Staff memo

Forecasting short-term movements in the Swedish krona

Ana Maria Ceh
Monetary Policy Department

november 2020

S V E R I G E S R I K S B A N K

Table of contents

Abstract

Ana Maria Ceh
Senior Economist, Macrofinancial Analysis Division, Monetary Policy Department

Nominal exchange rate dynamics are often difficult to relate to macroeconomic fundamentals. This fact has been documented by Meese and Rogoff (1983), who noted that exchange rate models are unlikely to beat the random-walk or no-change forecast, and has been surprisingly robust ever since. In this paper we evaluate whether the Meese and Rogoff (1983) result holds for the Swedish krona exchange rate by assessing the forecasting performance of a few exchange rate models, whose explanatory factors are available at high frequency. We find that random-walk remains the benchmark whose forecasting performance is difficult to beat (at least over the shorter horizons). In terms of out of sample forecasting power, simpler, calibrated models tend to perform better than estimated ones with many free parameters. Nevertheless, more complex estimated models tend to capture a significant portion of in-sample krona variation.[1]

Acknowledgements: The author would like to thank Jan Alsterlind, Mikael Apel, Paola Di Casola, Jesper Hansson, Jens Iversen, Stefan Laséen, Xin Zhang and the seminar participants at the Riksbank's Monetary Policy Department for their valuable comments. All remaining errors are my own. The views and opinions expressed in this staff memo are solely those of the author and do not necessarily reflect the views of Sveriges Riksbank.

[1] The partial results of the current study have been published as a part of Askestad et al. (2019).

1 Introduction

In their seminal paper Meese and Rogoff (1983) documented the failure of standard economic models, developed for the purpose of understanding exchange rate movements, to beat the forecast performance of a simple random-walk model. The failure of conventional economic models to beat the simple no-change or random-walk forecast is a well-known fact, the result which the international finance literature has attempted to overturn ever since, yet the superiority of the random-walk forecast remains surprisingly robust feature (Moosa and Burns (2015)).

The goal of this paper is to perform an evaluation of conventionally used benchmark models and standard methods for forecasting the Swedish krona nominal exchange rate in the short term. It is inspired by the forecast evaluation exercises performed in Rossi (2013), and while it draws similar conclusions, the main one being that *the krona random-walk remains the benchmark whose forecasting performance is difficult to beat*, the models whose forecasting ability we evaluate are motivated by the research agenda in various articles outlined below. We consider the models of the nominal exchange rate that will be mostly used for nowcasting purposes, and their forecasting ability is evaluated at shorter horizons, up to one year ahead in time.[2] Finally, the performance of the models is assessed over several dimensions – a number of different criteria are employed and different samples are evaluated.

While analyzing the data and reviewing the vast and often contradictory literature findings on exchange rate predictability, Rossi (2013) attempted to offer an answer to the question "Are exchange rates predictable?". The exercise performed throughout the paper is to try to beat the benchmark in terms of forecasting performance. The benchmark is set to be the random-walk without drift, motivated by the *Meese and Rogoff puzzle*, that has since become common in the literature.[3] The review employs *the horserace approach* in assessing the relative empirical content of different exchange rate models – it is *the evaluation of the performance of different models in predicting the actual exchange rate level when the determinants are assumed to be known*. The exchange rate predictability is found to depend on a number of selection criteria such as the *(i) choice of predictor*, *(ii) choice of model*, *(iii) the dataset used for testing* (both in terms of frequency and the training sample period), *(iv) the forecast horizon* and *(v) the forecast evaluation method*. In terms of the latter, two criteria are employed to evaluate the forecasts: *(i) (R)MSE metric*, measuring whether the variability of the predictions around the actual values is greater than or less than that obtained by the no-change prediction, ie. random-walk model; and *(ii) direction of change criterion*, measuring the proportion of forecasts that correctly predict the direction of change. In the short run, no model consistently outperforms the no-change or random-walk prediction. Over the longer horizon, a model specification incorporating the long-run relationship between the exchange rate level and the level of fundamentals outperforms specifications involving growth rates. The recent review by Cheung et al. (2019) further expands the set of commonly evaluated models to include Taylor-rule fundamentals, yield curve factors, and incorporate shadow rates and risk and liquidity factors, whose performance is then compared to the random-walk benchmark. They offer similar answers while also observing that accounting for the risk and liquidity in the recent period tends to improve the fit without much improvement in the forecasting performance. In addition, they find that the euro/dollar exchange rate appears to be the one particularly difficult to predict.

Nominal exchange rate dynamics are often difficult to relate to macroeconomic fundamentals. The random-walk superiority in (out-of-sample) forecasts suggests that exchange rate movements are orthogonal to (commonly used) fundamentals. Nominal exchange rate dynamics are as well difficult to relate to cross-country

[2]Given the delay in published macro outcomes and statistics, the first step in the general forecasting process (at quarterly frequency) is the assessment of the starting point for the forecast, called "nowcast".

[3]The concept of "puzzle" in economics is related to the fact that the implications of economic theory are at odds with the empirical evidence. This empirical pattern, of the inability of the exchange rate forecasting models to outperform the simple random-walk, is defined as a puzzle by Obstfeld and Rogoff (2000). The authors, though, consider it the manifestation of a broader puzzle, called the "exchange rate disconnect puzzle", according to which economic fundamentals cannot explain exchange rate movements. Other studies followed the seminal paper by Meese and Rogoff (1983), but could not fully overcome their conclusions.

differences in nominal interest rates. The *uncovered interest rate parity puzzle* suggests that high interest rate countries tend to have higher expected currency returns, at least in the short run, which is at odds with the theory. Much of the research literature is focused on establishing whether and why the uncovered interest parity fails in the data. Another strand of the international finance literature is focused on finding a "profitable" forecasting strategy. The failure to find fundamentals that co-move with exchange rates or forecasting models with even mild predictive power, facts broadly referred to as "exchange rate disconnect", is one of the most robust facts in international macroeconomics. Even though the theory of exchange rate determination managed to produce a number of plausible models, some of which generate sufficiently satisfactory in-sample explanatory power, in terms of out-of-sample forecasting tests, they generally fail to outperform the random-walk. According to Engel et al. (2007) many standard exchange rate models themselves imply near random-walk behavior of the exchange rate, so naturally their power to beat the random-walk in out-of-sample forecasts is low. The argument arises from the fact that short-term movements in the exchange rate are primarily driven by the shift in expectations (about future fundamentals), while in standard models the current economic fundamentals have relatively little weight in determining the exchange rate and therefore should not be expected to have much power in forecasting exchange rate movements.[4,5]

In line with the various literature findings, and given the evaluation of forecasting performance according to different criteria, as well as larger complexity levels and the inability of other models to significantly differentiate from the random-walk prediction, the random-walk model seem to be the best candidate to perform krona forecasts at short horizons, up to a few months ahead in time. However, even if some models can deliver a slight advantage over the random-walk projection under certain performance metrics, this is highly dependent on the training sample and is not uniform across different evaluation criteria.

The remainder of this paper is organized as follows. Section 2 introduces the set of models whose forecasting performance will be evaluated, while Section 3 describes the dataset with its limitations and proposes some selection criteria to discuss the performance of individual models. Section 4 then reports and evaluates forecasts from estimated models, while recognizing the importance of the training sample for the reported results. Finally, Section 5 concludes by summarizing findings and proposing *the best candidate* to perform the short-run forecast.

2 Model selection

This section introduces and describes the set of models whose forecasting performance will be subsequently evaluated. The choice of models introduced here is motivated by the more recent theoretical and empirical studies from the asset pricing literature (Cheung et al. (2019)), together with the availability of data at higher frequencies, which allows us to decompose krona movements in real time.

As pointed out in Engel et al. (2007), the random-walk superiority in (out-of-sample) exchange rate forecasting is not directly at odds with the theory since many economic models actually imply that the exchange rate is "nearly" a random-walk. Therefore, the natural benchmark for performance comparison between competing models is the random-walk forecast (see figure 1, left), stating that exchange rate changes are not predictable,

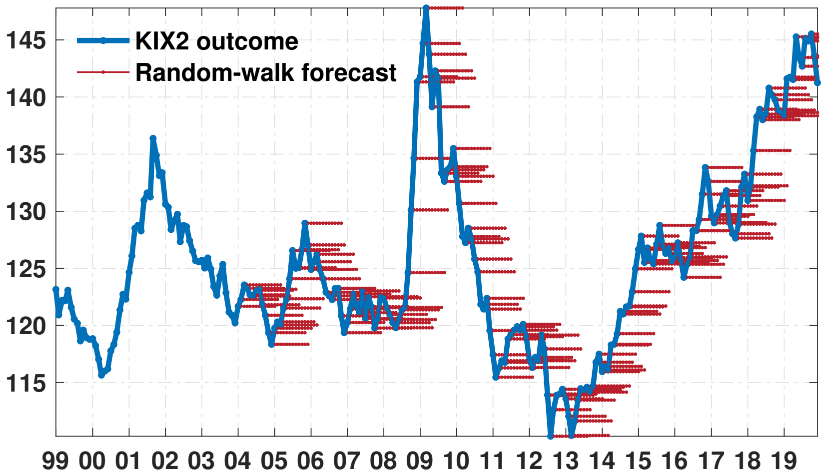
$$\mathbb{E}_t[\Delta s_{t+h+1}] = 0, \quad (1)$$

with S_t being the domestic price value of foreign currency, $s_t \equiv \log(S_t)$ stands for the logarithm of the krona exchange rate level so that $\Delta s_{t+h+1} = s_{t+h+1} - s_t = \log \frac{S_{t+h+1}}{S_t}$ denotes currency depreciation at horizon $h+1$ $(h = 0, 1, \ldots)$.

[4] Conversely, if exchange rates react to news about future economic fundamentals, then perhaps they can help forecast the (observed) fundamentals.

[5] More recently, Engel (2014) surveys current theoretical and empirical contributions on foreign exchange rate determination.

Figure 1: **Krona random-walk forecast**.



Note: This figure plots 12-months-ahead forecast for KIX2 coming from random-walk model. KIX2, or narrow KIX-index, is the euro and dollar competition-weighted krona exchange rate (see footnote 17 for more details).

An alternative version of the random-walk model is the random-walk with drift (figure 1, right), according to which exchange rate changes are predictable yet still independent of fundamentals,

$$\mathbb{E}_t[\Delta s_{t+h+1}] = \beta_0, \tag{2}$$

where $\beta_0 \neq 0$ is a non-zero constant.[6]

Motivated by the fact that the cross-country interest differential is carefully monitored by market participants and is often cited as one of the most important factors assumed to be driving the *krona level*,[7] we estimate the following empirical relationship and add it to the set of models to be evaluated:

$$\mathbb{E}_t[s_{t+h+1}] = \beta_0 + \beta_1 i_{t+h} - \beta_2 i^*_{t+h}, \tag{3}$$

where, as before, $s_t \equiv \log(S_t)$ denotes the log-level of the krona exchange rate and i_t, i^*_t stand for the domestic and foreign interest rate level, respectively.[8] In contrast, the uncovered interest parity (UIP) condition implies a strong and positive correlation between the currency depreciation and cross-country interest differentials

$$\mathbb{E}_t[\Delta s_{t+h+1}] = i_{t+h} - i^*_{t+h}, \tag{4}$$

and is another model to be evaluated. In a similar fashion, the covered interest parity (CIP) states that the forward currency rate equals the spot rate adjusted for the cross-country interest rate differential

$$\mathbb{E}_t[s_{t+h+1}] = f_{t+h+1}(\equiv s_t + i_t^{h+1} - i_t^{*h+1}), \tag{5}$$

where $f_t \equiv \log(F_t)$ denotes the log-level of the krona currency forward rate, while i_t^h, i_t^{*h} stand for domestic and foreign interest rate level at horizon h, respectively. However, it is an empirical regularity that the high interest rate currencies tend to subsequently appreciate rather than depreciate according to the uncovered interest parity condition. Therefore, it turns out that the forward rate is not a good forecaster of the future spot rate (see figure 2, right) because both the spot and the forward rate are determined simultaneously via covered interest parity, giving rise to the forward premium puzzle (see Fama (1984)). Nevertheless, the forward premium anomaly has been well documented and evaluated in the financial literature and we will as

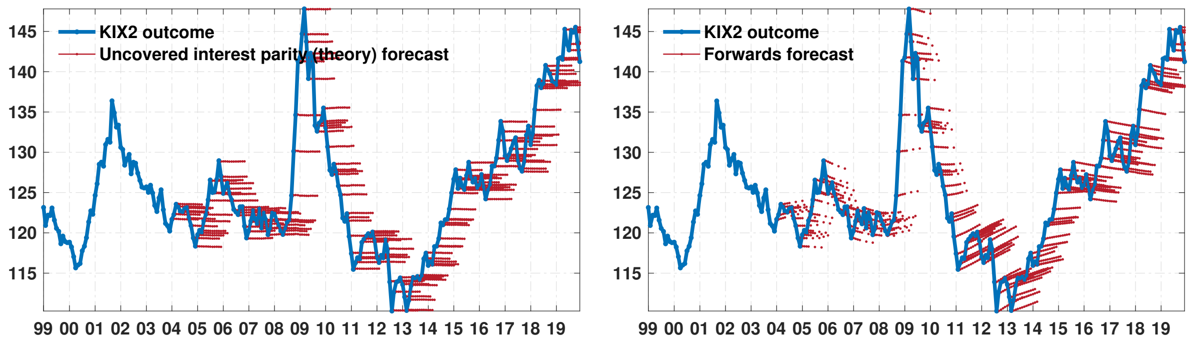
[6] The estimated non-zero constant is fairly close to zero, hence the forecasts coming from the two random-walk models, without and with drift, are almost indistinguishable.

