983	0.988	0.164
	2q	0.945	1.224	0.113	0.967	0.954	0.172	0.969	0.873	0.193	0.987	0.494	0.312
	4q	0.978	0.311	0.378	0.983	0.323	0.374	0.996	0.081	0.468	1.001	-0.031	0.512
MXN	1q	0.830	1.909	**0.031**	0.984	0.318	0.376	0.937	1.027	0.155	0.988	0.241	0.405
	2q	0.972	0.325	0.373	0.987	0.149	0.441	0.985	0.169	0.433	1.000	0.001	0.500
	4q	0.942	0.364	0.359	0.948	0.328	0.372	0.994	0.035	0.486	1.005	-0.029	0.511
Avg.	1q	0.868	-	-	0.940	-	-	0.994	-	-	1.029	-	-
	2q	1.035	-	-	0.990	-	-	1.059	-	-	1.094	-	-
	4q	1.087	-	-	1.085	-	-	1.199	-	-	1.220	-	-

BMA1: ϕ*: UIPr*, ρ *=0.5* uniform to all model sizes. *BMA2:* ϕ*: UIPr*, ρ *=0.5* for model sizes *0* and *1*. *BMA3:* ϕ*: 0.5*, ρ *=0.5* uniform to all model sizes. *BMA4:* ϕ*: 0.5*, ρ *=0.5* for model sizes *0* and *1*. The t-statistics and p-values were estimated using the Diebold-Mariano test described in section Results evaluation.

Table 6.
Model size: Priors vs. posteriors
Horizon BMA-1 BMA-2 BMA-3 BMA-4

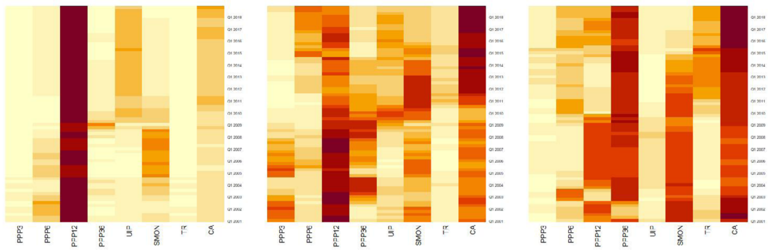
	Horizon	BMA-1	BMA-2	BMA-3	BMA-4
M prior	all	4	0.9	4	0.9
M post					
Developed	1q	1.8	0.8	4.7	0.9
	2q	3.2	0.9	5.2	1.0
	4q	3.9	1.1	5.4	1.0
Emerging	1q	5.3	5.3	5.8	1.1
	2q	5.8	5.8	5.7	1.0
	4q	6.6	6.6	5.9	1.0

M prior reports the prior beliefs implied by ρ. BMA-1 and BMA-3 assign equal probability to all model sizes leading to and expected modelsizeof4. BMA-2 and BMA4 penalise bigger models leading to an expected model size of 0.99. Models were trained using pool data by country group. Mpost is the average a posteriori model size over the whole testing period. BMA1: ϕ: UIPr, $\rho = 0.5$ uniform to all model sizes. BMA2: ϕ: UIPr, $\rho = 0.5$ for model sizes 0 and 1. BMA3: ϕ: 0.5, $\rho = 0.5$ uniform to all model sizes. BMA4: ϕ: 0.5, $\rho = 0.5$ for model sizes 0 and 1.

Another dimension of analysis focuses on which in-sample forecasts generated under theoretical models the BMA model incorporates. For this purpose, table 4 presents the average a posteriori inclusion probabilities for all cases. In the developed countries model, the price growth differential model containing information for the last 4 quarters stands out strongly, while for emerging countries the interest rate differential model is incorporated with high probability in the different estimation horizons. An interesting case to highlight is the model that includes information on the evolution of the current account. In this case, its probability of inclusion increases significantly at longer horizons.

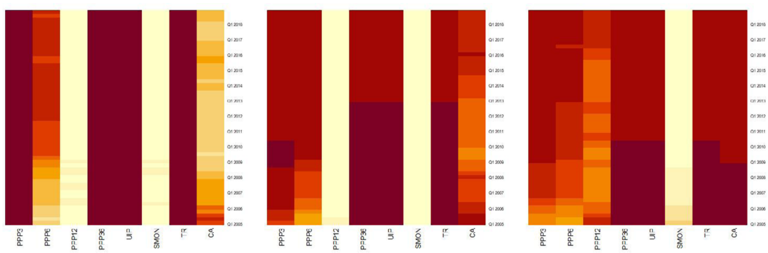
Again, to describe the changes in the trained model over the period we focus on the BMA1specification. Figure 2 presents the heat maps of a posteriori inclusion probabilities for the inputs of each theoretical model for each of the horizon combinations. Light colours refer to lower a posteriori inclusion probabilities. The higher presence of light colours in the first line of graphs refers to the more dispersed pattern of the trained model in developed countries. The opposite is seen in the second line of graphs for emerging countries. Although outside the main objective of this paper, one line of research that emerges from these results is to delve deeper into how changes in monetary policy, for example, relates to changes in a posteriori inclusion probabilities.

