Model Uncertainty and Exchange Rate Forecasting

Roy Kouwenberg, Agnieszka Markiewicz, Ralph Verhoeks, and
Remco C. J. Zwinkels*

Abstract

Exchange rate models with uncertain and incomplete information predict that investors focus on a small set of fundamentals that changes frequently over time. We design a model selection rule that captures the current set of fundamentals that best predicts the exchange rate. Out-of-sample tests show that the forecasts made by this rule significantly beat a random walk for 5 out of 10 currencies. Furthermore, the currency forecasts generate meaningful investment profits. We demonstrate that the strong performance of the model selection rule is driven by time-varying weights attached to a small set of fundamentals, in line with theory.

I.   Introduction

After decades of disappointing results, exchange rate economics has lately experienced a revival. Several empirical studies claim to have found a relation between exchange rates and macroeconomic fundamentals, although these relations are often unstable and short-lived (Rossi (2013)). Renewed interest in the field has also been triggered by the development of new exchange rate models in which investors face uncertainty about the factors driving currency movements and their relative importance (e.g., Bacchetta and van Wincoop (2004)). These new theoretical models predict that the relation between the exchange rate and

macro fundamentals is time varying, and a small number of fundamentals can receive "excessive" weights relative to a rational, perfect knowledge equilibrium.

In this article, we design a dynamic model selection rule that encompasses and tests these predictions. The model selection rule combines forecasts from different fundamental models with given weights, following the literature on forecast combinations (Timmermann (2006)). Although the empirical literature has shown that combinations of exchange rate forecasts perform better than individual fundamental-based models (e.g., Della Corte, Sarno, and Tsiakas (2009)), it is not clear why. This article fills in this gap. We derive the key properties of a forecast combination rule from incomplete information models, thereby providing explanations as to why this rule can predict exchange rates well empirically.

We allow model weights to vary through time to capture the dynamic relation between exchange rates and fundamentals. Furthermore, variables with insignificant forecasting power are eliminated by backward regression, so that only a few, the most relevant, fundamentals remain in the forecasting equation. As the model selection procedure is repeated at each point in time, the set of selected fundamental variables can adapt quickly to changes in the underlying exchange rate process.

We examine the out-of-sample performance of the forecast combination rule based on backward elimination (BE) in both statistical and economic terms. We use a quarterly real-time data set with observations between 1973Q1 and 2014Q1 for 10 currency pairs, with the U.S. dollar (USD) as numeraire, including the Australian dollar (AUD), Canadian dollar (CAD), New Zealand dollar (NZD), Brazilian real (BRL), Japanese yen (JPY), Mexican peso (MXN), Norwegian krone (NOK), Russian ruble (RUB), Swiss franc (CHK), and U.K. pound sterling (GBP).

We find that the BE rule systematically outperforms all benchmarks we compare it to, in both statistical and economic terms. It significantly beats a random walk for 5 out of 10 currencies, in both magnitude and direction. It yields a positive and significant excess return for 3 currency investment strategies based on the forecasts. The magnitude of the investment profits is considerable; for example, an equal-weight portfolio of 10 currency positions based on the BE rule achieves a Sharpe ratio of 0.81, whereas in our sample a carry-trade strategy with 3 long and short currencies delivers a Sharpe ratio of 0.34, and a random walk with drift earns no excess returns.

We examine in detail the sources of the strong forecasting performance of the BE rule. We find that this performance is driven by time-varying and excessive weights on fundamental models, consistent with the theories on model uncertainty.

This article is closely related to the theoretical literature on imperfect knowledge, recently introduced in the context of exchange rate modeling by Bacchetta and van Wincoop (2004), (2006), and (2013) and Markiewicz (2012). Bacchetta and van Wincoop propose a framework in which investors attribute exchange rate movements to a scapegoat variable, due to incomplete information. In Markiewicz, investors face model uncertainty, which leads them to employ statistical model learning. Model learning can also lead to excessive weight on one or more fundamentals. We use these two frameworks to derive our hypotheses.

Fratzscher, Rime, Sarno, and Zinna (2015) test the scapegoat theory for 12 currencies and report robust empirical evidence in sample. The out-of-sample test we carry out provides additional support for the theory's predictions of time-varying and excessive fundamental weights.

Our article is also related to the vast empirical literature on exchange rate forecasting using fundamental-based models. Early papers such as Meese and Rogoff (1983) and Cheung, Chinn, and Pascual (2005) claim that fundamental-based models have no systematic predictive ability. Recent studies, however, refute this fact. For example, Molodtsova and Papell (2009) show that the Taylor rule model of the exchange rate can beat a random walk in an out-of-sample forecasting test.[1] Sarno and Valente (2009) examine a large set of dynamic model selection criteria and find that the information content of macroeconomic and financial variables is relevant for exchange rate forecasting. They do not, however, find a model selection criterion that can successfully identify the correct model in advance and generate accurate out-of-sample forecasts. In this article, we implement a dynamic model selection scheme that is able to beat a random walk.

Only recently, combined forecasts and dynamic model selection criteria have been applied successfully to predict exchange rates.[2] Della Corte et al. (2009) provide the first evidence that combining fundamental-based forecasts, using Bayesian model averaging, can beat a random walk and generate economically meaningful profits. Subsequent studies by Della Corte and Tsiakas (2012), Morales-Arias and Moura (2013), and Garratt and Mise (2014) confirm that various other forecast combination rules also improve exchange rate predictions. Our work provides a better understanding of why these forecasting methods work well in the foreign currency market, based on theory.

The remainder of the article is organized as follows: Section II reviews empirical evidence and recent theoretical models of the time-varying relation between exchange rates and fundamentals. In Section II we also derive testable hypotheses from these models. In Section III we propose our empirical forecasting strategy and describe the fundamental-based models and the real-time data set. Section IV presents the main results, Section V shows robustness tests, and Section VI concludes.

II.   Time-Varying Fundamentals

Several studies in the empirical exchange rate literature document that the relation between exchange rates and macroeconomic fundamentals is unstable over time.[3] Findings from surveys of foreign exchange traders suggest a plausible

---

[1]In addition, Ferraro, Rogoff, and Rossi (2015) find that oil prices can predict the CAD exchange rate at a daily frequency. Bakshi and Panayotov (2013) show that carry-trade returns can be predicted with changes in volatility, commodity prices, and a liquidity factor.

[2]Meese and Rogoff (1983) apply the method of Granger and Newbold (1986) for combining individual model forecasts but find little improvement in forecasting results. Similarly, Wright (2008) uses Bayesian model averaging techniques and finds that they deliver only slightly better forecasts than a random walk.

[3]Schinasi and Swamy (1989) show that exchange rate models with time-varying parameters outperform a random walk in an out-of-sample forecasting test. Cheung et al. (2005), Rossi (2006), and

explanation for the time-varying link between fundamentals and exchange rates: The importance attached by market participants to individual macroeconomic indicators varies through time, as documented by Cheung and Chinn (2001). Every change in the fundamentals' importance (weight) can feed back into the exchange rate and change its behavior. This idea has been formalized only recently, first by Bacchetta and van Wincoop (2004), (2006), and (2013) and more recently by Markiewicz (2012).

Bacchetta and van Wincoop (2004), (2006), and (2013) propose the scapegoat theory in which investors have incomplete and heterogeneous information about underlying model parameters. In their model, the exchange rate movement is generated by a shift in an unobserved fundamental (e.g., liquidity trades). As the market participants search for an explanation of the exchange rate change, they may attribute this movement to an observed macro indicator that changed at the same time and is currently out of equilibrium. The chosen macro indicator then becomes a natural scapegoat and influences investors' trading strategies. Over time, different observed variables can be taken as scapegoats, so that the weights attributed to fundamentals become time varying.

Markiewicz (2012) proposes an alternative framework based on model learning theory. In this framework, investors do not know the set of fundamentals in the true exchange rate model, and they deal with this uncertainty by estimating a large model with all fundamentals and its parameters with the available historical data. They make forecasts based on the selected significant macro variables, which in turn feed back into the actual exchange rate process through investors' trading strategies. As a result, model learning can generate excessive weights on a subset of macro variables, depending on the model chosen by investors. As investors continuously attempt to discover the true model, the variables driving the exchange rate dynamics shift over time.

The theoretical and empirical studies suggest two main testable hypotheses. First, exchange rate returns are driven by a small subset of fundamentals selected by investors, and these fundamentals receive excessive weights relative to a full information equilibrium. Specifically, because of parameter uncertainty and incomplete information, investors focus only on a subset of the available fundamentals, namely, those that recently predicted the exchange rate well. Because of the self-referential structure of asset pricing models, the exchange rate itself is then driven by this subset of fundamentals.

Second, the relation between exchange rates and fundamentals is time varying. As new information becomes available, investors frequently change the set of fundamentals, for example, because new scapegoats have been identified or as a result of model learning. As investors switch their focus from one fundamental to another through time and adjust their trading behavior, the relation between the exchange rate and the fundamentals also changes frequently.