[7] For details see figure B.1a in the appendix, together with some simple summary statistics about the Swedish krona.

[8] Due to a better empirical fit, instead of the commonly used one-period interest rate, our preferred choice is the somewhat longer maturity one-year interest rate, spanning our forecast horizon of 12 months.

well evaluate the ability of the CIP identity to correctly predict the future krona level.[9]

Figure 2: **Krona interest parity forecasts**.



Note: This figure plots 12-months-ahead forecasts for KIX2 coming from uncovered interest rate parity condition (left) and implied KIX2 forwards (right) corresponding to covered interest parity condition.

Despite its empirical failings, the UIP is and remains an enduring benchmark in the foreign exchange literature. The reason behind this is the fact that the UIP possesses a handful of appealing properties. First, the UIP model is quite simple and determined by asset prices alone. Asset prices are frequently updated and observable at high frequencies, unlike infrequently updated and imperfectly measured macroeconomic data. Second, it has no free parameters to be estimated in-sample, which makes it well suited to out-of-sample forecasting.[10] Third, it has a straightforward and intuitively appealing interpretation as the expected exchange rate movement perceived by a risk-neutral investor. Therefore, while relying on the theory in the background, we will try to keep some of these appealing features and they will be the guidelines in the empirical model design.

Given the complexity and integration of financial markets, a portfolio consisting of only risk-free assets may seem too simplistic. Following the idea that *all* the financial flows in and out of an economy affect its exchange rate, one may augment the UIP to acknowledge the possibility of multiple assets in the portfolios of domestic investors. Therefore, the set of investment opportunities is expanded to include, apart from the bond market, equity market instruments as well. Typically, in expectation, equities (as risky assets) offer higher payoff than risk-free assets, and assuming the profitable strategies across domestic and foreign markets would be arbitraged away, we posit the following uncovered return parity (URP) condition [11]beginalign $\mathrm{E}_t[\Delta s_{t+h+1}]-(i_{t+h}-i^*_{t+h}) = \frac{1-\mu}{1+i_{t+h}}\,\mathbb{E}_t[r_{t+h+1}-i_{t+h}]-\frac{1-\mu^*}{1+i^*_{t+h}}\,\mathbb{E}_t[r^*_{t+h+1}-i^*_{t+h}]$, where $r_{t+h+1} = \log\frac{E_{t+h+1}}{E_t}$ defines the equity price return at horizon $(h{+}1)$ and, as before, Δs_{t+h+1} stands for horizon $(h{+}1)$ currency depreciation, while i_t corresponds to the domestic interest rate level. Starred letters denote foreign country variables. The excess currency return is defined as $rx_{t+1} = \Delta s_{t+1} - (i_t - i^*_t)$, often used as a proxy for risk premium (as it represents ex-post deviations from the UIP). The uncovered return parity relates expected excess currency return to the relative equity premiums across countries (in proportion to their respective shares in the portfolio) and nests the uncovered interest parity as a special case when no investments are made in equities, that is when $\mu = \mu^* = 1$.

In addition to the portfolio channel highlighted by the URP condition, another potential avenue for explaining the risk premium or, as previously stated, the deviations from the UIP measured as excess currency returns, is the information contained in the yield curve. The finding in Chen and Tsang (2013) suggests that augmenting the standard UIP relationship with longer maturity rates makes the UIP puzzle disappear. This finding is

[9]Unlike the UIP, until recently the CIP has been well supported empirically, though since the onset of the global financial crisis the deviations from the CIP relationship have significantly increased (see Cerutti et al. (2019)).

[10]When estimated at high frequencies, many models suffer from instability of coefficient estimates, as supported by a "scapegoat theory" of exchange rate fluctuations (see Bacchetta and Van Wincoop (2004)).

[11]See Djeutem and Dunbar (2018) for the motivation and derivation of a similar condition to the one below,

$$\mathbb{E}_t\left[\frac{S_{t+1}}{S_t}\left\{\mu^*(1+i^*_t)+(1-\mu^*)(1+r^*_{t+1})\right\}\right] = \mathbb{E}_t\left[\mu(1+i_t)+(1-\mu)(1+r_{t+1})\right],\ -\infty<\mu,\mu^*\leq\infty.$$

indicative of a puzzle being related to an omitted risk premium that is embodied in the rest of the yield curve. Hence, we will use yield curve factors to try to capture the systematic response of the risk premium to the yield curve shape. The yield curve shape in turn captures market expectations of future inflation, output growth and other macro indicators, as discussed in Chen and Tsang (2013). Since various structural exchange rate models can deliver the exchange rate determined by the expected future values of cross-country output, inflation, and interest rates (see Engel (2014)), Chen and Tsang (2013) result points to the exchange rate movements not being disconnected from fundamentals but instead tied to them via the present value asset pricing equation (to the extent that the yield curve is shaped by market expectations about future macroeconomic fundamentals).

While empirically assessing the validity of the above-proposed channels (interest parity, portfolio channel and expectations of future fundamentals via yield curve factors) in modelling the systematic response of the risk premium, we will evaluate each of them individually, as well as jointly in an empirical model designed to incorporate multiple channels at once. In that respect, we will separate the models into two groups, Simple and Complex, where the two categories are defined as single versus multiple channels at work. Finally, each of the models is estimated at time t delivering the model parameters' estimates that will be used, together with the future realizations of the model determinants, in order to produce the $t+h+1$-horizon-ahead forecast $(h=0,1,\dots)$.

An empirical version of the uncovered interest parity condition is designed to help test whether the uncovered interest parity condition holds in data,

$$\Delta s_{t+h+1} = \beta_0 + \beta_1(i_{t+h} - i^*_{t+h}) + e_{t+h+1}, \tag{6}$$

Likewise, to empirically assess the validity of the uncovered return parity condition, we will use the following empirical approximation to the URP model

$$rx_{t+h+1} = \delta_0 + \delta_1 oil_{t+h} + \delta_2 vol_{t+h} + \gamma_0 ex_{t+h+1} + \gamma_1 ep_{t+h+1} + \gamma_1^* ep^*_{t+h+1} + e_{t+h+1}, \tag{7}$$

where, as before, rx_{t+1} denotes the excess currency return, $ex_{t+1} = r_{t+1} - r^*_{t+1}$ defines relative equity returns and $ep_{t+1} = r_{t+1} - i_t$, $ep^*_{t+1} = r_{t+1} - i^*_t$ represent home and foreign equity premiums. oil_t denotes oil-price-returns, and vol_t represents log-differences in stock market option implied volatility (VIX).[12] The success of the yield curve factors in explaining the exchange rate movements will be tested via

$$\Delta s_{t+h+1} = \beta_0 + \beta_1 L^R_{t+h} + \beta_2 S^R_{t+h} + \beta_3 C^R_{t+h} + e_{t+h+1}, \tag{8}$$

where $\{L^R_t, S^R_t, C^R_t\}$ denote relative yield curve factors extracted from cross-country yield curve differences, that is the term structure of cross-country interest differentials. The joint success of the yield curve factors, together with the equity market returns, in explaining currency movements is assessed in the following model

$$\begin{aligned} rx_{t+h+1} = {} & \delta_0 + \delta_1 oil_{t+h} + \delta_2 vol_{t+h} + \gamma_0 ex_{t+h+1} + \gamma_1 ep_{t+h+1} + \dots \\ & + \gamma_1^* ep^*_{t+h+1} + \beta_1 L^R_{t+h} + \beta_2 S^R_{t+h} + \beta_3 C^R_{t+h} + e_{t+h+1}. \end{aligned} \tag{9}$$

Moreover, removing the constraint on the cross-country yield curve factors to be of equal weight and opposite sign results in a *risk premium* model

$$\begin{aligned} rx_{t+h+1} = {} & \delta_0 + \delta_1 oil_{t+h} + \delta_2 vol_{t+h} + \gamma_0 r^*_{t+h+1} + \gamma_1 (ep_{t+h+1} - ep^*_{t+h+1}) + \dots \\ & + \beta_1 L_{t+h} + \beta_2 S_{t+h} + \beta_3 C_{t+h} + \beta_1^* L^*_{t+h} + \beta_2^* S^*_{t+h} + \beta_3^* C^*_{t+h} + e_{t+h+1}, \end{aligned} \tag{10}$$

[12] Volatility and oil price changes are included as controls since in our specification we are using realized equity returns r_{t+1}. Including both s_{t+1} and r_{t+1} in the regression specification might result in an invalid specification due to correlated expectational errors for s_{t+1} and r_{t+1}. In case the controls are correlated with the expectational errors, the remaining errors might be orthogonal.

where $\{L_t, S_t, C_t\}$ denote level, slope and curvature factors of the domestic yield curve, while starred letters denote their foreign country counterparts.

All of the models outlined above are simply variations of the random-walk model with the addition of different factors identified in the theory (interest differentials, equity returns, equity premium and relative yield curve factors) and designed to capture and explain the observed currency in-sample dynamics. The model in levels is included due to frequently stated empirical regularity (see figure B.1a), often referenced by market participants.[13] Together with other candidates, we will assess its performance and compare it to the no-change model. In this way, we evaluate each model's forecasting performance against the commonly used benchmark, the random-walk.[14,15] The models are estimated as one period models (the case with $h = 1$), and their estimates are used to produce multi-period forecasts (with $h > 1$), under the assumption of perfect foresight pertaining to all of the explanatory factors.[16]

3 Dataset, sample selection and model evaluation criteria

In trying to model krona movements, we focus on euro and dollar competition-weighted krona (KIX2 or narrow KIX-index), as well as bilateral exchange rates versus krona (SEK per EUR and SEK per USD). Our bilateral nominal exchange rates are quoted as the number of units of domestic currency per unit of foreign currency, either euro or dollar, in which case an increase in the exchange rate corresponds to a depreciation of the domestic currency. As potential explanatory factors and the determinants of the observed krona movements we use financial data for Sweden, the Eurozone (or Germany as representative) and the United States. The dataset consists of stock market indices (OMX30, STOXX, SP500) price returns and measures of their implied volatility (VSTOXX, VIX), oil prices (Brent Crude as a benchmark price for oil purchases worldwide) and Nelson-Siegel-Svensson estimates of the government bond yield curve in the respective countries (Sweden, Germany, the United States).[17]

A natural start of our dataset sample is the beginning of 1999, at the inception of the euro currency. Data are collected at daily frequency and converted to monthly frequency as end-of-month values. To mimic the latest information available at the end of each period, we use daily observations of the last trading day of each month to represent the whole month.[18] The interest rates and equity returns used in the models are monthly holding period returns, as are subsequently computed excess currency returns. All data come either from Macrobond or are derived as the Riksbank's own calculations, and the sample runs monthly from January 1999 to December 2019.[19]

Evaluating predictive performance

In order to measure the forecasting performance across the set of different models, we employ a few commonly used selection criteria: the root mean square error and direction of change metrics. Let S_t denote the exchange

[13] Also, as mentioned in Rossi (2013), the models specified in levels perform better along long(er) horizons.

[14] All of these models rely on indicators that are available at high frequency, which is needed to inform and frequently update the short run forecast during the "nowcasting period."

[15] The models displayed above are listed together with their respective chart names in the appendix A.

[16] In a sense this kind of "forecast valuation design" gives an "unfair advantage" to all the models over the random-walk. Yet, as we will see, even under these conditions the random-walk maintains it's dominance.

[17] In the following analysis the foreign variables we consider are weighted averages of the corresponding variables for the Eurozone and the United States, with the same weights as in the narrow KIX index. Given the high correlation between the full and the narrow index, the conclusions from our analysis should hold for the full KIX index with minor differences.

[18] Since the purpose of this analysis is mainly to assess the model's performance in forecasting the krona at short horizons, the end-of-month data have been used as a proxy for the latest relevant information needed to timely update the view on the current quarter currency value, that is the exchange rate nowcast.

[19] Interest rates used are daily estimated zero-coupon yields coming from a Nelson-Siegel-Svensson fitted yield curve. Nelson-Siegel-Svensson estimates of the yield curve for Sweden on a regular daily frequency start in mid-1999, but were produced on a weekly basis prior to that period, so the latest observable value during the corresponding month was used as an end-of-period value. Nelson-Siegel-Svensson zero-coupon yields at $m = 1, 2, \ldots, 120$ durations are used to extract principal components which are used as level, slope and curvature factors in order to describe the yield curve shape.

rate outcome at time t, and let $\hat{S}_t$ denote the model produced forecast for the period t. Assume further that we evaluate the performance of the models in forecasting h-periods ahead, over the following $t = m+h, \ldots, n$ periods. The standard root mean square error metric is defined as

$$RMSE_m^{(h)} = \sqrt{\frac{1}{n-m-h+1} \sum_{t=m+h}^{n} \left(\hat{S}_t - S_t\right)^2} \, .$$

The RMSE measures deviations of forecasts from the true outcomes, and it depends on k, the forecast horizon, as well as m, the beginning of the evaluation sample. Larger deviations from the true outcomes are penalized more heavily, while this criterion places no value on true predictions of direction of change. For this purpose, the conventional measure of direction accuracy is given by

$$DA_m^{(h)} = \frac{1}{n-m-h+1} \sum_{t=m}^{n-h} \mathbf{1}\{(\hat{S}_{t+h} - S_t)(S_{t+h} - S_t) > 0\} \, .$$

The RMSE measures the prediction accuracy in terms of magnitude of the forecasting error, while the DA measures the proportion of model forecasts that correctly predict the direction of change, as at times the direction of change may be more important than the magnitude of error itself.[20]

In addition to measures of model accuracy, one may want to evaluate the model's tendency to produce systematic deviations over true outcomes. In order to measure the tendency of a model forecast to consistently over- or under-estimate the true values, we define a measure of bias

$$Bias_m^{(h)} = \frac{1}{n-m-h+1} \sum_{t=m+h}^{n} \left(\hat{S}_t - S_t\right),$$

as a mean deviation from the true outcome.