Figure 2.
Heat map of inclusion probabilities for the BMA1 specification



1 quarter **2 quarters** **4 quarters**

Panel A: Developed countries



1 quarter **2 quarters** **4 quarters**

Panel B: Emerging countries

Table 7.
A posteriori inclusion probabilities in BMA1 model (average)

		PPP3	PPP6	PPP12	PPP36	UIP	SMON	TR	CA
Panel A	1 quarter	0.05	0.14	0.80	0.10	0.20	0.21	0.05	0.22
	2 quarters	0.26	0.25	0.75	0.41	0.29	0.46	0.15	0.66
	4 quarters	0.10	0.31	0.54	0.89	0.13	0.66	0.31	0.94
Panel B	1 quarter	1.00	0.68	0.08	1.00	1.00	0.09	1.00	0.41
	2 quarters	0.97	0.89	0.08	1.00	1.00	0.07	1.00	0.79
	4 quarters	0.89	0.85	0.73	1.00	1.00	0.12	1.00	0.99

The table shows the a posteriori inclusion probabilities for each of the inputs belonging to the theoretical models entering the BMA1 model with pooled data by panel. The inclusion probabilities are averaged over the entire testing sample period. The maximum value includes values approximately equal to unity.

CONCLUSIONS

The objective of this paper was to evaluate the performance of exchange rate forecasting using the Bayesian model averaging methodology. A contribution of the exercise arises from the use of adjusted data obtained from different economic models supported by the theory as inputs of the BMA model. In this way, it is possible to study the forecasting gains generated by combining information from different theoretical models. Significant forecast improvements are found with respect to the random walk. These improvements particularly tend to stand out for developed countries and shorter horizons. Comparing the results at both the individual level and group pool level indicates that training models using data from other countries helps to improve the performance of BMA models. A regularity found for both country groups is that the BMA1 model tends to perform better. This model is the least restrictive in terms of both coefficient constraint and model size, indicating that the in-sample forecasts we generate under different theoretical models have relevant information for exchange rate estimation. This also indicates that forecast combination is a useful strategy for exchange rate estimation.

Finally, looking at the characteristics of BMA models indicates that the probability of inclusion of different in-sample forecasts in developed countries is more concentrated while, on average, BMA models from emerging countries tend to have a larger ex-post size. This may be a sign of the difficulty of learning models in highly volatile contexts, where the connection between exchange rates and its fundamentals may be less informative.

This work could be extended in many directions. As a methodological extension, it would be interesting to include nonlinear models and other technical improvements when we obtain fitted data from theoretical models. The use of nonlinearities could also be considered when combining information. Finally, it would be useful to deepen the relationship between the theoretical models' performance and the underlying behaviour of different economies over time. For example, to study the relationship in which different monetary policies result in better or worse forecasting results for each of the theoretical models. This would subsequently imply different inclusion probabilities throughout the sample period.

REFERENCES

1. Alquist, R., & Chinn, M. D. (2008). Conventional and unconventional approaches to exchange rate modelling and assessment. *International Journal of Finance & Economics*, *13*(1), 2-13.
2. Cassel, G. (1918). Abnormal deviations in international exchanges. *The Economic Journal*, *28*(112), 413-415.
3. Cheung, Y.-W., Chinn, M. D., & Pascual, A. G. (2005). Empirical exchange rate models of the nineties: Are any fit to survive? *Journal of International Money and Finance*, *24*(7), 1150-1175.