Sarno and Valente (2009) also find that fundamental exchange rate models have predictive power, but their performance depends strongly on the particular currency and forecast horizon considered.

III.   Model Selection by Backward Elimination

Our analysis starts with a large set of fundamental-based models for the exchange rate: for example, purchasing power parity, the canonical monetary model, and uncovered interest rate parity, among others. In Table 1 and Section III.B, we specify and describe these traditional fundamental exchange rate models in detail. After estimating each fundamental model individually on an initial holdout sample consisting of 25 years of quarterly data, we generate one out-of-sample exchange rate prediction with each model. We seek to efficiently combine these individual fundamental-based predictions into one overall combination forecast in a way that reflects our hypotheses.

There is a large pool of existing forecast combination methods to choose from in the literature (e.g., Newbold and Harvey (2002), Timmermann (2006)). For example, taking a simple average of all available model forecasts is a standard method shown to reduce forecast errors in many settings. However, we require a more dynamic method, as we aim to design a forecast combination rule that encompasses the two hypotheses of the previous section, namely, that the exchange rate is driven by a small, time-varying set of fundamentals.

The models developed by Bacchetta and van Wincoop (2004), (2006), and (2013) and Markiewicz (2012) suggest that the exchange rate is typically driven by only a subset of the available fundamentals, namely, those that recently predicted the exchange rate well. We therefore analyze the out-of-sample prediction errors that the exchange rate models made in the last 20 quarters, trying to select those models that are best performing. In practice, we regress the actual exchange rate on the predictions of all the models jointly, in the last 20 quarters. Based on the coefficients obtained from this regression, we eliminate insignificant forecasts by the so-called backward elimination of regressors method.[4] We choose a backward regression for removing insignificant fundamentals because this method follows the general-to-simple model-building strategy and therefore reduces omitted-variable bias (Hendry (1995)).

As the backward model selection procedure is repeated every period (quarter), the weights on fundamentals change through time, depending on their prediction performance. In sum, the proposed BE rule selects a small set of fundamental models based on its recent prediction performance and can account for i) time variation in the fundamentals driving the exchange rate and ii) excessive weights on a few selected fundamentals. The details of the forecast combination rule based on BE follow.

A.   The Exchange Rate Forecasting Procedure

Step 1. In-Sample Estimation of the Models

We estimate all the exchange rate models listed in Table 1 (see Section III.B) individually using data in the in-sample estimation period $[t_0, t_1]$ from time t_0 to t_1, where initially $t_0 = 1973Q1$ and $t_1 = 1998Q3$. The first estimation sample thus contains up to 102 quarterly observations, depending on data availability.

---

[4]A detailed description of backward regression and other stepwise regression methods can be found in Lai and Xing (2008).

Subsequently, the 14 exchange rate models are reestimated using an expanding window in each cycle.

Step 2. Generating Quasi Out-of-Sample Forecasts with Individual Models

At the end of the in-sample estimation period $[t_0, t_1]$, we generate one-step-ahead forecasts $E^i_{t_1}(\Delta s_{t_1+1})$ for time $t_1 + 1$, with $i = 1, 2, \ldots, N$ indicating the fundamental exchange rate model and $N = 14$.[5] Steps 1 and 2 are now repeated another 19 times, with t_1 moved forward 1 period in every iteration, until we have created 20 out-of-sample forecasts from $t_2 = 1998Q4$ until $t_3 = 2003Q3$. We refer to $[t_2, t_3]$ as the out-of-sample forecast evaluation period.

Step 3. Model Selection Based on Forecast Errors and BE

In this step, we compare the forecast errors of the 14 models in the evaluation period $[t_2, t_3]$, selecting the fundamental models that make the best predictions of the exchange rate. We do this by regressing the actual (realized) exchange rate returns in $[t_2, t_3]$ on the 14 model forecast series $E^i_{t-1}(\Delta s_t)$, using a multivariate regression model:

$$(1) \qquad \Delta s_t \quad = \quad \alpha + \sum_{i=1}^{N} \beta_{i,t} E^i_{t-1}(\Delta s_t) + \epsilon_t$$

for $t = t_2, t_2 + 1, \ldots, t_3$, and $E^i_{t-1}(\Delta s_t)$ is the one-step-ahead fundamental (i)-based forecast at time $t - 1$. In this step, we estimate the weights, $\beta_{i,t}$, that should be attributed to each fundamental-based forecast. We first include all 14 fundamental models in the regression equation. Second, we eliminate insignificant variables one by one (stepwise), effectively setting the estimated coefficients to 0 ($\widehat{\beta}_{i,t} = 0$), using a p-value of 5% as the cutoff for significance. Convergence is reached when either all remaining fundamental-based forecasts have significant coefficients or a random walk model with drift α remains. Given that the evaluation period $[t_2, t_3]$ covers 20 quarterly observations in the first cycle, this procedure is initially applied to 20 observations and subsequently to an expanding window.

Step 4. Generate One Out-of-Sample Combination Forecast

In this final step, we create one out-of-sample combination forecast for time $t_3 + 1$ by using the estimated $\widehat{\beta}_{i,t}$ coefficients from the final regression model in step 3:

$$(2) \qquad E^{BE}_t(\Delta s_{t+1}) \quad = \quad \sum_{i=1}^{N} \widehat{\beta}_{i,t} E^i_t(\Delta s_{t+1}),$$

where E^{BE}_t stands for the one-period-ahead forecast based on the BE weights at time $t = t_3$. Note that potentially several of the $\widehat{\beta}_{i,t}$ coefficients are equal to 0, depending on the outcome of the BE procedure.

Steps 1, 2, 3, and 4 are now repeated, each time adding one more observation to the in-sample estimation window $[t_0, t_1]$ by increasing t_1 by 1, and adding one

---

[5]Although the 14 fundamental-based models can include multiple macroeconomic and financial variables, each model produces only one forecast i.

more observation to the out-of-sample model selection period $[t_2, t_3]$ by increasing t_3 by 1. Eventually, the BE procedure makes 43 out-of-sample forecasts of the exchange rate, covering the period from 2003Q4 through 2014Q2. We use these 43 forecasts to evaluate the performance of the BE rule.

B.   Pool of Models

We select a large set of macroeconomic and financial fundamentals commonly used in the literature to predict the exchange rate. The entire set of exchange rate models is summarized in Table 1, where Panel A shows the macroeconomic fundamental-based models, and Panel B shows models based on financial factors such as liquidity. All variables shown in Table 1 are in natural logarithms, such that Δs_t represents the period t exchange rate return. An increase in Δs_t corresponds to a depreciation of the U.S. dollar.

TABLE 1

Macroeconomic and Financial Exchange Rate Models

In Table 1, * denotes foreign variables. $i_t^{US} - i_t^*$ is the interest rate differential. $f_t^m = (m_t^* - m_t^{US}) - h(y_t^* - y_t^{US})$, where m is the log money supply and y is log output (gross domestic product (GDP)). Following Molodtsova and Papell (2009), we set h equal to 0.5. $(p_t^{US} - p_t^*) - s_t$ stands for the deviation of the nominal exchange rate from the purchasing power parity (PPP) level, where p_t^{US} and p_t^* are the U.S. and foreign price levels, respectively. CG is (yearly) consumption growth, and TB is trade balance. PCM (SCM) stands for the total purchases (sales) of securities from the foreign country to the United States. $\tilde{y}_t$ is the output gap, and π_t^{US} and π_t^* represent the U.S. and foreign inflation rates. COM$_t$ stands for commodity price index. OIL$_t$ stands for the oil price. TED$_t$ is the U.S. TED spread, which is computed as the spread between the 3-month interbank rate and the 3-month Treasury yield. UIP stands for uncovered interest rate parity, and FX stands for foreign exchange.

Model	Variables	Specification
Panel A. Macroeconomic Fundamental-Based Models		
UIP	Interest rate	$\alpha_{UIP} + \beta_{UIP}\,(i_t^{US} - i_t^*)$
Monetary model	Money supply, real output	$\alpha_{MM} + \beta_{MM}\,(f_t^m - s_t)$
Diff. monetary model	Growth in money and output	$\alpha_{\Delta m} + \beta_{\Delta m}\,(f_t^m - f_{t-1}^m)$
PPP	Deviation from PPP	$\alpha_{PPP} + \beta_{PPP}[(p_t^{US} - p_t^*) - s_t]$
Risk-sharing model	Consumption growth	$\alpha_{CG} + \beta_{CG}\,(CG_t^{US} - CG_t^*)$
Trade balance model	Trade balance	$\alpha_{TB} + \beta_{TB1}\left(TB_t^{US}/GDP_t^{US}\right)$ $+ \beta_{TB2}(TB_t^*/GDP_t^*)$
Capital flows	Net foreign assets	$\alpha_{NFA} + \beta_{NFA}\,(SCM_t - PCM_t)$
Taylor rule model	Output gap, inflation	$\alpha_{TR} + \beta_{TR1}(\tilde{y}_t^{US} - \tilde{y}_t^*) + \beta_{TR2}(\pi_t^{US} - \pi_t^*)$
Asymmetric Taylor rule model	Output gap, inflation	$\alpha_{AT} + \beta_{AT1}(\tilde{y}_t^{US} - \tilde{y}_t^*) + \beta_{AT2}(\pi_t^{US} - \pi_t^*)$ $+ \beta_{AT3}(s_t - (p_t^{US} - p_t^*))$
Panel B. Financial Factor Models		
Momentum model	AR(1)	$\alpha_{AR} + \beta_{AR}\Delta s_t$
Commodity prices	Commodity price index	$\alpha_{COM} + \beta_{COM}\Delta COM_t$
Oil price	Oil price index	$\alpha_{OIL} + \beta_{OIL}\Delta OIL_t$
Exchange rate volatility	Std. dev. of FX returns	$\alpha_v + \beta_v \sigma_t$
Liquidity	TED spread	$\alpha_{TED} + \beta_{TED}TED_t$

The exchange rate models displayed in Panel A of Table 1 are based on a set of macroeconomic fundamentals that is expected to drive and therefore predict exchange rate movements. The first model is uncovered interest rate parity (UIP), which is derived from a no-arbitrage condition. The value of β_{UIP} is unrestricted so that its estimated value can reflect either UIP (positive) or the carry trade (negative).