Finally, one may want to assess the ability of the model forecast to outperform the random-walk benchmark in a statistically significant manner. For that reason, we run the Diebold-Mariano test (Diebold and Mariano (2002)) of the null hypothesis that the random-walk forecast and the forecast coming from the competitor model perform equally well.[21] Formally, we test the null hypothesis that the forecast errors coming from the competitor model are not significantly different than the ones coming from the random-walk projection. With quadratic loss function, in practice this corresponds to the equality test of the two forecast error variances (here we additionally use the adjustment proposed by Harvey et al. (1997)).

4 Results

All the models discussed in section 2 are initially estimated on the $1999:1 - 2003:12$ sample and subsequently re-estimated by recursively adding one additional end-of-month observation until $2019:12$. Each model estimate is used to produce forecasts up to 12-months ahead in time and forecasting errors across different horizons are used to produce the relevant statistics.

In-sample fit

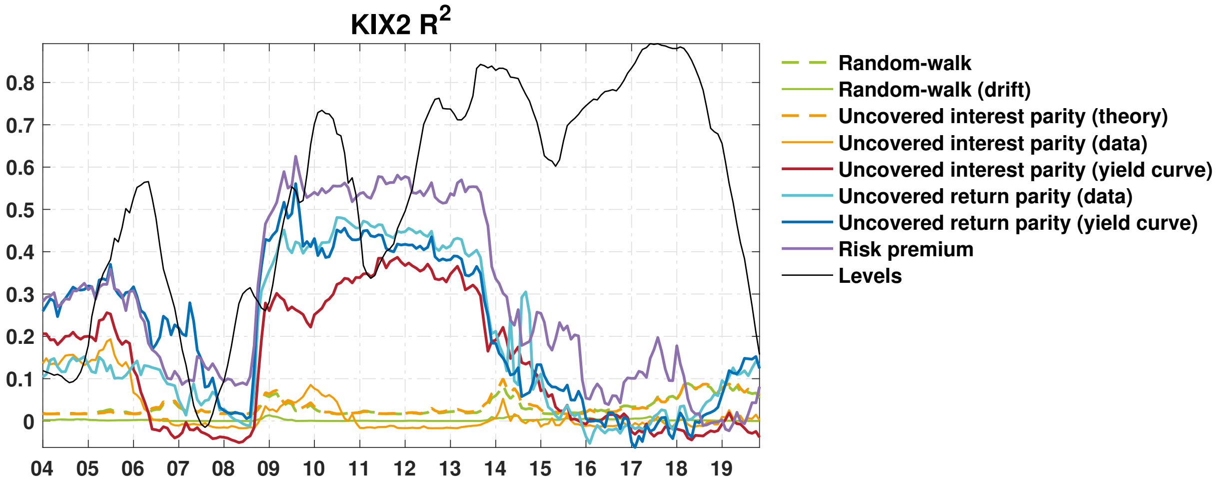
Before evaluating the forecasting performance, it is instructive to examine the ability of the models to explain the observed movements within the sample. The outcome of this exercise is reported in figure 3, which plots the

[20] Sometimes both magnitude of error and the ability of the model to predict direction correctly matter, in which case the two criteria may be merged into a single measure. The two of the above-defined measures could be combined together into adjusted root mean squared error as follows $ARMSE_m^{(h)} = \sqrt{\frac{1-DA_m^{(h)}}{n-m-h+1} \sum_{t=m+h}^{n} \left(\hat{S}_t - S_t\right)^2}$.

[21] In Diebold (2015) the author points out that the purpose of the Diebold-Mariano test is to compare forecasts, not models. In what follows, we will use the Diebold-Mariano test to compare the (pseudo)-out-of-sample forecasts, not evaluating the model fit.

time-varying share of fluctuations in krona explained by the model's determinants, across all of the evaluated models. Even though we assume the perfect predictability of explanatory factors, most of these models seem to capture very little of variation observed in data. Overall, the subset of models carrying the forward looking components embedded in the equity markets or the yield curve shape seem to separate themselves from the standard model benchmarks, suggesting that *Complex models* have more power than *Simple* ones in explaining the in-sample movements.

Figure 3: **Krona variation explained in-sample**. The percentage of krona variation explained by the model determinants.



Note: The figure shows recursive window adjusted R^2 estimates across different models for the narrow krona effective exchange rate. All models are estimated on a $1999:1-2003:12$ sample and subsequently re-estimated by sequentially adding one new (end-of-month) observation until $2019:12$. Each model estimate is used to produce a time-varying R^2 coefficient.

The share of the in-sample variability in krona movements explained by the relative country yield curve factors (specification 8) is an order of magnitude higher than the one explained solely by the interest differentials, at least during the crisis era. Despite the fact that the uncovered interest parity normally fails empirically, Chen and Tsang (2013) find that it holds in the longer run, suggesting that an omitted risk premium captured by the yield curve shape is an important driver of excess currency returns. As it turns out, adding relative yield curve factors offers further improvements and explains the additional in-sample variation, which confirms the findings in Chen and Tsang (2013) for the Swedish krona case. In recent years, though, these added gains seem to disappear, coinciding with the period when the term premiums are compressed due to massive asset purchase programs, either still continuing, or being only gradually tapered.

On the other hand, adding relative equity returns and equity premiums, as is suggested by the uncovered return parity condition (model 7), captures a substantially larger portion of the in-sample variation. The uncovered return parity condition appears to have relatively strong empirical support and is suggestive of the importance of financial capital flows and equity risk premiums for currency risk premium determination (as suggested by Hau and Rey (2005), and empirically supported by Cappiello and De Santis (2007)).

Merging explanatory factors implied by both conditions into a risk premium model (model 10) further improves the fit. Even though time-varying, the in-sample variation explained by the risk premium model is mainly above the uncovered return parity model (model 7).[22] Moreover, the uncovered return parity model and its variants (yield curve factors augmented version and risk premium model) seem to outperform all the others by a significant margin. We note that these are fairly high numbers, even though the performance varies over

[22]In the risk premium specification (model 10), yield curve factors in individual countries turn out to be strong determinants across all currency pairs, in comparison to the yield curve augmented uncovered interest parity specification (model 8) where they are immediately taken in relative terms as cross-country factor differentials, and instead turn out to have less of explanatory power. In a similar way as the empirical UIP, coefficient estimates are at odds with the theoretical ones, this is probably due to the fact that the individual factor-country coefficients, although of opposite sign (as expected) are not exactly of the same magnitude.

time.[23] Additionally, financial factors contributing to greater in-sample explanatory power in the case of the krona-dollar relative to the case of krona-euro, with a fairly significant increase following the global financial crisis.This result is in line with the recent findings in Lilley et al. (2019), who find the exchange rate reconnect to U.S. foreign bond purchases since the global financial crisis and offer suggestive evidence that these flows pick up changes in risk premiums (supporting the narrative that the US dollar's role as an international and safe-haven currency has surged since the global financial crisis). Here too, as already argued above, financial factors seem to proxy for the unobservable risk.[24]
The additional non-trivial power to explain the krona variation, at least in-sample, likely brought on by the perfect foresight of otherwise highly unpredictable financial factors, though unrealistic, plays an important role in this in-sample fit exercise. In what follows we will describe the out-of-sample forecasting performance of the models in comparison to the out-of-sample forecasting abilities of the random-walk.

Models' Out-Of-Sample Forecasting Performance

Detailed results of the previously described measures of predictive performance (and statistical tests) for the narrow krona effective exchange rate are reported in table B.2 in the appendix and are summarized in figures 4, 5, 6 below.[25] The accuracy measure of error magnitude (RMSE) is computed as a ratio relative to random-walk model performance. Therefore, the models with a ratio above one are performing worse relative to the random-walk forecast, while the models with a ratio below one are performing better than the random-walk benchmark. The bias measure is expressed as a percentage deviation from the true value, while the direction accuracy measures the share of accurately predicted model forecasts. By definition of predicting no change, the random-walk benchmark has zero direction accuracy. The p-values from the Diebold-Mariano test for the performance of the model relative to the random-walk report the probability that the model forecast is no different from the random-walk projection (ie. a low p-value corresponds to the model likely being different from the random-walk benchmark).

Figures 4, 5, 6 below report the out-of-sample forecasting performance of the krona exchange rate models, while simultaneously inspecting the model's performance across two dimensions. By plotting the RMSE ratio against the p-values from the Diebold-Mariano test (figure 4), we can inspect the model's ability to statistically outperform the random-walk benchmark (as in Lilley et al. (2019)). Likewise, by plotting the RMSE ratio against the direction accuracy, we can check whether the improvements in RMSE accuracy are coming at the expense of correctly predicting the direction of change (figure 5). Finally, by plotting the RMSE ratio against the mean error (bias), we can see how the model's systematic forecasting error (bias) relates to its accuracy relative to the random-walk (figure 6). As before, the models are classified into either a Simple or a Complex model group.The solid dots represent the statistic values at a one-month horizon. Lighter shades represent the statistic pair value further ahead in time. Finally, clear dots represent the statistic pair at a one-year horizon.

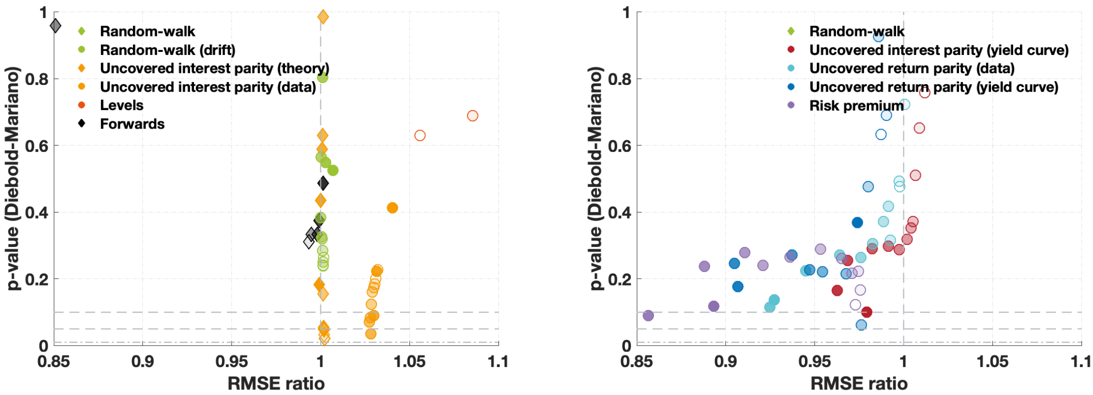
[23]The same holds true for the krona against the euro and dollar exchange rates (see figures C.7 and D.7 in the appendix), with the main difference being the role that crises have played in the bilateral krona exchange rate determination. Financial factors have had higher predictive ability during the Global financial crisis in the case of krona-dollar, while the European sovereign debt crisis seems to have played larger role in the krona-euro determination.

[24]Conventional wisdom and observed historical correlations (see table B.1 in the appendix) suggest that krona is to a large extent affected by foreign shocks and is following common global trends. This is somewhat confirmed within this model. Foreign equity markets are important as financial shocks gets amplified and spilled over to Sweden, in turn affecting krona excess currency returns. Krona-dollar exchange rate, in comparison to krona-euro, is more strongly influenced by equity return and equity premium movements.

[25]All of the described measures of forecasting accuracy for krona against euro and dollar bilateral exchange rates are reported in tables C.1 and D.1 in the appendix. Similar summary figures for the corresponding bilateral exchange rates can be found in the appendix as well.

Figure 4: **Beating the benchmark vs. error magnitude accuracy.**

(a) Simple models. (b) Complex models.

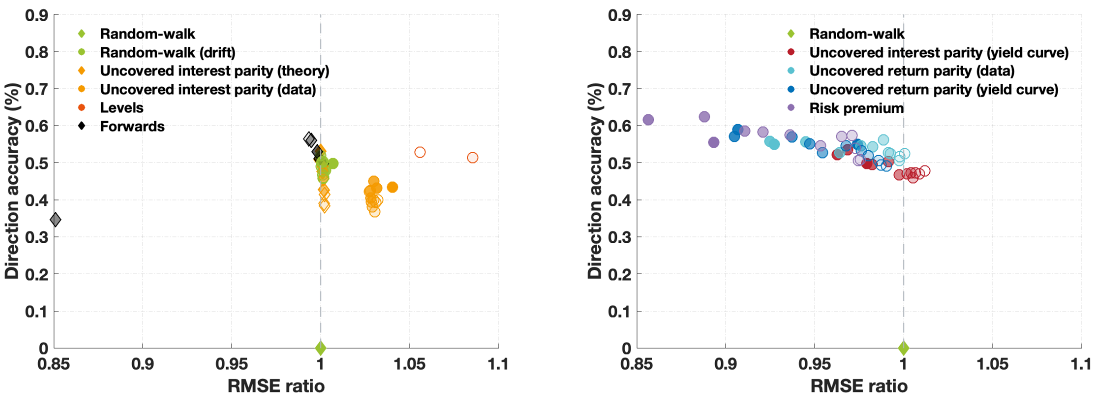


Note: This figure reports the performance of various exchange rate models relative to a random-walk over different forecast horizons. Each dot reports the p-value of Diebold-Mariano test statistic (y-axis) and the ratio of the model's root mean squared forecast error relative to a random-walk (x-axis). The solid dots represent the statistic values at a one-month horizon. Lighter shades represent the statistic pair value further ahead in time. Clear dots represent the statistic pair at a one-year horizon.