4. Cheung, Y.-W., Chinn, M. D., Pascual, A. G., & Zhang, Y. (2019). Exchange rate prediction redux: New models, new data, new currencies. *Journal of International Money and Finance*, 95, 332-362.
5. Chinn, M. D., & Meredith, G. (2004). Monetary policy and long-horizon uncovered interest parity. *IMF Staff Papers*, *51*(3), 409-430.
6. Ciccone, A., & Jarociński, M. (2010). Determinants of economic growth: Will data tell? *American Economic Journal: Macroeconomics*, *2*(4), 222-246.
7. Della Corte, P., Sarno, L., & Sestieri, G. (2012). The predictive information content of external imbalances for exchange rate returns: How much is it worth? *Review of Economics and Statistics*, *94*(1),100-115.
8. Della Corte, P., Sarno, L., & Tsiakas, I. (2009). An economic evaluation of empirical exchange rate models. *The Review of Financial Studies*, 22(9), 3491-3530.
9. Diebold, F. X., & Mariano, R. S. (1995). Com paring predictive accuracy. *Journal of Business and Economic Statistics*, *13*(3), 253-263.
10. Dornbusch, R. (1976). Expectations and exchange rate dynamics. *Journal of Political Economy*, *84*(6), 1161-1176.
11. Fernandez, C., Ley, E., & Steel, M. F. (2001). Benchmark priors for Bayesian model averaging. *Journal of Econometrics*, *100*(2), 381-427.
12. Fisher, I. (1896). *Appreciation and interest: A study of the influence of monetary appreciation and depreciation on the rate of interest with applications to the bimetallic controversy and the theory of interest*. American Economic Association.
13. Frankel, J. A. (1979). On the mark: A theory of floating exchange rates based on real interest differentials. *The American Economic Review*, *69*(4), 610-622.
14. Frenkel, J. A. (1976). A monetary approach to the exchange rate: Doctrinal aspects and empirical evidence. *Scandinavian Journal of Economics*, *75*(2), 200-224.
15. Giannone, D., Lenza, M., & Primiceri, G. E. (2021). Economic predictions with big data: The illusion of sparsity. *Econometrica*, *89*(5), 2409-2437.
16. Gourinchas, P.-O., & Rey, H. (2007). International financial adjustment. *Journal of Political Economy*, *115*(4), 665-703.
17. Hodrick, R. J., & Prescott, E. C. (1997). Postwar US business cycles: An empirical investigation. *Journal of Money, Credit, and Banking*, *29*(1), 1-16.
18. Li, J., & Chen, W. (2014). Forecasting macroeconomic time series: Lasso-based approaches and their forecast combinations with dynamic factor models. *International Journal of Forecasting*, *30*(4), 996-1015.
19. Mark, N. C. (1995). Exchange rates and fundamentals: Evidence on long-horizon predictability. *The American Economic Review*, 85(1), 201-218.

20. Mark, N. C., & Sul, D. (2001). Nominal exchange rates and monetary fundamentals: Evidence from a small post-bretton woods panel. *Journal of International Economics*, *53*(1), 29-52.
21. Medel, C., Camilleri, G., Hsu, H.-L., Kania, S., & Touloumtzoglou, M. (2016). *Robustness in foreign exchange rate forecasting models: Economics-based modelling after the financial crisis* (Working Paper 784). Central Bank of Chile.
22. Meese, R. A., & Rogoff, K. (1983). Empirical exchange rate models of the seventies: Do they fit out of sample? *Journal of International Economics*, *14*(1-2), 3-24.
23. Molodtsova, T., & Papell, D. H. (2009). Out-of-sample exchange rate predictability with Taylor rule fundamentals. *Journal of International Economics*, *77*(2), 167-180.
24. Mussa, M. (1982). A model of exchange rate dynamics. *Journal of Political Economy*, *90*(1), 74-104.
25. Pesaran, M. H., Schuermann, T., & Smith, L. V. (2009). Forecasting economic and financial variables with global vars. *International Journal of Forecasting*, *25*(4), 642-675.
26. Rossi, B. (2013). Exchange rate predictability. *Journal of Economic Literature*, *51*(4), 1063-1119.
27. Stock, J.H., & Watson, M.W. (2005). An empirical comparison of methods for forecasting using many predictors. *Manuscript, Princeton University*, 46.
28. Taylor, J. B. (1993). Discretion versus policy rules in practice. *Carnegie-Rochester Conference Series on Public Policy*, 39, 195-214.
29. Timmermann, A. (2006). Forecast combinations. In G. Elliott, C. Granger, & A. Timmermann (eds.), *Handbook of Economic Forecasting* (vol. 1, pp. 135-196). Elsevier.
30. Wright, J. H. (2008). Bayesian model averaging and exchange rate forecasts. *Journal of Econometrics*, *146*(2), 329-341.
31. Zellner, A. (1971). *An introduction to Bayesian inference in econometrics*. Wiley.

ANNEX

Table 8.
Details of dropped observations for each country

Country	Dropped obs.	N of obs
AUD		132
CAD		132
JPY		132
KRW		132
NZD		132
SGD		132
SEK	Q1 1986–Q4 1992	104
CHF	Q1 2013–Q4 2014	124
GBP		132
ARS	Q1 1991- Q4 2001	72
BRL		116
COP		116
PHP		116
IDR	Q1 1990–Q4 1997	84
MYR	Q1 1998–Q3 2005	85
MXN	Q1 1990–Q3 1994	97

Note: The sample covers the periods of Q1 1986-Q4 2018 and Q1 1990-Q4 2018 for developed and emerging countries, respectively. Observations related to periods of fixed exchange rate regimes were not included in the sample. Identification of those exchange rate regimes was based on the IMF Annual Report.

Since Q4 2011, Argentina has strict exchange controls. Results presented here include this period. However, robustness exercises were carried out eliminating these observations and we found that results are not altered.

92

CUADERNOS DE ECONOMÍA

ISSN 0121-4772

ARTÍCULOS

ISSN 0121-4772
9 770121 477005 92