The second and third models in Table 1 belong to the class of canonical monetary models, which describe how the exchange rate reacts to changes in real

output and money supply. Any shift in these two variables is expected to change the demand for the currency and therefore change its price.

The fourth model is based on purchasing power parity (PPP), the relation between prices and nominal exchange rates derived from the law of one price. The PPP model postulates that in the absence of impediments to trade, goods' prices should be the same across borders. Any deviation of the nominal exchange rate s_t from the relative price level $(p_t^{US} - p_t^*)$ should be corrected.[6]

The international risk-sharing model, the fifth model in Table 1, predicts a relation between relative consumption growth and the exchange rate (see, e.g., Benigno and Thoenissen (2008), Sarkissian (2003)). The trade balance model and the capital flows model, the sixth and seventh models, both derive from the assumption that the exchange rate changes its value in response to imbalances in the markets for goods and capital.

Finally, in Taylor rule models, the exchange rate is driven by shifts in monetary policy in response to changing economic conditions, represented by the output gap and inflation. Furthermore, the asymmetric Taylor rule model assumes that one of the central banks adjusts the interest rate to accommodate exchange rate fluctuations.[7]

Financial models are displayed in Panel B of Table 1 and they incorporate variables that may influence the investment strategies of foreign currency market participants and, thus, predict the exchange rate itself. The first model is an AR(1) process, reflecting possible momentum trading strategies.[8] The second and third models are based on commodity price movements, which have been shown to be strongly related to commodity currencies (see Chen and Rogoff (2003), Chen, Rogoff, and Rossi (2010)). The fourth model is based on shifts in exchange rate volatility, and the fifth model includes a measure for global market liquidity. Bakshi and Panayotov (2013) show that the last four models can predict the return of carry-trade strategies.

C.   Forecast Accuracy Measures

Forecast accuracy is evaluated using several metrics. First, we measure the forecast accuracy of model i by the mean square prediction error (MSPE) relative to the MSPE of the random walk model. In comparing the forecast errors, we use the Clark and West (CW) (2007) adjustment and statistic, given by

$$(3) \qquad \mathrm{CW}_i \quad = \quad \frac{1}{l} \sum_{t=1}^{l} \frac{(E_{t-1}^i(\Delta s_t) - \Delta s_t)^2 - E_{t-1}^i(\Delta s_t)^2}{(\Delta s_t)^2},$$

in which l is the total number of forecasts. We employ the CW statistic to test the null hypothesis that the BE-based forecasts have the same predictive ability as a random walk without drift benchmark ("no predictability").

---

[6]Mark (1995) shows that PPP holds in the long run. Pojarliev and Levich (2010) demonstrate that currency fund managers rely on the PPP model.

[7]Molodtsova and Papell (2009) show evidence of exchange rate predictability with a Taylor rule model for 11 out of 12 currencies in the post–Bretton Woods period.

[8]See Menkhoff, Sarno, Schmeling, and Schrimpf (2012b) for momentum strategies in the foreign exchange markets.

In addition, we calculate the nonparametric sign test of Pesaran and Timmermann (PT) (1992), which tests the ability of the fundamental model to forecast the direction of change correctly, again relative to the random walk forecast. The PT statistic is given by

$$(4) \qquad \mathrm{PT}_i \quad = \quad \left(\frac{p^*(1-p^*)}{l} \right)^{-1/2} (\widehat{p} - p^*),$$

in which p^* is the benchmark proportion of correct sign predictions and $\widehat{p}$ is the observed proportion of correct sign predictions. We set $p^*=0.5$ to represent the random walk without drift model. The PT statistic is asymptotically distributed as $N(0,1)$.

Apart from the statistical forecast evaluation measures, we examine the economic value added of the model selection rules. First, we calculate the returns of an investment strategy that buys (sells) 1 unit of the foreign currency vis-à-vis the U.S. dollar when the model predicts an appreciation (depreciation) of the foreign currency. In case of subsequent buy signals, the long position is rolled over. We calculate two types of returns to this strategy: with and without the interest rate differentials. The latter does not take deposit interest rates into account when calculating investment strategy returns, such that it captures only the foreign exchange rate returns. The raw foreign-exchange return of the strategy i is given by

$$(5) \qquad r_{1,t}^i \quad = \quad \frac{E_{t-1}^i(\Delta s_t)}{\left| E_{t-1}^i(\Delta s_t) \right|} \Delta s_t.$$

The second measure does include the interest rate differential and thus implies borrowing in the currency on the short end and lending in the currency on the long end of the investment. This is equivalent to buying (selling) a foreign currency forward when the model predicts an appreciation (depreciation) of the foreign currency vis-à-vis the U.S. dollar. The return is given by

$$(6) \qquad r_{2,t}^i \quad = \quad \frac{E_{t-1}^i(\Delta s_t)}{\left| E_{t-1}^i(\Delta s_t) \right|} (s_t - f_{t-1}),$$

where f_{t-1} is the log 3-month forward exchange rate in period $t-1$.

Second, we form equal-weighted and volatility-weighted portfolios of all 10 currencies in our sample. Volatility is measured and forecasted using the exponentially weighted moving average method, given by

$$(7) \qquad \sigma_{t+1}^2 \quad = \quad \lambda \sigma_t^2 + (1-\lambda)\Delta s_t^2,$$

in which we follow the RiskMetrics approach and set $\lambda=0.94$.[9] The return of the volatility-weighted portfolio is given by

$$(8) \qquad r_t^{\mathrm{vw}} \quad = \quad \left(\sum_i \frac{1}{\sigma_{i,t+1}^2} \right)^{-1} \sum_i \frac{1}{\sigma_{i,t+1}^2} r_{l,t}^i,$$

---

[9]The RiskMetrics variance model was developed by J. P. Morgan to measure portfolio risk.

where i denotes the currency and l denotes the type of return (raw or interest rate adjusted).

Note that the equal-weighted and volatility-weighted portfolios of currencies can contain both long and short positions, depending on the signs of the return forecasts for the 10 currencies in our sample. We also form one pure long–short portfolio based on the forecasts. Specifically, we construct a portfolio that goes long in the currency with the highest forecasted appreciation and short in the currency with the lowest forecasted appreciation (i.e., the largest deprecation). The profitability of all currency investment strategies is assessed by considering the average annualized return and the annualized Sharpe (1966) ratio.[10]

D.   Reality Check Procedure

We evaluate all the exchange rate forecasts out of sample, using real-time macroeconomic data, which rules out look-ahead bias. However, when several model forecasts are compared out of sample, tests of significance also need to be adjusted to take into account the number of forecasts evaluated. For this purpose, we apply the reality check procedure based on White's (2000) test, as implemented by Sullivan, Timmermann, and White (1999) and Sarno and Valente (2009) in the context of time-varying models.

The reality check procedure tests whether a given forecasting model has predictive superiority over a benchmark (random walk without a drift in our case), after accounting for the fact that multiple forecast evaluations are made. The p-values presented in the remainder of the article are calculated with White's (2000) procedure, described in more detail in the Appendix.

E.   Data

We use quarterly data from 1973Q1 to 2014Q1 for Australia, Brazil, Canada, Japan, Mexico, New Zealand, Norway, Russia, Switzerland, the United Kingdom, and the United States from the Organisation for Economic Co-Operation and Development's (OECD) Main Economic Indicators data set. The real-time data consist of the values of the macroeconomic variables as they were originally published by the OECD, before any subsequent revisions. We have four vintages of real-time data per year, starting from 1998Q3, giving a total of 43 vintages. For the sample until 1998Q2, we use revised data.

We use seasonally adjusted values of gross domestic product (GDP), the trade balance, and the commodity price index (CPI) for all countries. For the money supply, when possible, we use seasonally adjusted money supply M1. For the United Kingdom we use M1 plus quasi-money, and for Switzerland we use nonseasonally adjusted money. All variables are expressed in home currencies. Interest rates are 3-month euro deposit rates, retrieved from Datastream. The potential output $\widetilde{y}$ for calculation of the output gap is determined using the Hodrick–Prescott (1997) filter, as in Molodtsova and Papell (2009).