In terms of the ability to outperform the random-walk benchmark (model 1), we see that complex models are generally more successful compared to simple ones, as they have relatively smaller RMSE. Unlike Figure 4a, where we can see almost all the models aligning themselves along the vertical line at one, Figure 4b shows a grouping in the lower left corner. This grouping indicates the ability of complex models to outperform the random-walk, though mainly at shorter horizons (darker shaded dots), since some of them perform significantly better than the random-walk along the RMSE dimension (as measured by low Diebold-Mariano test p-values).

Figure 5: **Direction vs. error magnitude accuracy.**

(a) Simple models. (b) Complex models.



Note: This figure reports the performance of various exchange rate models relative to a random-walk over different forecast horizons. Each dot reports the direction accuracy statistic (y-axis) and the ratio of the model's root mean squared forecast error relative to a random-walk (x-axis). The solid dots represent the statistic values at a one-month horizon. Lighter shades represent the statistic pair value further ahead in time. Clear dots represent the statistic pair at a one-year horizon.

In Figure 5b we can see that the above-mentioned out-performance of complex models does not come at the expense of poorer direction accuracy. In fact, these models are positioned slightly higher at the upper left end of the unit vertical line, suggesting the dominance along the direction prediction too. Since by definition the random-walk predicts no change from the latest outcome, it effectively does not choose any direction of change. Therefore, the random-walk has the direction accuracy of zero. Unlike the simple random-walk, the other simple models, including the random-walk with a drift, get the direction right about half of the time, and are therefore outperforming the random-walk along the direction accuracy dimension.

Among the complex models, the risk premium model (model 10) comes across to combine the success of the models augmented with yield curve factors that seem to get the direction right (see figure 5b), together with the uncovered return parity model that appears to capture the magnitudes observed in data well (see figure 3). This mix in the end offers superior performance in terms of jointly capturing both the direction and magnitude and with roughly the same amount of bias.[26] At very short horizons, up to a few months ahead, the risk premium model (model 10) succeeds in outperforming the random-walk model in terms of error magnitudes, yet by a very narrow margin. This can be seen in figure 4b where darker shaded purple dots (RIP data) are placed in the far-left lower quadrant. These dots become lighter as we approach the unit vertical line from the left, suggesting that, as the horizon extends, the positive performance gap to the random-walk diminishes. On the other hand, even though the performance gap between the risk premium model and the random-walk decreases over time, in terms of getting the level direction right, the risk premium model outperforms all the tested models (purple shaded dots are clustered in the upper-left corner in the figure 5b), but it does so at the expense of increased bias (in figure 6b purple shaded dots stand lower compared to the others, indicating more of a negative bias in the model's forecasts), though the amount of added bias is relatively small.

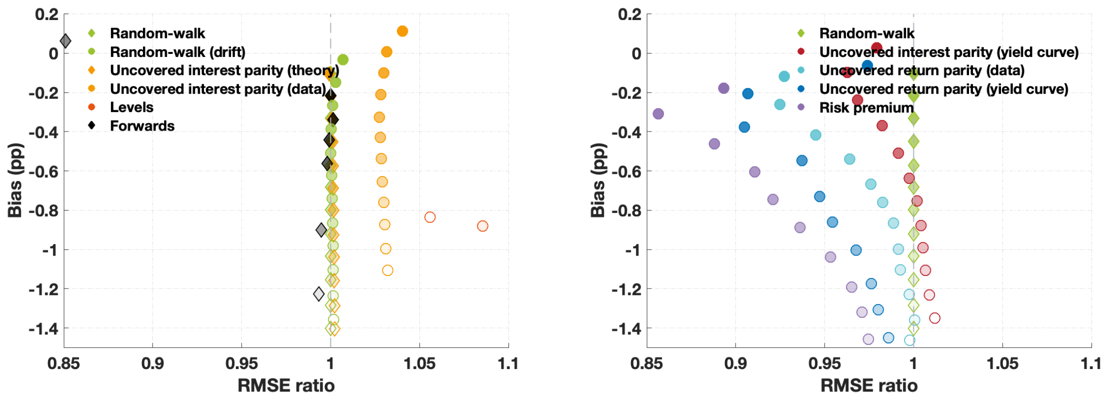
Another interesting result is that the uncovered interest parity condition, as predicted by theory (model 4), performs on a par with the random-walk when measuring the magnitude of forecasting error (RMSE). In figure 4a we can see that the yellow dots align over the unit vertical line. Surprisingly, this comes with the ability to correctly predict the direction more frequently and with no larger bias than the one implied by the random-walk model. This can be seen in figure 5a, where we have UIP data (yellow dots) positioned at the higher end of unit vertical line, signalling superior direction accuracy. At the same time, in figure 6a we see the random-walk data (green dots) mostly overlapping the UIP data (yellow dots), indicating approximately the same amount of bias implied by the two models. In particular, given that the empirically estimated uncovered interest parity relationship gets the direction "wrong" on average, the theoretical relationship performs much better in terms of error magnitude and comes quite close to the random-walk in terms of forecasting performance (figure 5a). The empirically estimated relationship has a different slope sign than that predicted by the theoretical condition, but at the same time it has a non-zero constant. The non-zero constant seems, at least in-sample, to partially cancel out the "wrong" slope effect, which in turn makes its forecasting performance close (or slightly worse) to the random-walk in terms of error magnitudes, yet better than the random-walk in terms of getting the direction of change right. Still, the empirical relationship performs slightly worse compared to the theoretical one since the error magnitudes seem to be larger, which in turn strongly affects the variability of predictions around the true values, and hence the mean square error measure of predictive performance. Overall, it seems that the predictions made with calibrated simple models are, at the very least, not performing significantly worse than those coming from their empirically estimated counterparts, as indicated by the Diebold-Mariano test statistic (figure 4a). Even though the estimated models are by assumption better suited to describing the data, and they do provide higher in-sample fit, it is not clear that they are superior in predicting the future outcomes. Hence, they are not necessarily better suited to out-of-sample forecasting.[27]

[26] The bias that is present throughout the whole set of evaluated models, approximately in the same amount, likely comes from the inability of the models to explain the continued depreciation observed in the latter part of our sample.

[27] This result is likely due to the high volatility and instability of estimated coefficients.

Figure 6: **Forecasting bias vs. error magnitude accuracy.**

(a) Simple models. (b) Complex models.



Note: This figure reports the performance of various exchange rate models relative to a random-walk over different forecast horizons. Each dot reports the mean error (bias) statistic (y-axis) and the ratio of the model's root mean squared forecast error relative to a random-walk (x-axis). The solid dots represent the statistic values at a one-month horizon. Lighter shades represent the statistic pair value further ahead in time. Clear dots represent the statistic pair at a one-year horizon.

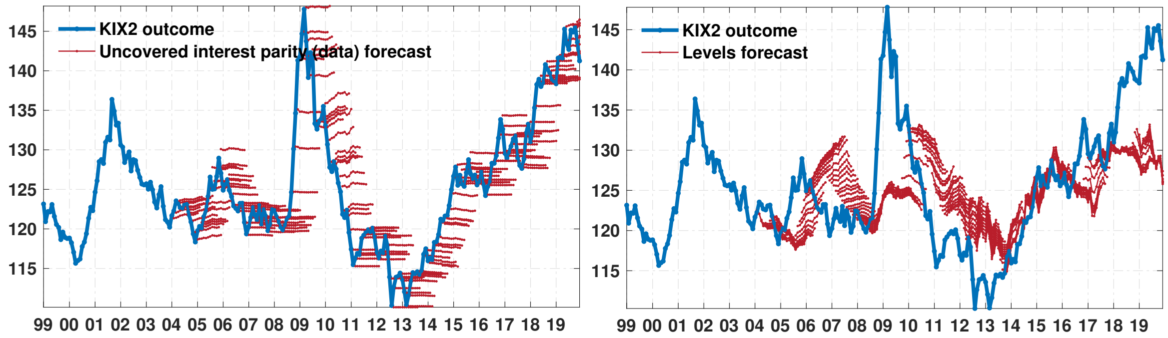
In terms of bias, among the simple models, the krona level model (model 3) and the empirical version of the UIP (model 6), depicted by orange and yellow dots, perform worse than the others at short horizons yet better at longer ones (for easier readability most of the orange dots and some of the black stars are out of the figure focus). The remaining random-walk and interest parity models (green dots and yellow diamonds) in figure 6a seem to overlap across all horizons (lighter shaded dots and diamonds), indicating no clear outperformance by any model. Among the complex models in figure 6b, most can be seen to almost align with the random-walk measure of bias across different horizons (green colored dots), yet still performing slightly worse along that dimension. Still, the yield curve augmented uncovered interest parity model, represented by the red colored dots, manages to separate itself slightly from the others, indicating somewhat less negative and marginally lower bias in the model forecasts. Therefore, in terms of forecasting error bias, the uncovered interest parity condition augmented by the information contained in the yield curve shape (model 8) seems to perform marginally better than the others, and it does so consistently across all horizons.

A larger and more negative measured bias points to inability of the models to explain the continued depreciation observed in recent years, since it indicates that model forecasts have consistently under-predicted the krona outcomes.[28] This difficulty of the estimated model projections to point in the direction of the observed depreciation, which is common across all the estimated models, can be seen in figures 7, 8, 9.

[28]For the case of the krona against dollar bilateral exchange rate, the uncovered return parity models are slightly more biased than the random-walk benchmark, which appears to be almost unbiased. The better test performance of unbiasedness of forecasting errors for the other two exchange rate candidates might be driven by a slight depreciation trend in the krona against the euro bilateral exchange rate (figures C.10 and D.10 in the appendix).

Figure 7: **Out-of-sample forecasts coming from empirically estimated interest parity models**.

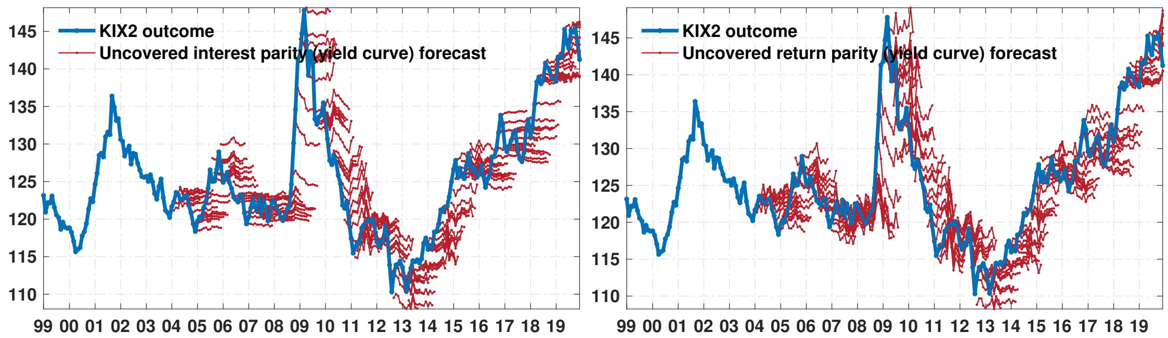
(a) Uncovered interest parity. (b) Krona level.



Note: This figure plots 12-months-ahead forecasts from a set of different models. All models are estimated on a $1999:1-2003:12$ sample and subsequently re-estimated by sequentially adding one new (end-of-month) observation until $2019:12$. Each model estimate is used to produce forecasts up to 12 months ahead in time.

Figure 8: **Out-of-sample forecasts coming from empirically estimated parity models augmented by yield curve factors**.

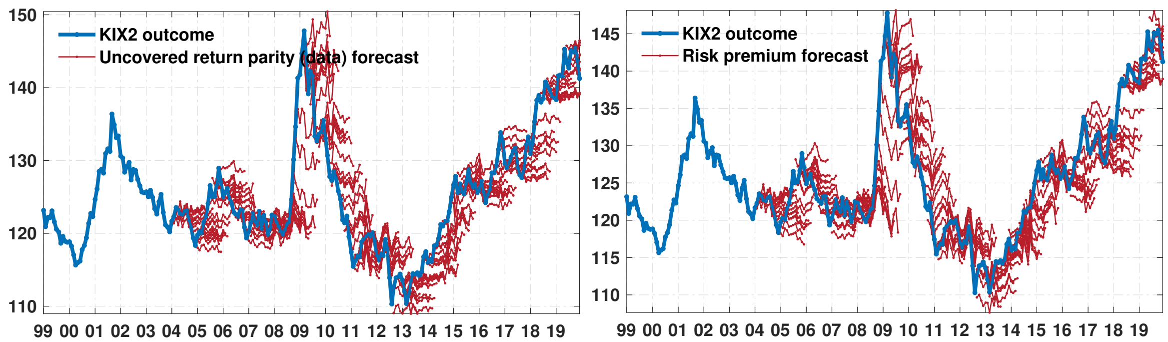
(a) Yield curve augmented uncovered interest parity. (b) Yield curve augmented uncovered return parity.



Note: This figure plots 12-months-ahead forecasts from a set of different models. All models are estimated on a $1999:1-2003:12$ sample and subsequently re-estimated by sequentially adding one new (end-of-month) observation until $2019:12$. Each model estimate is used to produce forecasts up to 12 months ahead in time.

Figure 9: **Out-of-sample forecasts coming from empirically estimated return parity models**.

(a) Uncovered return parity. (b) Risk premium model.



Note: This figure plots 12-months-ahead forecasts from a set of different models. All models are estimated on a $1999:1-2003:12$ sample and subsequently re-estimated by sequentially adding one new (end-of-month) observation until $2019:12$. Each model estimate is used to produce forecasts up to 12 months ahead in time.

In terms of prediction accuracy, the model of the exchange rate level performs poorly at short horizons across all reported measures. Due to the strong underperformance, it falls out of the scales plotted in the previous summary figures. Therefore, statistic values of different performance measures are reported in table B.2 in the appendix B. The level model appears to be the only model that consistently separates itself from the random-walk benchmark, by performing worse than the benchmark in a statistically significant manner (unlike any of the previously highlighted models). For the reason that it clearly fails to perform on all of the inspected performance criteria (see the reported statistics in table B.2), the level model (model 3) is not a good candidate for explaining krona movements in the short term. On the other hand, not taking a stand of a future direction at the time of a forecast decision (the assumption behind the random-walk forecast) seems to give rise to smaller errors in terms of error magnitudes but at the cost of the low direction accuracy.

Adding additional explanatory factors mainly seems to improve the model performance across the RMSE dimension. Nevertheless, most of the other models' predictions are statistically indistinguishable from the no-change projection in terms of error magnitude, yet they come at the expense of increased model complexity. Since the improvement in the combined accuracy comes only at the longer horizons, the added benefit of higher direction accuracy is likely not enough to overcome the difficulty of having to forecast all of the highly unpredictable financial factors (as one would have to do in the real-time forecasting exercise), and potentially induce significant errors in the model determinants.

Throughout the forecasting exercises performed in this paper, perfect foreseeability of explanatory factors is assumed and the exchange rate forecasts (across all horizons) have been made by using their true outcomes. Otherwise, these financial factors would have been likely predicted with significant forecasting errors, in particular during turbulent crisis periods, thus in turn affecting forecasting performance. In order to explore the importance of a perfect foresight assumption brought on by using the realizations of explanatory factors in a pseudo out-of-sample forecasting exercises performed, we will analyze an impact of the training sample (the parameter estimation sample) on the forecasting performance of our models. The sample dependence becomes the relevant proxy considering that our sample includes both the Global financial crisis and the European sovereign debt crisis periods. The following section will examine whether the models' performance statistics are significantly affected by the change of training sample.

Sample choice matters

In order to examine the importance of the training sample, we have repeated the forecast evaluation exercise, that now starts after the two crises periods, the Global financial crisis and the European sovereign debt crisis. Our models are initially estimated on a $1999:1 - 2013:12$ sample and subsequently re-estimated by sequentially adding one new observation until $2019:12$. Each model estimate is used to produce a forecast up to 12 months ahead and the forecasting errors across different horizons are used to produce the relevant statistics.[29]

The comparison with the previous results can be seen in figures 10, 11, 12 below. These figures show the performance of the models by comparing statistics at different horizons (one-month and one-year) in *before* and *after* samples. The *before sample* (solid dots) stands for the exercise where models' forecasting performance evaluation starts in 2004, before the emergence of the Global financial crisis, indicating that models have the advantage of perfect foreseeability of the crises captured through the lens of explanatory financial factors (to the extent that these factors are able to foresee the crises). Similarly, the *after sample* (clear dots) denotes the exercise in which forecasting performance evaluation starts in 2014, in the aftermath of the European sovereign debt crisis, meaning that these models did not have the difficult task of forecasting turbulent and extremely unpredictable crises periods (that otherwise might have been captured through the financial determinants).

The solid dots in the figures below represent the statistic values computed using the before sample (model forecasting performance evaluation starting in 2004, before the beginning of the Global financial crisis). Clear dots represent the statistic pair computed using the after sample (model forecasting performance evaluation starting in 2014, after the end of the European sovereign debt crisis).

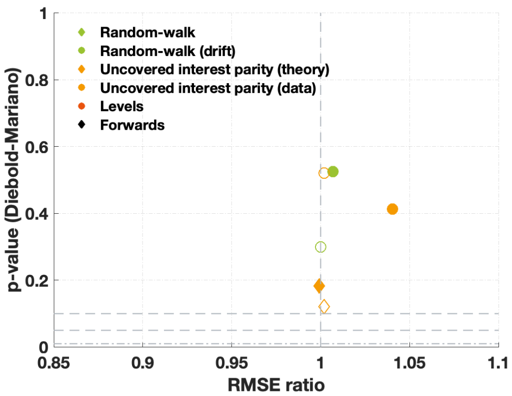
As anticipated, and in line with the findings in the literature (Rossi (2013), Cheung et al. (2019)), the

[29] All of the performance statistics for the candidate models re-evaluated on the new sample are summarized in table B.3 and figures B.2, B.3, B.4 in the appendix.

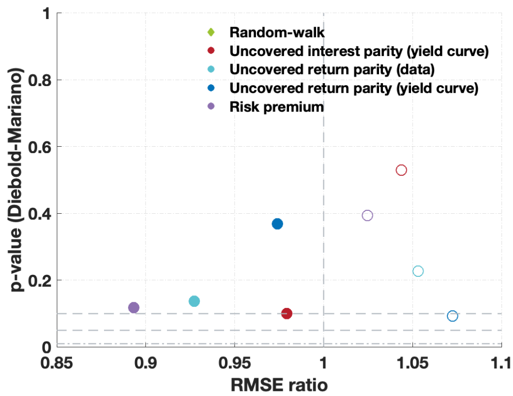
sample choice turns out to matter substantially for the krona forecasting performance. Unless we have a perfect foresight of financial markets during the crises periods captured in the *before sample*, the random-walk projection seems to outperform almost all of the models (in the short-run) in terms of error magnitudes in the *after sample* (compare the *before sample* results reported in table B.2 in the appendix B, with the *after sample* results in table B.3).

Figure 10: **Beating the benchmark vs. error magnitude accuracy (sample comparison).**

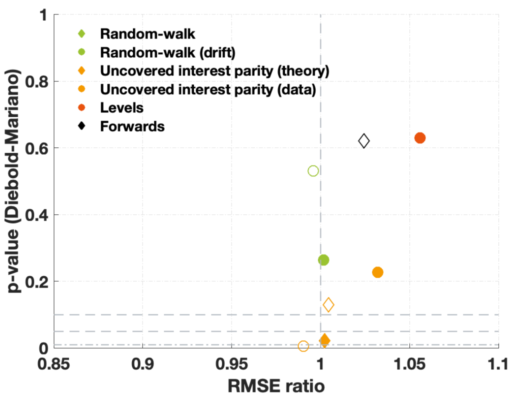
(a) Simple models, one-month forecasting horizon.



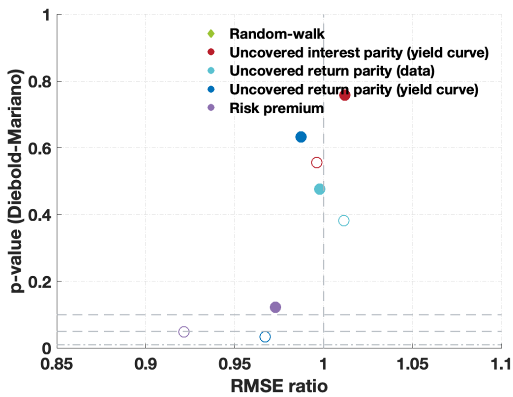
(b) Complex models, one-month forecasting horizon.



(c) Simple models, one-year forecasting horizon.



(d) Complex models, one-year forecasting horizon.

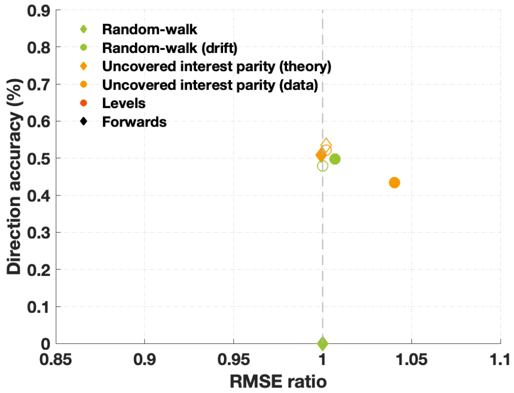


Note: This figure reports the performance of various exchange rate models relative to a random-walk over different sample periods. Each dot reports the p-value of the Diebold-Mariano test statistic (y-axis) and the ratio of the model's root mean squared forecast error relative to a random-walk (x-axis). The solid dots represent the statistic values computed over the *before sample* (model forecasting performance evaluation starting in 2004, before the beginning of the Global financial crisis). Clear dots represent the statistic pair computed over the *after sample* (model forecasting performance evaluation starting in 2014, after the end of the European sovereign debt crisis).

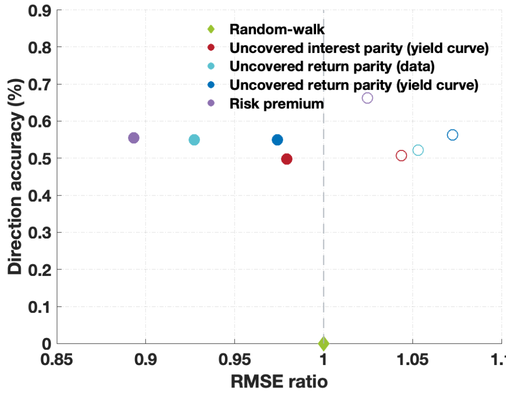
Among the simple models, we see no clear gain of any model's short-run forecast over the random-walk. As figures 10a and 10c show, all of the simple models tend to align close to, or to the right of, the unit vertical line. On the other hand, figure 10b shows that better accuracy of the complex models at short horizons in the *before sample* seems to disappear in the *after sample*. The *before sample* represented by the solid shaded dots has a low RMSE ratio, while the *after sample* represented by clear dots has a RMSE ratio above one. This result suggests that the reason behind the good predictive ability of the risk premium model might be the sample that captures two big crises (Global financial crisis and European sovereign debt crisis) combined with a perfect foresight of financial factors that are difficult to predict (particularly in terms of direction, but also magnitudes) just ahead of the crisis periods.

Figure 11: **Direction vs. error magnitude accuracy (sample comparison).**

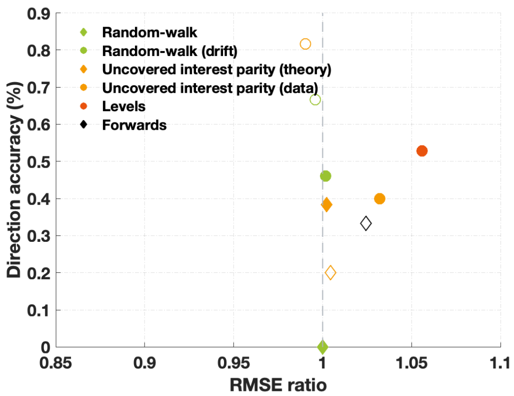
(a) Simple models, one-month forecasting horizon.



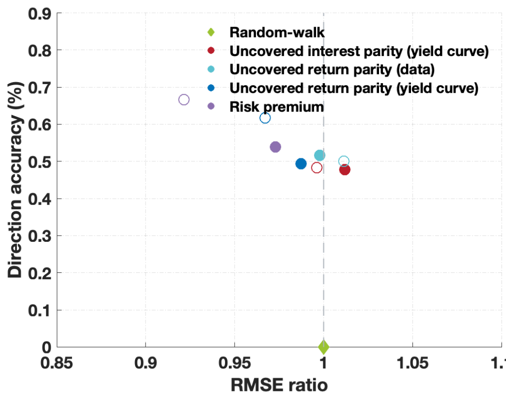
(b) Complex models, one-month forecasting horizon.



(c) Simple models, one-year forecasting horizon.



(d) Complex models, one-year forecasting horizon.



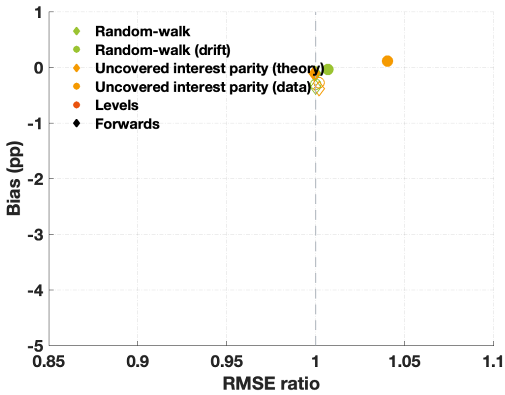
Note: This figure reports the performance of various exchange rate models relative to a random-walk over different sample periods. Each dot reports the direction accuracy statistic (y-axis) and the ratio of the model's root mean squared forecast error relative to a random-walk (x-axis). The solid dots represent the statistic values computed over the *before sample* (model forecasting performance evaluation starting in 2004, before the beginning of the Global financial crisis). Clear dots represent the statistic pair computed over the *after sample* (model forecasting performance evaluation starting in 2014, after the end of the European sovereign debt crisis).