We use two commodity prices. The first is the composite commodity price index of the Commodity Research Bureau. The second is simply the oil price,

---

[10]For simplicity, we ignore transaction costs when calculating strategy returns. Given the high liquidity of foreign exchange markets and the relatively low quarterly rebalancing frequency of the investment strategies, it is unlikely that transaction costs will affect the results much.

measured by the price per barrel of West Texas Intermediate. The U.S. TED spread is computed as the spread between the 3-month interbank rate and the 3-month Treasury yield. Ideally, we would like to have a separate liquidity proxy for each currency, but this is not possible because of data limitations. Instead, we use the TED spread as a proxy for global liquidity. All financial data are retrieved from Datastream.

We estimate each of the 14 models shown in Table 1 in sample for each available OECD data vintage and then create one out-of-sample forecast. For example, the first in-sample period runs from 1973Q1 to 1998Q3 (data permitting) and yields the first out-of-sample forecast for 1998Q4.[11] The second in-sample period runs from 1973Q1 to 1998Q4 and yields an out-of-sample forecast for 1999Q1, and similarly for the following periods. In other words, we create rolling forecasts using an expanding window for each of the 14 models separately. Next, at a given time t when we have at least 20 forecasts per fundamental model, the 14 one-step-ahead forecasts by the individual models are weighted by the BE rule to create one exchange rate forecast for time $t + 1$.

IV.   Main Results

This section describes the performance of the BE mechanism from both statistical and economic perspectives. Specifically, we test whether the BE rule can beat the random walk benchmark in the out-of-sample period in terms of mean prediction errors as well as risk-adjusted returns. A random walk without drift is the most natural benchmark, as it implies a no-change forecast and corresponds to a passive strategy earning a risk-free return (see Qi and Wu (2006)).

A.   Performance of the BE Rule

Table 2 reports the out-of-sample forecasting results when the BE rule is used. We find strong support that combining forecasts using this approach can reduce exchange rate prediction errors substantially. All except two MSPE ratios (third column) in Table 2 are less than 1, indicating that the BE rule delivers better predictions than the random walk model for 8 currencies. On average, the BE rule reduces forecast errors by 24.8%. The last two columns report the CW (2007) and PT (1992) test results: The forecasts made with the BE rule have significantly lower prediction errors than a random walk for 5 out of 10 currencies, according to the reality-check-based p-values for at least one of the tests.

In addition, Table 3 shows that the investment strategies based on the BE rule have a positive and significant mean return for all three currency portfolios. The magnitude of the investment profits is considerable for a market-timing strategy such as ours; for example, the annualized mean return is 3.1% for the volatility-weighted portfolio, with an annualized volatility of 3.2% and a Sharpe ratio of 0.95. All currency portfolios based on the BE scheme in Table 3 have an annualized Sharpe ratio of at least 0.8. The risk-adjusted performance is high compared to other investments in the same period (2003Q3–2014Q2). For example, the Barclays Global Bond Composite Index (a global bond portfolio) achieved

---

[11]The data for this first in-sample period are not real time, whereas the data in each subsequent period are.

TABLE 2

Prediction Performance of the BE Rule

Table 2 reports the out-of-sample forecasting performance of the backward elimination (BE) rule. The column labeled "MSPE Ratio" represents the Clark–West (CW) (2007) adjusted ratio of forecast errors. The columns labeled "CW p-Value" and "PT p-Value" show p-values for the CW statistic and the Pesaran–Timmermann (PT) (1992) statistic, respectively. The p-values test the null hypothesis that the BE forecast is as accurate as a random walk forecast; they are computed using the reality check procedure described in the Appendix. AUD stands for Australian dollar, BRL stands for Brazilian real, CAD stands for Canadian dollar, JPY stands for Japanese yen, MXN stands for Mexican peso, NZD stands for New Zealand dollar, NOK stands for Norwegian krone, RUB stands for Russian ruble, CHF stands for Swiss franc, and GBP stands for U.K. pound sterling.

Currency	MSPE Ratio	CW p-Value	PT p-Value
AUD	0.904	0.240	0.189
BRL	0.261	0.023	0.009
CAD	0.901	0.290	0.117
JPY	0.512	0.016	0.271
MXN	1.166	0.932	0.933
NZD	0.786	0.042	0.185
NOK	1.064	0.660	0.386
RUB	0.211	0.046	0.072
CHF	0.862	0.276	0.191
GBP	0.848	0.190	0.031
BE Avg. Performance	24.8%		

TABLE 3

Investment Strategies Based on the BE Rule Forecasts

Table 3 reports the out-of-sample returns of an investment strategy that goes long (short) in the currency that the backward elimination (BE) model selection rule forecasts to appreciate (depreciate). The rows labeled "P-EW," "P-VW," and "P-LS" represent the performance of the equal-weighted, volatility-weighted, and long–short portfolios, respectively. The annualized average return ("Avg."), annualized return standard deviation ("Std. Dev."), and Sharpe ratio ("Sharpe") of the investment strategies are shown. The returns in Panel A do not take deposit interest rates into account and represent the exchange rate returns. The returns in Panel B include deposit interest rates (long and short) in addition to the exchange rate returns. The p-values reported are used to evaluate the null hypothesis that the BE-based strategy performs better than the random walk strategy and are computed using the reality check procedure described in the Appendix.

	Panel A. FX Returns				Panel B. FX + Interest Returns			
Portfolio	Avg.	Std. Dev.	Sharpe	p-Value	Avg.	Std. Dev.	Sharpe	p-Value
P-EW	0.028	0.034	0.809	0.008	0.029	0.035	0.811	0.014
P-VW	0.031	0.032	0.955	0.004	0.027	0.032	0.841	0.011
P-LS	0.097	0.123	0.789	0.010	0.105	0.122	0.863	0.005

a Sharpe ratio of 0.76 in this period and the MSCI (formerly Morgan Stanley Capital International) World index (a global equity portfolio) achieved a Sharpe ratio of 0.45.

B.   Alternative Model Averaging Criteria

If investors attach excessive weights to a subset of fundamentals, the BE scheme should perform better than other model averaging criteria that do not eliminate irrelevant fundamentals.

First, as a simple benchmark for comparison, we take an equal-weighted average of the 14 fundamental-based forecasts. The second weighting scheme is the smoothed Bayesian information criterion (BIC) model averaging rule, which is closely related to the Bayesian model averaging method applied for exchange rate forecasting by Wright (2008) and Della Corte et al. (2009).[12] The investor is assumed to have a prior distribution about which model is the true data-generating

[12]Given diffuse priors and equal model prior probabilities, BIC rule weights and Bayesian model averaging weights are approximately the same.

process and computes posterior probabilities after additional information becomes available. An often-applied estimator of the marginal likelihood is the exponent of the BIC. Following Buckland, Burnham, and Augustin (1997), we define the BIC combination forecast of the exchange rate as

$$
(9a) \qquad E_t^{\text{BIC}}(\Delta s_{t+1}) \quad = \quad \sum_{i=1}^{N} w_{i,t} E_t^i(\Delta s_{t+1}),
$$

$$
(9b) \qquad w_{i,t}^{\text{BIC}} \quad = \quad \frac{\exp(-k\,\Delta\text{BIC}_{i,t-1})}{\sum_{j=1}^{N} \exp(-k\,\Delta\text{BIC}_{j,t-1})},
$$

in which $\Delta\text{BIC}_{i,t} = \text{BIC}_{i,t} - \min_j \text{BIC}_{j,t}$ gives the posterior odds that model i is the best predictive model.[13] The parameter $k \geq 0$ controls the relative importance of the best-fitting model. For $k=0$, the model weights are equal, and the higher k is, the heavier the weight attributed to the best performing model (the model with the lowest BIC). As a starting point, we use $k=10$, which implies a smooth adjustment of the model weights toward the model with the best recent in-sample fit. More precisely, with $k=10$, the best performing fundamental receives a weight of up to 25% in the averaged forecast.

As a third forecast combination rule, we apply an approach that assigns weights to the models based on their recent out-of-sample forecasting performance, also known as cross-validation. Della Corte and Tsiakas (2012) apply this method for exchange rate forecasting and find that it can improve performance, beating a random walk. In general, the use of recent out-of-sample information can make forecasts more robust to structural breaks and model misspecifications (Newbold and Harvey (2002)). To be specific, the MSPE rule weighs the 14 forecasts based on each model's MSPE during the last τ periods:

$$
(10a) \qquad E_t^{\text{MSPE}}(\Delta s_{t+1}) \quad = \quad \sum_{i=1}^{N} w_{i,t} E_t^i(\Delta s_{t+1}),
$$

$$
(10b) \qquad w_{i,t}^{\text{MSPE}} \quad = \quad \frac{\left(\sum_{j=1}^{\tau} \text{MSPE}_{t-j}^i\right)^{-1}}{\sum_{j=1}^{N}\left(\sum_{j=1}^{\tau} \text{MSPE}_{t-j}^j\right)^{-1}},
$$

in which τ is the lookback period over which forecasting performance is evaluated. MSPE_t^i is determined using model i's one-step-ahead out-of-sample forecasts during the last τ periods, compared to the actual exchange rate returns. In the empirical analyses, we use $\tau = 5$; the results are not sensitive to this choice.[14]

We note that the three alternative rules described above give all available fundamental models nonzero weights, contrary to the first hypothesis in Section II.