The predictive ability of the simple models, in terms of direction of change (figure 11a) and mean error (figure 12a), does not seem to change much between the *before* and *after sample* at short horizons, as they are all clustered around roughly the same y-axis level. In contrast, the direction accuracy of complex models at the one-year horizon (figure 11d) seems to improve somewhat in the *after sample*, yet with significantly more negative bias (figure 12d). A significant increase in bias in the *after sample* at longer horizons is present across both Simple and Complex model groups (figures 12c, 12d).[30] As already discussed previously, the increase in bias in the *after sample* likely comes from the difficulty of the models in capturing the depreciation trend observed in the latter part of the sample.

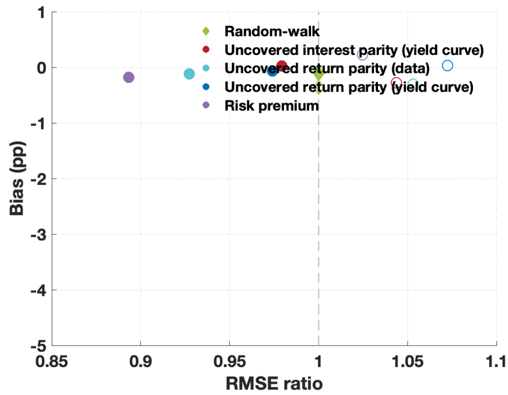
[30]A similar conclusion can be drawn from the performance of the bilateral exchange rate models (figures C.14-C.16 and D.14-D.16 in the appendix) with the exception that the direction accuracy gains over the expanded sample are now smaller, though still remain positive.

Figure 12: **Forecasting bias vs. error magnitude accuracy (sample comparison).**

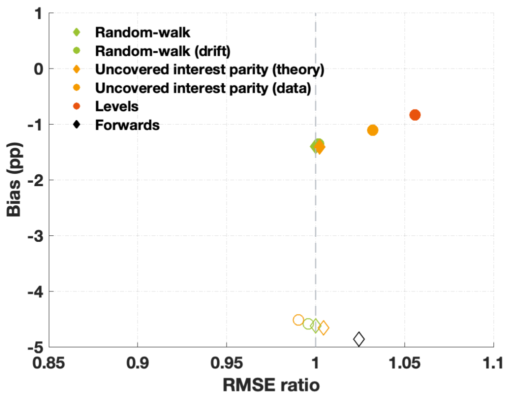
(a) Simple models, one-month forecasting horizon.



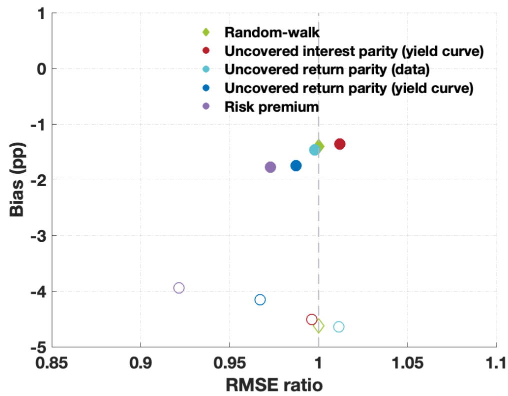
(b) Complex models, one-month forecasting horizon.



(c) Simple models, one-year forecasting horizon.



(d) Complex models, one-year forecasting horizon.



Note: This figure reports the performance of various exchange rate models relative to a random-walk over different sample periods. Each dot reports the mean error (bias) statistic (y-axis) and the ratio of the model's root mean squared forecast error relative to a random-walk (x-axis). The solid dots represent the statistic values computed over the *before sample* (model forecasting performance evaluation starting in 2004, before the beginning of the Global financial crisis). Clear dots represent the statistic pair computed over the *after sample* (model forecasting performance evaluation starting in 2014, after the end of the European sovereign debt crisis).

The reliance on the training period sample observed in the forecast evaluation exercise is also noticeable in the models' parameter estimates. They turn out to be volatile and unstable throughout the sample.[31] In the case of the uncovered interest parity, estimates feature the wrong signs and magnitudes (in line with the UIP puzzle), that with time turn out to be statistically indistinguishable from zero. A similar pattern with quite volatile parameter estimates can be seen as well throughout other inspected models.[32]

Finally, even though simple and calibrated models are not strongly inferior to the more complex ones (that are coming with significantly more instability and sample dependence), they explain very little of the observed in-sample variation. On the other hand, more complex models like the risk premium model could pick up some of the krona variation in-sample, beyond what could be attributable to pure interest differential. With the

[31] One potential theory supporting such a finding is a "scapegoat theory", stating that market participants being unsure of a true model driving the exchange rate movements (and/or having ever-changing views of a true model determining the currency movements) are frequently shifting focus between different economic fundamentals (see Bacchetta and Van Wincoop (2004)), which are being picked as "scapegoats", to rationalize the observed currency fluctuations at times when exchange rates are driven by unobservable shocks.

[32] Another notable peculiarity common to recursive coefficient estimates coming from these models is the level shift around the global financial crisis period suggesting a potential nonlinearity that has not been fully captured with the explanatory factors used in respective models.

added interpretation of the yield curve factors summarizing the market expectations of future fundamentals it suggests that excess currency returns are not perfectly orthogonal to fundamentals, at least not ex-post, and can therefore offer some insights into the fundamental drivers of krona movements.

5 Conclusion

This paper evaluates the forecasting performance in the short run (up to one year ahead) of many commonly used models for the Swedish krona exchange rate(s). According to different criteria, many models have difficulty in significantly differentiating themselves from the random-walk prediction. The very few models that under certain performance metrics can deliver a slight advantage over the random-walk projection typically come with a higher complexity levels and with a performance that is highly dependent on the training sample and the perfect forecastability of the explanatory factors. In addition, the outperformance is not uniform across different evaluation criteria.

As the analysis performed throughout this paper suggests, simple and calibrated models do not have significantly inferior out-of-sample predictive ability with respect to the random-walk, though they can only explain little of the observed in-sample variation. To the extent that they are not significantly different from the random-walk prediction, they have an advantage of not having an increased complexity of many (potentially unstable) parameters to be estimated. On the other hand, complex models could explain quite some of the krona in-sample variation, beyond what could be attributable to pure interest differential. Even though the risk premium model does not uniformly succeed in outperforming the predictive ability of the random-walk, it offers some ex-post insights into (fundamental) drivers of the observed krona movements. Nevertheless, the standard finding in the literature, corroborated as well here, is that the random-walk remains a benchmark whose forecasting performance is difficult to beat, at least over a short horizon. Moreover, the analysis of forecasting performance over different samples suggests that, at moments when additional information is available, it should be exploited by the forecaster, as indicated by the risk premium model's outperformance during crisis periods with perfect foresight of financial factors.[33] However, switching towards longer horizons, beyond a year and further, some more fundamentals driven models might be taking over the forecasting performance of the random-walk, but long horizon performance falls beyond the scope of this paper.[34]

[33] Private information available to the central banker setting the policy rate and having the advantage of knowing when to incorporate the policy rate announcement effects, at least in terms of direction if not magnitude, is one such example of a proprietary information.

[34] Askestad et al. (2019) offer a summary of the forecasting performance evaluation of a few commonly used krona models across both the shorter and longer time horizon.

List of references

Askestad, Emil, Ana Maria Ceh, Paola Di Casola and Annukka Ristiniemi. 2019. "Forecasting the krona." *Sveriges Riksbank: Economic Commentaries* (12): 1–15.

Bacchetta, Philippe and Eric Van Wincoop. 2004. "A Scapegoat Model of Exchange-Rate Fluctuations." *American Economic Review* 94 (2): 114–118.

Cappiello, Lorenzo and Roberto A. De Santis. 2007. "The uncovered return parity condition." Working Paper Series 812, European Central Bank.

Cerutti, Eugenio M, Maurice Obstfeld and Haonan Zhou. 2019. "Covered Interest Parity Deviations: Macrofinancial Determinants." Working Paper 26129, National Bureau of Economic Research.

Chen, Yu-Chin and Kwok Ping Tsang. 2013. "What Does the Yield Curve Tell Us about Exchange Rate Predictability?" *The Review of Economics and Statistics* 95 (1): 185–205.

Cheung, Yin-Wong, Menzie D. Chinn, Antonio Garcia Pascual and Yi Zhang. 2019. "Exchange rate prediction redux: New models, new data, new currencies." *Journal of International Money and Finance* 95 332–362.

Diebold, Francis X. 2015. "Comparing Predictive Accuracy, Twenty Years Later: A Personal Perspective on the Use and Abuse of Diebold–Mariano Tests." *Journal of Business & Economic Statistics* 33 (1): 1–9.

Diebold, Francis X and Robert S Mariano. 2002. "Comparing Predictive Accuracy." *Journal of Business & Economic Statistics* 20 (1): 134–144.

Djeutem, Edouard and Geoffrey R. Dunbar. 2018. "Uncovered Return Parity: Equity Returns and Currency Returns." Staff Working Papers 18-22, Bank of Canada.

Engel, Charles. 2014. "Chapter 8 - Exchange Rates and Interest Parity." In *Handbook of International Economics*, edited by Gopinath, Gita, Elhanan Helpman and Kenneth Rogoff Volume 4. 453 – 522, Elsevier.

Engel, Charles, Nelson C. Mark, Kenneth D. West, Kenneth Rogoff and Barbara Rossi. 2007. "Exchange Rate Models Are Not as Bad as You Think [with Comments and Discussion]." *NBER Macroeconomics Annual* 22 (2): 381–473.

Fama, Eugene F. 1984. "Forward and spot exchange rates." *Journal of Monetary Economics* 14 (3): 319–338.

Harvey, David, Stephen Leybourne and Paul Newbold. 1997. "Testing the equality of prediction mean squared errors." *International Journal of Forecasting* 13 (2): 281–291.

Hau, Harald and Hélène Rey. 2005. "Exchange Rates, Equity Prices, and Capital Flows." *The Review of Financial Studies* 19 (1): 273–317.

Lilley, Andrew, Matteo Maggiori, Brent Neiman and Jesse Schreger. 2019. "Exchange Rate Reconnect." *National Bureau of Economic Research, Working Paper Series*.

Meese, Richard A. and Kenneth Rogoff. 1983. "Empirical exchange rate models of the seventies: Do they fit out of sample?" *Journal of International Economics* 14 (1): 3–24.

Moosa, Imad A. and Kelly Burns. 2015. *Demystifying the Meese-Rogoff Puzzle*. Palgrave Macmillan UK, , 1st edition.

Obstfeld, Maurice and Kenneth Rogoff. 2000. "The Six Major Puzzles in International Macroeconomics: Is There a Common Cause?" *NBER Macroeconomics Annual* 15 (2): 339–390.

Rossi, Barbara. 2013. "Exchange Rate Predictability." *Journal of Economic Literature* 51 (4): 1063 – 1119.

Appendix A – List of models

In search for "the model" to deliver the best forecast and potential explanatory factor candidates we were inspired by the research and reasoning outlined in section 2. Therefore, those models and their variations have been included in the selection of models whose forecasting performance we are evaluating here. The following set of models is estimated and evaluated throughout the paper

(a) Random walk, equation (1),

$$\mathbb{E}_t[\Delta s_{t+h+1}] = 0,$$

(b) Random walk with drift, equation (2),

$$\mathbb{E}_t[\Delta s_{t+h+1}] = \beta_0,$$

(c) Level model, equation (3),

$$\mathbb{E}_t[s_{t+h+1}] = \beta_0 + \beta_1 i_{t+h} - \beta_2 i^*_{t+h},$$

(d) Uncovered interest parity, equation (4),

$$\mathbb{E}_t[\Delta s_{t+h+1}] = i_{t+h} - i^*_{t+h},$$

(e) Currency forwards (covered interest parity), equation (5),

$$\mathbb{E}_t[s_{t+h+1}] = f_{t+h+1}(\equiv s_t + i_t^{h+1} - i_t^{*h+1}),$$

(f) Empirical uncovered interest parity, equation (6),

$$\Delta s_{t+h+1} = \beta_0 + \beta_1(i_{t+h} - i^*_{t+h}) + e_{t+h+1},$$

(g) Uncovered return parity, equation (7),

$$rx_{t+h+1} = \delta_0 + \delta_1 oil_{t+h} + \delta_2 vol_{t+h} + \gamma_0 ex_{t+h+1} + \gamma_1 ep_{t+h+1} + \gamma_1^* ep^*_{t+h+1} + e_{t+h+1},$$

(h) Yield curve factors augmented uncovered interest parity, equation (8),

$$\Delta s_{t+h+1} = \beta_0 + \beta_1 L^R_{t+h} + \beta_2 S^R_{t+h} + \beta_3 C^R_{t+h} + e_{t+h+1},$$

(i) Yield curve factors augmented uncovered return parity, equation (9),

$$\begin{aligned} rx_{t+h+1} = {} & \delta_0 + \delta_1 oil_{t+h} + \delta_2 vol_{t+h} + \gamma_0 ex_{t+h+1} + \gamma_1 ep_{t+h+1} + \dots \\ & + \gamma_1^* ep^*_{t+h+1} + \beta_1 L^R_{t+h} + \beta_2 S^R_{t+h} + \beta_3 C^R_{t+h} + e_{t+h+1}. \end{aligned}$$

(j) Risk premium model, equation (10),

$$\begin{aligned} rx_{t+h+1} = {} & \delta_0 + \delta_1 oil_{t+h} + \delta_2 vol_{t+h} + \gamma_0 r^*_{t+h+1} + \gamma_1(ep_{t+h+1} - ep^*_{t+h+1}) + \dots \\ & + \beta_1 L_{t+h} + \beta_2 S_{t+h} + \beta_3 C_{t+h} + \beta_1^* L^*_{t+h} + \beta_2^* S^*_{t+h} + \beta_3^* C^*_{t+h} + e_{t+h+1}. \end{aligned}$$

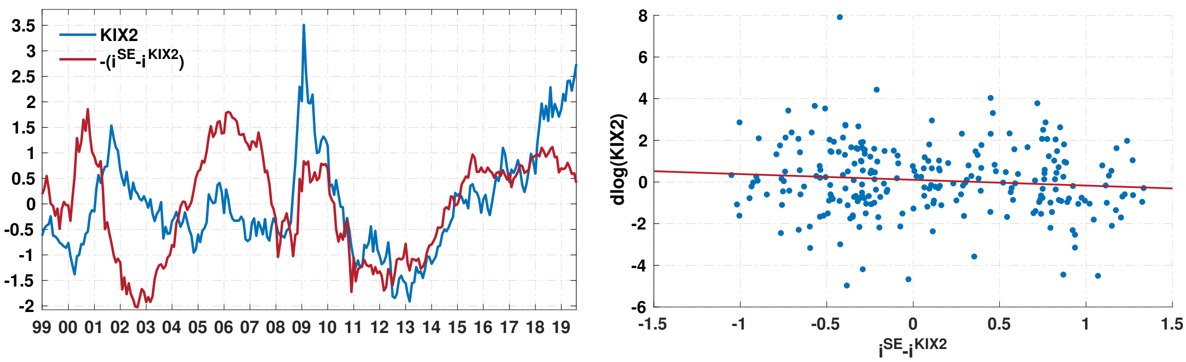
Appendix B – KIX2

Two stylized krona facts

A stylized fact for the krona, as well as many other currencies, is the fact that the uncovered interest parity condition does not hold empirically. Not only the correlation between the depreciation rate of currency and interest differentials is far from the theoretical but it as well turns out to be of the wrong sign, meaning negative. The figure B.1b documents this fact for the narrow krona effective exchange rate, KIX2, while similar holds for the bilateral exchange rates of euro and dollar against krona, as is reported in the appendix. The figure B.1b indicates a weak relationship and implies that the standard uncovered interest parity relationship seem to fail in data. On the other hand, the figure B.1a seems to emphasize the (strong) krona level connection to interest differentials, as the *model* (3) actually proposes, though it suggests that this correlation changed with the Global financial crisis.

Figure B.1: **Krona exchange rate.**

(a) **Krona level and interest differential**. (b) **Uncovered Interest Parity**.



Note: The LHS figure plots krona exchange rate level next to interest differential as standardized timeseries. The RHS figure shows scatter plots of the krona exchange rate (KIX2) depreciation versus the corresponding interest differentials.

Additionally, the table B.1 reports the equity price movements displaying much stronger correlation to krona depreciation than do interest rates themselves, suggesting that the equity returns might have better success in explaining the krona movements than does the interest parity relationship.

Table B.1: **Depreciation correlation.** The table reports the correlation between the krona exchange rate(s) depreciation (measured as log-difference) and different risk-free (small) and risky asset returns and premiums (large).

	Risk-free asset return			Risky asset return			Risk premium
	i	i^*	$i-i^*$	r	r^*	$r-r^*$	r^*-i^*
$\Delta \log EURSEK$	0.04	0.05	−0.04	−0.26	−0.36	0.04	−0.32
$\Delta \log USDSEK$	−0.02	0.04	−0.07	−0.17	−0.41	0.24	−0.41
$\Delta \log KIX2$	0.03	0.04	−0.05	−0.26	−0.41	0.11	−0.37

Out-Of-Sample Forecasting Performance

Before sample: start evaluation process before the Global financial crisis

Table B.2: **Out-Of-Sample Forecasting Performance** of various krona exchange rate models, KIX2.

	RMSE ratio			
	1 month	3 months	6 months	12 months
	Benchmark			
Random-walk	1.00	1.00	1.00	1.00
	Simple models			
Random walk (drift)	1.01	1.00	1.00	1.00
Uncovered interest parity (theory)	1.00	1.00	1.00	1.00
Uncovered interest parity (data)	1.04	1.03	1.03	1.03
Levels	3.58	1.97	1.40	1.06
Forwards	65.15	1.00	0.84	0.72
	Complex models			
Uncovered interest parity (yield curve)	0.98	0.97	1.00	1.01
Uncovered return parity (data)	0.93	0.95	0.98	1.00
Uncovered return parity (yield curve)	0.97	0.90	0.95	0.99
Risk premium	0.89	0.89	0.94	0.97

Note: This table reports the ratio of model's root mean square error (RMSE) relative to a random-walk, over different forecasting horizons.

	Direction accuracy			
	1 month	3 months	6 months	12 months
	Benchmark			
Random walk	0.00	0.00	0.00	0.00
	Simple models			
Random walk (drift)	0.50	0.50	0.50	0.46
Uncovered interest parity (theory)	0.51	0.48	0.46	0.38
Uncovered interest parity (data)	0.43	0.45	0.42	0.40
Levels	0.51	0.48	0.51	0.53
Forwards	0.47	0.50	0.36	0.23
	Complex models			
Uncovered interest parity (yield curve)	0.50	0.53	0.47	0.48
Uncovered return parity (data)	0.55	0.56	0.54	0.52
Uncovered return parity (yield curve)	0.55	0.57	0.53	0.49
Risk premium	0.55	0.62	0.58	0.54

Note: This table reports the percentage of model's forecasts that correctly predicted the direction of change (direction accuracy), over different forecasting horizons.

	p-value (Diebold-Mariano)			
	1 month	3 months	6 months	12 months
	Benchmark			
Random walk	0.00	0.00	0.00	0.00
	Simple models			
Random walk (drift)	0.53	0.80	0.33	0.26
Uncovered interest parity (theory)	0.18	0.59	0.16	0.02
Uncovered interest parity (data)	0.41	0.09	0.08	0.23
Levels	0.00	0.04	0.11	0.63
Forwards	0.00	0.49	0.97	0.35
	Complex models			
Uncovered interest parity (yield curve)	0.10	0.26	0.29	0.76
Uncovered return parity (data)	0.14	0.22	0.31	0.48
Uncovered return parity (yield curve)	0.37	0.25	0.22	0.63
Risk premium	0.12	0.24	0.27	0.12

Note: This table reports the p-values of Diebold-Mariano test for the model's performance relative to random-walk, over different forecasting horizons.

	Bias			
	1 month	3 months	6 months	12 months
	Benchmark			
Random walk	−0.10	−0.33	−0.68	−1.40
	Simple models			
Random walk (drift)	−0.03	−0.27	−0.62	−1.36
Uncovered interest parity (theory)	−0.10	−0.34	−0.69	−1.41
Uncovered interest parity (data)	0.11	−0.10	−0.43	−1.11
Levels	−1.30	−1.26	−1.14	−0.83
Forwards	−118.54	−0.34	0.05	−2.57
	Complex models			
Uncovered interest parity (yield curve)	0.03	−0.24	−0.64	−1.35
Uncovered return parity (data)	−0.12	−0.42	−0.76	−1.46
Uncovered return parity (yield curve)	−0.06	−0.38	−0.86	−1.74
Risk premium	−0.18	−0.46	−0.89	−1.77

Note: This table reports the model's mean error (bias), over different forecasting horizons.

Note: The table shows various performance criteria and statistics at different forecasting horizons for the narrow krona effective exchange rate, KIX2. All models are estimated on $1999:1 - 2003:12$ sample and subsequently re-estimated by sequentially adding one new (end-of-month) observation until $2019:12$. Each model estimate is used to produce up to 12-month ahead forecast and the forecasting errors across different horizons are used to produce the relevant statistics.

Out-Of-Sample Forecasting Performance

After sample: start evaluation process after the European sovereign debt crisis

Table B.3: **Out-Of-Sample Forecasting Performance** of various krona exchange rate models.

	RMSE ratio			
	1 month	3 months	6 months	12 months
	Benchmark			
Random walk	1.00	1.00	1.00	1.00
	Simple models			
Random walk (drift)	1.00	1.00	1.00	1.00
Uncovered interest parity (theory)	1.00	1.00	1.00	1.00
Uncovered interest parity (data)	1.00	1.00	0.99	0.99
Levels	4.79	2.83	2.20	1.68
Forwards	78.14	1.00	1.01	1.02
	Complex models			
Uncovered interest parity (yield curve)	1.04	1.03	1.01	1.00
Uncovered return parity (data)	1.05	1.05	1.02	1.01
Uncovered return parity (yield curve)	1.07	1.03	0.98	0.97
Risk premium	1.02	0.96	0.92	0.92

Note: This table reports the ratio of models' root mean squared error (RMSE) relative to a random-walk, over different forecasting horizons.

	Direction accuracy			
	1 month	3 months	6 months	12 months
	Benchmark			
Random walk	0.00	0.00	0.00	0.00
	Simple models			
Random walk (drift)	0.48	0.48	0.52	0.67
Uncovered interest parity (theory)	0.54	0.48	0.39	0.20
Uncovered interest parity (data)	0.52	0.62	0.68	0.82
Levels	0.48	0.39	0.35	0.25
Forwards	0.42	0.49	0.45	0.33
	Complex models			
Uncovered interest parity (yield curve)	0.51	0.42	0.47	0.48
Uncovered return parity (data)	0.52	0.52	0.56	0.50
Uncovered return parity (yield curve)	0.56	0.52	0.61	0.62
Risk premium	0.66	0.71	0.70	0.67

Note: This table reports the percentage of model's forecasts that correctly projected the direction of change (direction accuracy), over different forecasting horizons.

	p-value (Diebold-Mariano)			
	1 month	3 months	6 months	12 months
	Benchmark			
Random walk	0.00	0.00	0.00	0.00
	Simple models			
Random walk (drift)	0.30	0.32	0.29	0.53
Uncovered interest parity (theory)	0.12	0.19	0.17	0.13
Uncovered interest parity (data)	0.52	0.25	0.06	0.01
Levels	0.00	0.03	0.12	0.22
Forwards	0.00	0.26	0.26	0.62
	Complex models			
Uncovered interest parity (yield curve)	0.53	0.32	0.49	0.56
Uncovered return parity (data)	0.23	0.05	0.14	0.38
Uncovered return parity (yield curve)	0.09	0.30	0.38	0.03
Risk premium	0.39	0.08	0.07	0.05

Note: This table reports the p-values of a Diebold-Mariano test for the model's performance relative to random-walk, over different forecasting horizons.

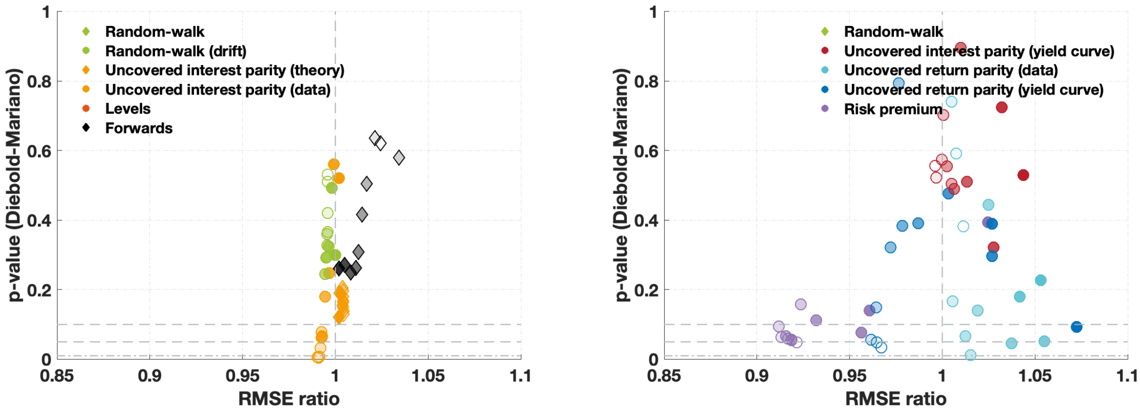
	Bias			
	1 month	3 months	6 months	12 months
	Benchmark			
Random walk	−0.36	−1.19	−2.40	−4.62
	Simple models			
Random walk (drift)	−0.30	−1.13	−2.36	−4.58
Uncovered interest parity (theory)	−0.38	−1.21	−2.43	−4.65
Uncovered interest parity (data)	−0.27	−1.10	−2.31	−4.52
Levels	−6.16	−6.79	−7.68	−9.29
Forwards	−123.27	−1.21	−2.51	−4.86
	Complex models			
Uncovered interest parity (yield curve)	−0.27	−1.09	−2.30	−4.50
Uncovered return parity (data)	−0.30	−1.13	−2.35	−4.63
Uncovered return parity (yield curve)	0.04	−0.76	−1.93	−4.15
Risk premium	0.23	−0.56	−1.73	−3.94

Note: This table reports the model's mean error (bias), over different forecasting horizons.

Note: The table shows various performance criteria and statistics at different forecasting horizons for the krona trade weighted exchange rate, KIX2. All models are estimated on $1999:1 - 2013:12$ sample and subsequently re-estimated by sequentially adding one new (end-of-month) observation until $2019:12$. Each model estimate is used to produce up to 12-month ahead forecast and the forecasting errors across different horizons are used to produce the relevant statistics.