Table 4 reports the forecasting performance of the alternative model averaging criteria: The first column shows the forecasted currency, the second column reports the MSPE ratios, and the final two columns show the CW (2007) and PT (1992) statistics. The table is divided into three panels. Panel A reports the results

---

[13]Alternatively, the Aikaike information criterion (AIC) could be used. Results are similar for BIC and AIC but the AIC results are not reported here for brevity.

[14]We also applied a version of MSPE forecasting based on the forecast rank of each model. Results are qualitatively similar to the MSPE results and are available from the authors.

TABLE 4

Prediction Performance of Other Model Averaging Rules

Table 4 reports the out-of sample prediction performance of three alternative model averaging rules, using equal weights (Panel A), Bayesian information criterion (BIC) weights (Panel B), and mean squared prediction error (MSPE)-based weights (Panel C). The MSPE ratio (shown in columns labeled "MSPE") represents the Clark–West (CW) (2007)-adjusted ratio of forecast errors relative to a random walk model. The columns labeled "CW p" and "PT p" show p-values for the CW test and the Pesaran–Timmermann (PT) (1992) test. AUD stands for Australian dollar, BRL stands for Brazilian real, CAD stands for Canadian dollar, JPY stands for Japanese yen, MXN stands for Mexican peso, NZD stands for New Zealand dollar, NOK stands for Norwegian krone, RUB stands for Russian ruble, CHF stands for Swiss franc, and GBP stands for U.K. pound sterling.

	Panel A. Equal-Weighted			Panel B. BIC Weights			Panel C. MSPE Weights		
Currency	MSPE	CW p	PT p	MSPE	CW p	PT p	MSPE	CW p	PT p
AUD	1.019	0.969	0.757	1.038	0.962	0.500	1.025	0.980	0.757
BRL	0.812	0.613	0.938	0.910	0.376	0.870	0.576	0.272	0.398
CAD	1.012	0.845	0.895	1.037	0.672	0.713	1.011	0.842	0.836
JPY	0.995	0.389	0.445	1.012	0.373	0.389	1.014	0.514	0.287
MXN	1.084	0.356	0.662	1.338	0.945	0.801	1.106	0.477	0.662
NZD	0.969	0.030	0.018	1.003	0.120	0.045	0.961	0.041	0.035
NOK	0.981	0.527	0.713	1.005	0.505	0.100	0.983	0.576	0.500
RUB	1.250	0.576	0.662	1.130	0.423	0.921	1.276	0.721	0.199
CHF	0.996	0.432	0.555	1.011	0.382	0.389	1.001	0.439	0.338
GBP	0.980	0.116	0.338	1.001	0.158	0.500	0.977	0.118	0.445
Avg.	−1.0%			−4.9%			0.7%		

for the equal-weighted forecast combination, Panel B shows the findings for the Bayesian model averaging rule, and Panel C reports the MSPE weighting rule based on recent forecast errors.

The results in Table 4 show that none of the alternative forecast combination rules meaningfully outperforms a random walk without drift. Although on average the forecast errors are lower than the random walk benchmark (ratio < 1) by 0.7% when using MSPE weights, the performance difference is significant only for the NZD. The outperformance of the MSPE-based forecasts is 0.7% on average for the 10 currencies, close to 0, and substantially lower than the 24.8% improvement achieved by the BE rule in Table 2. We also note that the forecasts based on the Bayesian model weighting rule perform substantially worse than a random walk, by 4.9% on average.

Table 5 displays the returns of investment strategies based on the forecasts made by the three alternative model selection criteria. In addition, we report the results of an investment strategy based on forecasts from a random walk model with drift.[15] For each of the strategies, we construct an equal-weighted portfolio of currencies, a volatility-weighted portfolio, and a long–short portfolio (one currency long, one short). Panel A of Table 5 reports raw currency returns without interest, and Panel B reports returns that take interest rate differentials into account.

The results in Table 5 show that none of the three alternative model averaging methods has a mean return that is positive and significant. Some of the Sharpe ratios are positive, but they are always less than half the size of those reported for the BE rule in Table 3.[16] The results demonstrate that alternative model selection

[15]We regress the exchange rate returns on a constant, and the sign of the constant's coefficient indicates a buy or a sell signal.

[16]Please note that the average returns in Panel B of Table 5 (including interest) are frequently more negative than in Panel A (without interest). There are two reasons: Most fundamental-based forecasts

TABLE 5

Investment Returns of Other Model Averaging Rules and Benchmarks

Table 5 reports the out-of-sample returns of an investment strategy that goes long (short) in the currency that the model selection rule forecasts to appreciate (depreciate). The annualized average return ("Avg."), annualized return standard deviation ("Std. Dev."), and Sharpe ratio ("Sharpe") of the investment strategies are shown. The three rows labeled "Equal" display returns of an investment strategy based on forecasts with equal model weights; "BIC" denotes Bayesian information criterion weights, and "MSPE" denotes mean squared prediction error (MSPE)-based weights. The three rows labeled "RWD" show the investment returns for a strategy using currency forecasts from a random walk with drift model as an alternative benchmark. The returns in Panel A do not take deposit interest rates into account and represent the exchange rate returns. The returns in Panel B include deposit interest rates (long and short) in addition to the exchange rate returns.

Weights	Portfolios	Panel A. FX Returns				Panel B. FX + Interest Returns			
		Avg.	Std. Dev.	Sharpe	p-Value	Avg.	Std. Dev.	Sharpe	p-Value
Equal	P-EW	−0.001	0.054	−0.010	0.974	−0.017	0.054	−0.313	0.316
	P-VW	0.003	0.051	0.052	0.865	−0.013	0.050	−0.256	0.412
	P-LS	−0.042	0.145	−0.289	0.343	−0.051	0.148	−0.344	0.259
BIC	P-EW	0.016	0.044	0.376	0.142	−0.008	0.032	−0.254	0.421
	P-VW	0.003	0.035	0.082	0.783	−0.008	0.032	−0.244	0.440
	P-LS	0.010	0.150	0.067	0.827	0.010	0.150	0.067	0.827
MSPE	P-EW	0.010	0.054	0.181	0.558	−0.003	0.053	−0.050	0.873
	P-VW	0.011	0.052	0.204	0.508	−0.002	0.050	−0.031	0.920
	P-LS	0.008	0.103	0.075	0.805	0.008	0.103	0.075	0.805
RWD	P-EW	0.001	0.036	0.030	0.922	−0.009	0.037	−0.250	0.424
	P-VW	0.005	0.036	0.148	0.627	−0.004	0.036	−0.121	0.699
	P-LS	−0.030	0.173	−0.174	0.568	−0.113	0.179	−0.630	0.039

rules that give nonzero weights to all 14 fundamental model forecasts do not perform well, as suggested by our hypothesis that exchange rates are driven by a small subset of fundamentals.

C.   Comparison with Carry Trades

Carry trades generated substantial investment profits until the financial crisis of 2008–2009 (Menkhoff, Sarno, Schmeling, and Schrimpf (2012a)). It is therefore relevant to compare the performance of the BE-based investment strategy with carry-trade returns, shown in Table 6. Taking the perspective of a U.S. investor, we compute the returns of long–short portfolios that go long in the currency with the highest 3-month interbank rate and short in the currency with the lowest 3-month interbank rate.[17] First, we construct two portfolios with one and three currencies per short and long leg. Second, we calculate the returns of an investment in one leg only (long and short separately), again with one and three currencies. The portfolio weights of the currencies are always equal, and rebalancing takes place every quarter.

Table 6 shows that the Sharpe ratios of carry-trade strategies are lower than the Sharpe ratio of our investment strategy based on BE rule forecasts, especially for raw returns. Although the Sharpe ratios after including interest rates are reasonably high, in nearly all cases they are statistically insignificant (the one exception is a portfolio that always goes long in the highest yielding currency). These results clearly show that the fundamental-based predictions of the BE scheme

---

ignore interest rate differentials, and the investment strategies have a tendency to short currencies with high interest rates, such as the RUB and the BRL.

[17]We effectively use long and short positions in a 3-month forward contract on a foreign currency versus the USD.

TABLE 6

Investment Returns of Carry-Trade Strategies

Table 6 reports the out-of-sample returns of carry-trade strategies that go long (short) in currencies yielding high (low) interest rates. The annualized average return ("Avg."), annualized return standard deviation ("Std. Dev."), and Sharpe ratio ("Sharpe") of the investment strategies are shown. The returns in Panel A do not take deposit interest rates into account and represent the exchange rate returns. The returns in Panel B include deposit interest rates (long and short) in addition to the exchange rate returns.

Carry-Trade Strategies	Panel A. FX Returns				Panel B. FX + Interest Returns			
	Avg.	Std. Dev.	Sharpe	p-Value	Avg.	Std. Dev.	Sharpe	p-Value
Long − short (1 pair)	0.010	0.174	0.059	0.847	0.088	0.177	0.494	0.105
Long (1 long)	0.026	0.147	0.180	0.556	0.097	0.149	0.652	0.033
Short (1 short)	0.016	0.104	0.155	0.611	0.009	0.100	0.093	0.761
Long − short (3 pairs)	−0.020	0.105	−0.186	0.543	0.037	0.109	0.342	0.262
Long (3 long)	−0.001	0.112	−0.009	0.975	0.051	0.110	0.465	0.127
Short (3 short)	0.018	0.072	0.256	0.402	0.014	0.069	0.202	0.507

exploit additional information that goes beyond simple carry-trade strategies. In the next section, we analyze the features that make the BE rule so successful.

Finally, we note that Sharpe ratios for carry-trade strategies reported in the literature are usually slightly higher. We believe that the main reason for this difference lies in the lower, quarterly frequency of our data. Bakshi and Panayotov (2013) show that only dynamically rebalanced carry-trade portfolios generate statistically significant returns when using monthly data. In contrast, our quarterly portfolio rebalancing might be too infrequent to fully exploit all potential carry-trade profits. Our main aim is to compare performance with the BE rule, which is based on quarterly macroeconomic data.

D.   What Drives Out-of-Sample Performance?

The results presented so far clearly indicate that the BE rule performs best. Here we explore in greater detail why this is the case and whether the reasons for its good performance are consistent with the hypotheses drawn from theory.

The first hypothesis states that exchange rate returns are driven by a small subset of fundamentals. The second hypothesis indicates that the relation between exchange rates and fundamentals is time varying. We first verify the consistency of our main results with these hypotheses by studying the nature of the exchange rate forecasting models selected by the BE mechanism and the number of predictor changes through time. Next, we carry out additional forecasting tests to validate the hypotheses.

1.   Does Time Variation in the Selected Fundamentals Drive the Results?

Table 7 reports how frequently each model is selected by the BE rule, separately for each currency. Clearly, for each currency, a different set of fundamentals matters. For example, models selected for forecasting the JPY often include interest rate differentials (UIP) and global market liquidity (TED spread), both of which are related to carry trades. The commodity index and oil prices are frequently selected to predict the currency of commodity exporters, such as Australia, New Zealand, Brazil, and Mexico. For the CHF, a safe-haven currency that tends to appreciate in times of turmoil, the volatility index matters most, whereas other fundamentals have a minor role.

TABLE 7

Model Selection Frequencies of the BE Rule

Table 7 reports how often 14 individual fundamental models were selected by the backward model selection rule. The model selection rule was applied in 43 consecutive quarters from 2003Q3 to 2014Q1. "Model" indicates the individual macroeconomic and financial models used to predict the exchange rate (see Table 1). "AR1" is the momentum model. "Cons growth" is the risk-sharing model. "Monetary model 1" and "Monetary model 2" stand for the monetary model using levels (1) and first differences (2) of money supply and real output. "NFA" stands for the capital flows model using net foreign assets (NFA). "PPP" stands for the purchasing power parity model. "Taylor rule 1" and "Taylor rule 2" stand for the Taylor rule model (1) and its asymmetric version (2). "UIP" stands for the uncovered interest parity model. "FX volatility" stands for the exchange rate volatility model. The row "Avg. no. of models" displays the average number of fundamental models retained (and used) in the final forecasting equation of the backward regression rule. The row "Avg. no. of changes" shows the average number of changes per quarter in the set of fundamental models selected to predict a currency. AUD stands for Australian dollar, BRL stands for Brazilian real, CAD stands for Canadian dollar, JPY stands for Japanese yen, MXN stands for Mexican peso, NZD stands for New Zealand dollar, NOK stands for Norwegian krone, RUB stands for Russian ruble, CHF stands for Swiss franc, and GBP stands for U.K. pound sterling.

Model	AUD	BRL	CAD	JPY	MXN	NZD	NOK	RUB	CHF	GBP
AR1	0%	7%	12%	0%	23%	9%	5%	21%	5%	0%
Cons growth	0%	7%	0%	21%	0%	0%	0%	12%	0%	5%
Monetary model 1	2%	14%	2%	42%	12%	0%	5%	19%	0%	5%
Monetary model 2	88%	5%	0%	0%	0%	37%	0%	14%	2%	12%
NFA	2%	37%	23%	0%	14%	40%	0%	7%	0%	9%
PPP	7%	12%	0%	16%	26%	72%	2%	23%	0%	5%
Taylor rule 1	9%	2%	0%	19%	0%	14%	2%	9%	0%	56%
Taylor rule 2	33%	9%	0%	19%	0%	65%	0%	12%	0%	56%
Trade balance	40%	2%	2%	2%	7%	12%	0%	7%	0%	12%
UIP	23%	23%	30%	47%	33%	42%	2%	28%	5%	19%
FX volatility	30%	5%	0%	0%	2%	5%	0%	16%	19%	35%
Commodity index	33%	19%	2%	21%	26%	35%	7%	23%	2%	28%
Oil price	0%	21%	0%	0%	21%	30%	0%	9%	0%	0%
TED spread	9%	7%	0%	60%	2%	19%	2%	56%	0%	0%
Avg. no. of models	2.77	1.70	0.72	2.47	1.65	3.79	0.26	2.56	0.33	2.40
Avg. no. of changes	0.44	0.42	0.30	0.47	0.54	0.81	0.19	0.44	0.14	0.44

The last row of Table 7 shows how often changes occur in the set of fundamentals selected to predict a currency. For example, for the CAD, the set of fundamentals changes 13 times in 43 quarters, or roughly one-third of the time (0.30). Overall, the average number of predictor changes is 0.4 per quarter, or 1.6 times per year. Thus, although there is some persistence in the selected predictors, they change frequently over time, in line with the second hypothesis.

2.   Do Excessive Weights Drive Performance?

We now focus on our first hypothesis, namely, excessive weights on a small subset of fundamentals. The second-to-last row in Table 7 shows that, on average, only a small subset of 2 out of 14 fundamentals is selected to predict the currency.

If investors focus excessively on a small subset of fundamentals, then not only the BE scheme but also other model selection criteria that exploit this fact should deliver good forecasting performance. We test this hypothesis by shifting the value of k in the smoothed BIC model averaging rule described by equation (9b). Higher values of k imply heavier weights for the most relevant fundamentals (the best-fitting models in sample) and lower weights for other exchange rate predictors.

We note that the performance of the original BIC rule in Table 4 was poor (-4.9%). In results not reported to save space (but available from the authors), we show that the original BIC rule performs poorly because it uses BIC values reflecting in-sample model fit, not out-of-sample prediction performance, and model weights that are bounded between 0 and 1. We therefore make two changes: The BIC values of the 14 models are based on regressions of actual exchange rates on recent out-of-sample model forecasts, similar to the BE and MSPE rules, and we remove the non-negativity constraint on the BIC weights.

Figure 1 shows the average MSPE ratios for 10 currencies of the adjusted BIC rule for $k = 0, 1, 10, 100, 1,000$. We observe that the adjusted BIC rule's prediction performance gradually improves as k increases. This is in line with the first hypothesis, predicting that investors focus on a small subset of fundamentals.

FIGURE 1

Performance of BIC Averaging Rule for Different Values of k

Figure 1 shows the mean squared prediction error (MSPE) ratio of the adjusted Bayesian information criterion (BIC) rule relative to the random walk benchmark on the y-axis, for different values of parameter $k = 0, 1, 10, 100, 1,000$ on the x-axis. A decrease in the value of the MSPE ratio implies improved performance of the BIC rule relative to the random walk.
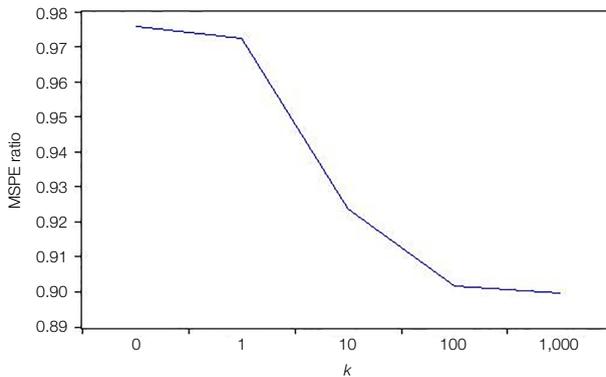
V.   Robustness Checks

The BE procedure involves several empirical choices concerning the base currency, parameter values, sample length, and order of the fundamentals' inclusion. In this section, we demonstrate with several robustness tests that our findings do not depend on most of these choices.

A.   Base Currency

To eliminate the influence of the base currency (USD) and to compare the BE rule's performance with often-implemented carry trades, we evaluate its statistical performance using fixed currency pairs with high and low interest rates: AUD/JPY, AUD/CHF, NZD/JPY, NZD/CHF, NOK/JPY, and NOK/CHF. The results reported in Table 8 show that the BE procedure is also successful when applied to the usual carry-trade currency pairs. The BE rule is significantly better than a random walk according to at least one statistic for three out of six currency pairs, and forecasting errors are reduced by 26.7% on average compared to a random walk. We note that two adjusted MSPE ratios in Table 8 appear extreme (AUD/JPY and NOK/CHF), but closer inspection reveals that this is mainly the result of high currency volatility during the 2008–2009 crisis. Excluding the 2008–2009 crisis period, the MSPE ratios for AUD/JPY and NOK/CHF are 0.982 and 0.114, respectively. Overall, the results indicate that the BE rule's good forecasting performance is not limited to the original set of exchange rates with the USD as the base currency.

TABLE 8

BE Performance for Carry-Trade Currencies

Table 8 presents the out-of-sample forecasting performance of the backward elimination (BE) rule. The column labeled "MSPE Ratio" represents the Clark–West (CW) (2007)-adjusted ratio of forecast errors. The columns labeled "CW *p*-value" and "PT *p*-value" show *p*-values for the CW test statistic and the Pesaran–Timmermann (PT) (1992) test statistic. The *p*-values test the null hypothesis that the BE forecast is as accurate as a random walk forecast; they are computed using the reality check procedure described in the Appendix. AUD stands for Australian dollar, JPY stands for Japanese yen, CHF stands for Swiss franc, NZD stands for New Zealand dollar, and NOK stands for Norwegian krone.

Currency Pair	MSPE Ratio	CW *p*-Value	PT *p*-Value
AUD/JPY	2.589	0.858	0.757
AUD/CHF	0.761	0.170	0.338
NZD/JPY	0.692	0.156	0.062
NZD/CHF	0.545	0.003	0.002
NOK/JPY	0.046	0.001	0.002
NOK/CHF	−0.232	0.056	0.000
BE Avg. Performance	26.7%		

B. Elimination Criterion

The BE procedure drops insignificant fundamental models using a *p*-value of 5% as the threshold for significance. We check the robustness of this choice by changing the significance levels to 1% and 10%. The results are presented in the second and third columns of Table 9. We find that the BE rule outperforms the random walk by 26.4% and 19.2% on average, confirming our previous findings.

By definition, the BE process first includes all regressors and subsequently eliminates those that are insignificant, helping reduce omitted-variable bias. However, it may lead to overparametrization. Therefore, we also test an alternative approach based on forward inclusion: Start with no regressors and then include significant variables one by one, starting from the most significant one. The results for the forward inclusion rule are shown in the fourth column of Table 9. This model selection procedure improves over a random walk by an average of only 2.9%.

TABLE 9

Robustness Checks

Table 9 reports the mean squared prediction error (MSPE) ratios for several robustness checks. We change the significance level for the backward elimination step from 5% to 1% ("$\alpha = 1\%$") and 10% ("$\alpha = 10\%$"); we change the model selection method from backward elimination to forward inclusion of significant variables ("Forward"); instead of backward elimination, we maximize R^2 to select the best two fundamental models ("Max R^2"); instead of an expanding estimation window, we use a 30-quarter rolling window ("Rolling"); and we split the full sample period into two halves. The MSPE ratio is the Clark–West (2007)-adjusted ratio of forecast errors. * and ** indicate significant predictability based on either the CW statistic or the Pesaran–Timmermann (PT) (1992) statistic at the 5% and 1% levels, respectively. AUD stands for Australian dollar, BRL stands for Brazilian real, CAD stands for Canadian dollar, JPY stands for Japanese yen, MXN stands for Mexican peso, NZD stands for New Zealand dollar, NOK stands for Norwegian krone, RUB stands for Russian ruble, CHF stands for Swiss franc, and GBP stands for U.K. pound sterling.

Currency	$\alpha = 1\%$	$\alpha = 10\%$	Forward	Max R^2	Rolling	1st Half	2nd Half
AUD	0.923	0.983	1.105	1.195	2.264	0.852*	0.960
BRL	0.131**	0.225*	1.257	1.107	0.261**	—	—
CAD	0.942	1.245	0.975	1.083	1.238	0.942	0.836
JPY	0.640*	0.384*	0.691	0.726*	0.941	0.215*	0.922
MXN	1.048	1.235	1.035	1.017	0.636	1.105	1.244
NZD	0.914	0.579	0.968	1.020	0.238*	0.984	0.582*
NOK	1.016	1.280	1.016	1.077	0.921	1.097	1.008
RUB	0.032**	0.451**	0.984	0.888	−0.735**	—	—
CHF	1.036	0.915	0.829	0.889	1.103	0.688	1.099
GBP	0.681*	0.778*	0.851	0.848	0.907	1.022	0.525*
Avg.	26.4%	19.2%	2.9%	1.5%	22.2%	13.7%	10.3%

As an alternative model selection criterion, we select the combination of two fundamental models that maximizes the R^2. Because R^2 always increases when the number of regressors increases, we need to fix the maximum number of fundamental-based models in the regression equation in advance. We choose two, as this is the average number of models included by the BE procedure (see Table 7). The fifth column of Table 9 shows that this method produces only a minor improvement of 1.5% over a random walk.

We thus find that both forward inclusion and restricting the number of fundamental models leads to relatively poor results compared to the BE rule. A potential explanation is that both alternative model selection rules are prone to omitted-variable bias and thus may fail to take all available relevant information for exchange rate prediction into account.

C.   Rolling Window

The second-stage regression in step 3 of the BE rule (see Section III) is based on an expanding window, starting with 20 observations in the first cycle and expanding to 62 observations in the last cycle. We run an alternative specification with a rolling window of 30 observations.[18] The findings, shown in the sixth column of Table 9, indicate that the BE scheme with a rolling window performs well, 22.2% better than the random walk benchmark on average. Hence, the results are robust to changing from an expanding to a rolling window.

D.   Sample Choice

We split the out-of-sample period into two halves to study the robustness of the results to the sample period. We exclude the BRL and RUB in these two tests because limited historical data for these currencies prevent splitting the sample. The seventh and eighth columns of Table 9 display the findings of this robustness check. The BE rule beats the random walk benchmark in both subsamples. In the first-half subsample, the BE scheme outperforms a random walk by 13.7% on average, and in the second-half subsample by 10.3%. The baseline performance of the BE rule in this limited group of eight currencies (excluding BRL and RUB) is 12.0% in the full sample, close to the performance in either half sample. Hence, our results do not depend strongly on the sample period.

VI.   Conclusion

We study the time-varying relation between the exchange rate and economic fundamentals. Our empirical analysis builds on recently developed theoretical models of the exchange rate in which investors have imperfect information. These models suggest that the weights attached to the macroeconomic fundamentals in the exchange rate process are time varying and often excessive, focusing on only a small subset of fundamentals.

---

[18]We choose not to use 20 observations because of the low number of observations (the BE process starts with 14 regressors). We choose to start the original specification with 20 observations to maximize the number of out-of-sample observations. Starting with 30 observations and expanding the window yield similar results.

We test the theoretical framework via a dynamic model selection rule, based on backward elimination of fundamentals. This approach encompasses the two main predictions regarding the weights. Using real-time economic data for 10 major currencies in an out-of-sample forecasting test, we find that the BE scheme outperforms the random walk benchmark by 24.9% on average. In addition, investment strategies based on these forecasts have positive and significant average returns and Sharpe ratios of 0.8 and higher. We find that the forecasting performance of the BE rule is driven by time-varying and excessive weights on the fundamentals, in line with the recent theoretical models of exchange rate determination.

Appendix. Reality Check Procedure

In this Appendix, we describe the reality check procedure implemented to correct the p-values used in the out-of-sample tests, following White (2000) and Sarno and Valente (2009). Let $\overline{f}$ denote the $N \times 1$ vector of performance statistics:

$$\overline{f} \quad = \quad R^{-1} \sum_{t=L}^{T} \widehat{f}_{t+1},$$

where $N = 14$ is the number of fundamental-based models used to make the forecast, R is the number of prediction periods, and $\widehat{f}_{t+1}$ is the performance measure relative to the benchmark. In our application, we use four performance measures: the CW (2002)-adjusted MSPE, the PT (1992) statistic, raw investment strategy returns, and interest-rate-adjusted returns. Our objective is to test the null hypothesis that the model selected by BE has no predictive superiority over the random walk benchmark forecast. For the two statistical performance metrics, we test the null hypothesis:

$$H_0^1 : E\left(f^{\mathrm{BE}}\right) \quad \geq \quad 0,$$

where f^{BE} is the performance measure of the BE rule. Rejection of H_0^1 implies that the BE-rule-based forecast performs better than a random walk (no change forecast).

For the Sharpe ratios of the investment strategies, we test the null hypothesis:

$$H_0^2 : E\left(f^{\mathrm{BE}}\right) \quad \leq \quad 0.$$

Rejection of H_0^2 implies that the BE-based investment strategy performs better than the random walk strategy. White (2000) proposes the reality check procedure to test both null hypotheses. We implement a version of the routine used by Sarno and Valente (2009) to compute the p-values for each of our four performance measures.

i) We assume that the exchange rate follows a driftless random walk, and we simulate $j = 1, \ldots, 1,000$ return series of the 10 currencies, with the length equal to the empirical number of observations. The return variance of each currency and the correlation between the currency returns are replicated exactly by taking the Cholesky decomposition of the historical variance–covariance matrix of the exchange rate returns and multiplying it by the matrix of $R \times 10$ random draws. In addition, to determine the interest-adjusted Sharpe ratios of investment strategies, we simulate forward premia as well. The characteristics of the simulated forward premia are matched to those of the empirically observed premia.

ii) For each of the simulated series, using the actual fundamental data, we estimate all models and generate out-of-sample forecasts based on the BE rule.

iii) Using these forecasts, we compute each of four performance measures, $\overline{f}^* = R^{-1}\sum_{t=L}^{T}\widehat{f}^*_{t+1}$, where asterisks denote the fact that these measures are based on the simulated data.

iv) We construct the following statistics: $V_j^{1,*} = \sqrt{R}(\overline{f}^{*\text{BE}})_j$ for statistical metrics and $V_j^{2,*} = \sqrt{R}(\overline{f}^{*\text{BE}})_j$ for Sharpe ratios, where $j = 1,\ldots,1{,}000$ denote replications of test statistics. We know now the distribution of the performance measures under the null.

v) To obtain the p-values, we compute the percentage of times the performance measure in $V_j^{1,*}$ $\left(V_j^{2,*}\right)$ is smaller (larger) than the empirical value of the corresponding performance measure.

References

Bacchetta, P., and E. van Wincoop. "A Scapegoat Model of Exchange Rate Determination." *American Economic Review, Papers and Proceedings*, 94 (2004), 114–118.

Bacchetta, P., and E. van Wincoop. "Can Information Heterogeneity Explain the Exchange Rate Determination Puzzle?" *American Economic Review*, 96 (2006), 552–576.

Bacchetta, P., and E. van Wincoop. "On the Unstable Relationship between Exchange Rates and Macroeconomic Fundamentals." *Journal of International Economics*, 91 (2013), 18–26.

Bakshi, B., and G. Panayotov. "Predictability of Currency Carry Trades and Asset Pricing Implications." *Journal of Financial Economics*, 110 (2013), 139–163.

Benigno, G., and C. Thoenissen. "Consumption and Real Exchange Rates with Incomplete Markets and Non-Traded Goods." *Journal of International Money and Finance*, 27 (2008), 926–948.

Buckland, S. T.; K. P. Burnham; and N. H. Augustin. "Model Selection: An Integral Part of Inference." *Biometrics*, 53 (1997), 603–618.

Chen, Y., and K. S. Rogoff. "Commodity Currencies." *Journal of International Economics*, 60 (2003), 133–160.

Chen, Y.; K. S. Rogoff; and B. Rossi. "Can Exchange Rates Forecast Commodity Prices?" *Quarterly Journal of Economics*, 125 (2010), 1145–1194.

Cheung, Y.-W., and M. D. Chinn. "Currency Traders and Exchange Rate Dynamics: A Survey of the U.S. Market." *Journal of International Money and Finance*, 20 (2001), 439–471.

Cheung, Y.-W.; M. D. Chinn; and A. G. Pascual. "Empirical Exchange Rate Models of the Nineties: Are Any Fit to Survive?" *Journal of International Money and Finance*, 24 (2005), 1150–1175.

Clark, T., and K. D. West. "Approximately Normal Tests for Equal Predictive Accuracy in Nested Models." *Journal of Econometrics*, 138 (2007), 291–311.

Della Corte, P.; L. Sarno; and I. Tsiakas. "An Economic Evaluation of Empirical Exchange Rate Models." *Review of Financial Studies*, 22 (2009), 3491–3530.

Della Corte, P., and I. Tsiakas. "Statistical and Economic Methods for Evaluating Exchange Rate Predictability." In *Handbook of Exchange Rates*, J. James, I. W. Marsh, and L. Sarno, eds. Hoboken, NJ: Wiley (2012), 221–263.

Ferraro, D.; K. S. Rogoff; and B. Rossi. "Can Oil Prices Forecast Exchange Rates? An Empirical Analysis of the Relationship between Commodity Prices and Exchange Rates." *Journal of International Money and Finance*, 54 (2015), 116–141.

Fratzscher, M.; D. Rime; L. Sarno; and G. Zinna. "The Scapegoat Theory of Exchange Rates: The First Tests." *Journal of Monetary Economics*, 70 (2015), 1–21.

Garratt, A., and E. Mise. "Forecasting Exchange Rates Using Panel Model and Model Averaging." *Economic Modelling*, 37 (2014), 32–40.

Granger, C. W. J., and P. Newbold. *Forecasting Economic Time Series*, 2nd ed. New York, NY: Academic Press (1986).

Hendry, D. *Dynamic Econometrics*. Oxford, UK: Oxford University Press (1995).

Hodrick, R. J., and E. C. Prescott. "Postwar U.S. Business Cycles: An Empirical Investigation." *Journal of Money, Credit and Banking*, 29 (1997), 1–16.

Lai, T. L., and H. Xing. *Statistical Models and Methods in Financial Markets*. New York, NY: Springer (2008).

Mark, N. C. "Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability." *American Economic Review*, 85 (1995), 201–218.

Markiewicz, A. "Model Uncertainty and Exchange Rate Volatility." *International Economic Review*, 53 (2012), 815–844.

Meese, R. A., and K. S. Rogoff. "Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?" *Journal of International Economics*, 14 (1983), 345–373.

Menkhoff, L.; L. Sarno; M. Schmeling; and A. Schrimpf. "Carry Trades and Global Foreign Exchange Volatility." *Journal of Finance*, 67 (2012a), 681–718.

Menkhoff, L.; L. Sarno; M. Schmeling; and A. Schrimpf. "Currency Momentum Strategies." *Journal of Financial Economics*, 106 (2012b), 660–684.

Molodtsova, T., and D. H. Papell. "Out-of-Sample Exchange Rate Predictability with Taylor Rule Fundamentals." *Journal of International Economics*, 77 (2009), 167–180.

Morales-Arias, L., and G. V. Moura. "Adaptive Forecasting of Exchange Rates with Panel Data." *International Journal of Forecasting*, 29 (2013), 493–509.

Newbold, P., and I. H. Harvey. "Forecast Combination and Encompassing." In *A Companion to Economic Forecasting*, M. P. Clements and D. F. Hendry, eds. New York, NY: Blackwell (2002), 268–283.

Pesaran, M. H., and A. Timmermann. "A Simple Nonparametric Test of Predictive Performance." *Journal of Business and Economic Statistics*, 10 (1992), 461–465.

Pojarliev, M., and R. M. Levich. "Are All Currency Managers Equal?" *Journal of Portfolio Management*, 64 (2010), 42–53.

Qi, M., and Y. Wu. "Technical Trading-Rule Profitability, Data Snooping, and Reality Check: Evidence from the Foreign Exchange Market." *Journal of Money, Credit and Banking*, 30 (2006), 2135–2158.

Rossi, B. "Are Exchange Rates Really Random Walks? Some Evidence Robust to Parameter Instability." *Macroeconomic Dynamics*, 10 (2006), 20–38.

Rossi, B. "Exchange Rate Predictability." *Journal of Economic Literature*, 51 (2013), 1063–1119.

Sarkissian, S. "Incomplete Consumption Risk Sharing and Currency Risk Premiums." *Review of Financial Studies*, 16 (2003), 983–1005.

Sarno, L., and G. Valente. "Exchange Rates and Fundamentals: Footloose or Evolving Relationship?" *Journal of European Economic Association*, 7 (2009), 786–830.

Schinasi, G. J., and P. A. V. B. Swamy. "The Out-of-Sample Forecasting Performance of Exchange Rate Models When Coefficients Are Allowed to Change." *Journal of International Money and Finance*, 8 (1989), 375–390.

Sharpe, W. F. "Mutual Fund Performance." *Journal of Business*, 39 (1966), 119–138.

Sullivan, R.; A. Timmermann; and H. White. "Data-Snooping, Technical Trading Rule Performance, and the Bootstrap." *Journal of Finance*, 54 (1999), 1647–1692.

Timmermann, A. "Forecast Combinations." In *Handbook of Economic Forecasting*, G. Elliott, C. W. J. Granger, and A. Timmermann, eds. Amsterdam, Netherlands: Elsevier (2006), 135–196.

White, H. "A Reality Check for Data Snooping." *Econometrica*, 68 (2000), 1097–1127.

Wright, J. H. "Bayesian Model Averaging and Exchange Rate Forecasting." *Journal of Econometrics*, 146 (2008), 329–341.