Figure B.2: **Beating the benchmark vs. error magnitude accuracy.**

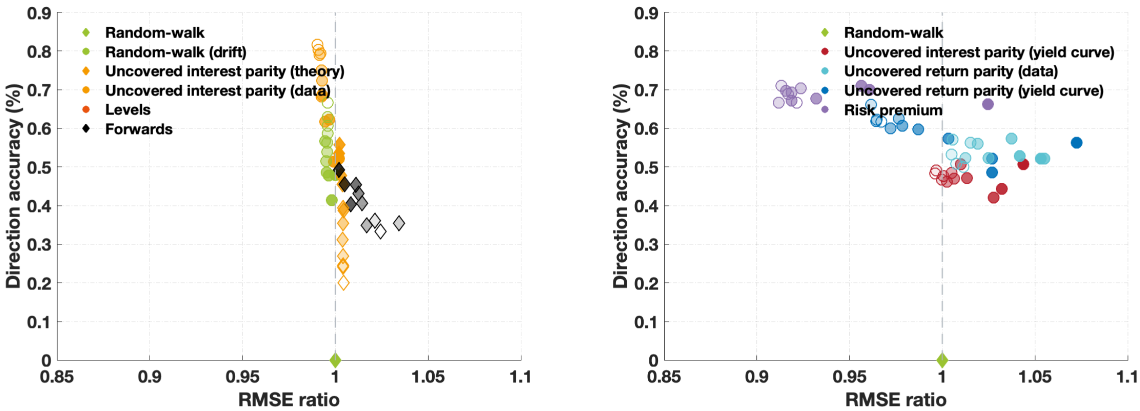
(a) Simple models. (b) Complex models.



Note: This figure reports the performance of various exchange rate models relative to a random-walk over different sample periods. Each dot reports the p-value of the Diebold-Mariano test statistic (y-axis) and the ratio of the model's root mean squared forecast error relative to a random-walk (x-axis). The solid dots represent the statistic values at a one-month horizon. Lighter shades represent the statistic pair value further ahead in time. Clear dots represent the statistic pair at a one-year horizon.

Figure B.3: **Direction vs. error magnitude accuracy.**

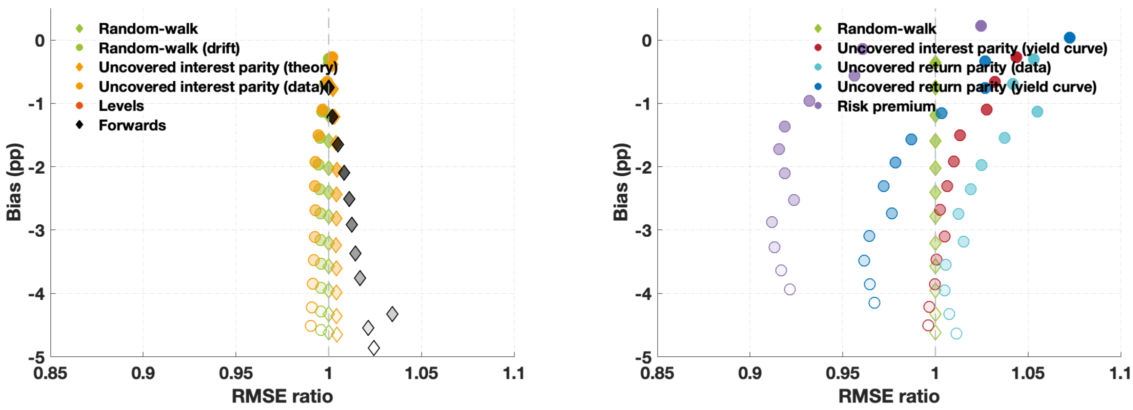
(a) Simple models. (b) Complex models.



Note: This figure reports the performance of various exchange rate models relative to a random-walk over different sample periods. Each dot reports the direction accuracy statistic (y-axis) and the ratio of the model's root mean squared forecast error relative to a random-walk (x-axis). The solid dots represent the statistic values at a one-month horizon. Lighter shades represent the statistic pair value further ahead in time. Clear dots represent the statistic pair at a one-year horizon.

Figure B.4: **Forecasting bias vs. error magnitude accuracy.**

(a) Simple models. (b) Complex models.

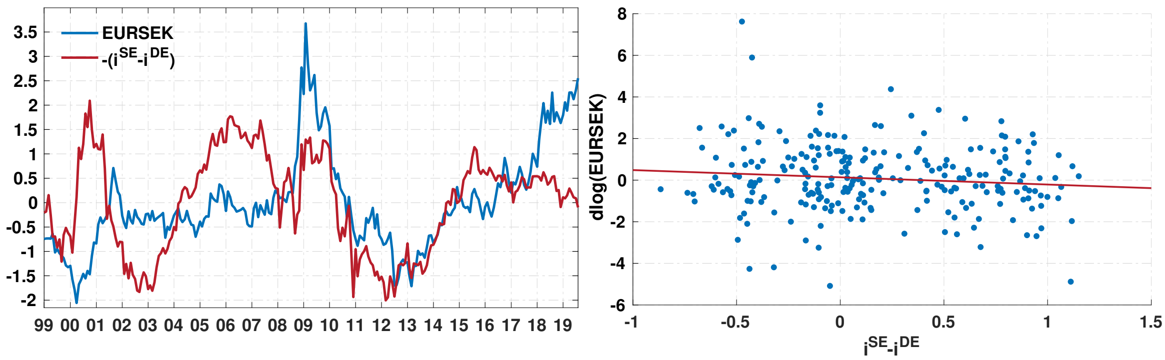


Note: This figure reports the performance of various exchange rate models relative to a random-walk over different sample periods. Each dot reports the mean error (bias) statistic (y-axis) and the ratio of the model's root mean squared forecast error relative to a random-walk (x-axis). The solid dots represent the statistic values at a one-month horizon. Lighter shades represent the statistic pair value further ahead in time. Clear dots represent the statistic pair at a one-year horizon.

Appendix C – EURSEK

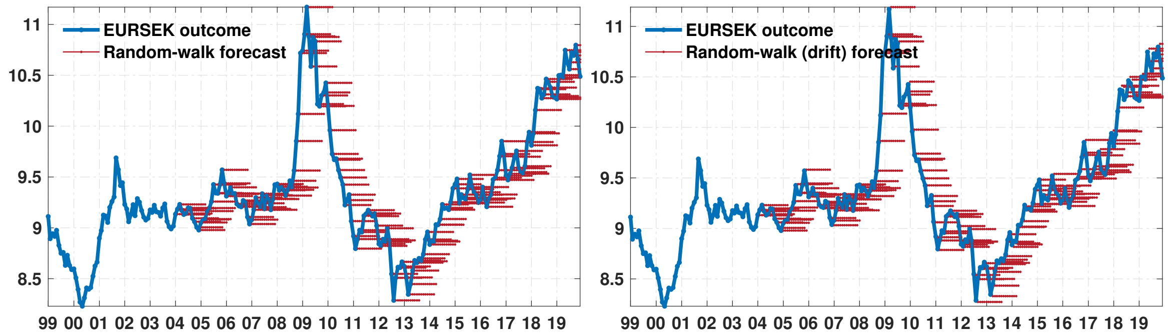
Figure C.1: **Krona exchange rate.**

(a) **Krona level and interest differential.** (b) **Uncovered Interest Parity.**



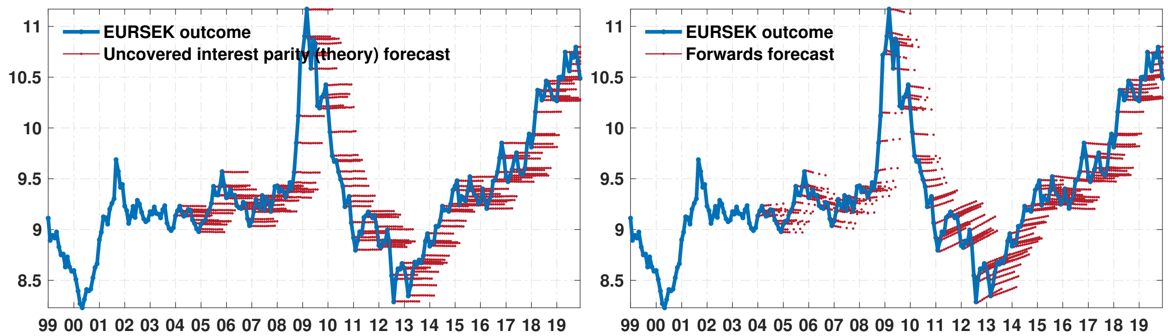
Note: The LHS figure plots EURSEK exchange rate level next to interest differential as standardized timeseries. The RHS figure shows scatter plots of the EURSEK exchange rate depreciation versus the corresponding interest differentials.

Figure C.2: **EURSEK random-walk forecasts.**



Note: This figure plots 12-months-ahead forecast for EURSEK coming from the random-walk model without drift (left) and the random walk model with drift (right).

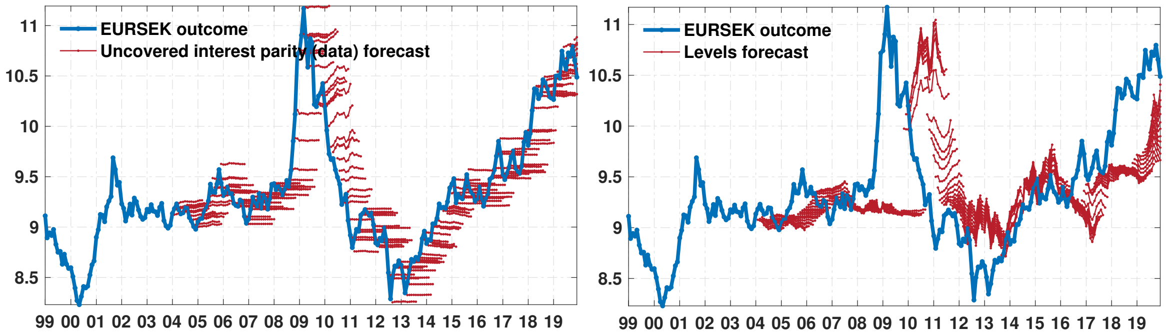
Figure C.3: **EURSEK interest parity forecasts.**



Note: This figure plots 12-months-ahead forecasts for EURSEK coming from the uncovered interest rate parity condition (left) and implied EURSEK forwards (right) corresponding to the covered interest parity condition.

Figure C.4: **Out-of-sample forecasts coming from empirically estimated interest parity models.**

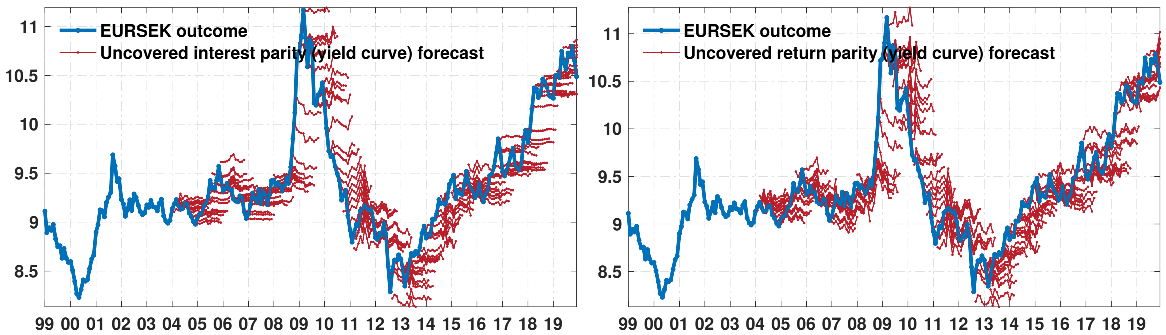
(a) Uncovered interest parity. (b) Krona level.



Note: This figure plots 12-months-ahead forecasts from set of different models. All models are estimated on a $1999:1-2003:12$ sample and subsequently re-estimated by sequentially adding one new (end-of-month) observation until $2019:12$. Each model estimate is used to produce forecasts up to 12 months ahead in time.

Figure C.5: **Out-of-sample forecasts coming from empirically estimated parity models augmented by yield curve factors.**

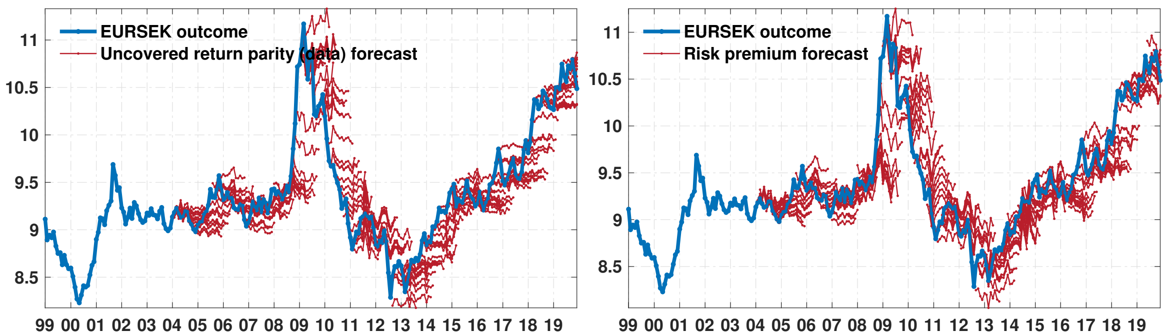
(a) Yield curve augmented uncovered interest parity. (b) Yield curve augmented uncovered return parity.



Note: This figure plots 12-months-ahead forecasts from set of different models. All models are estimated on a $1999:1-2003:12$ sample and subsequently re-estimated by sequentially adding one new (end-of-month) observation until $2019:12$. Each model estimate is used to produce forecasts up to 12 months ahead in time.

Figure C.6: **Out-of-sample forecasts coming from empirically estimated return parity models.**

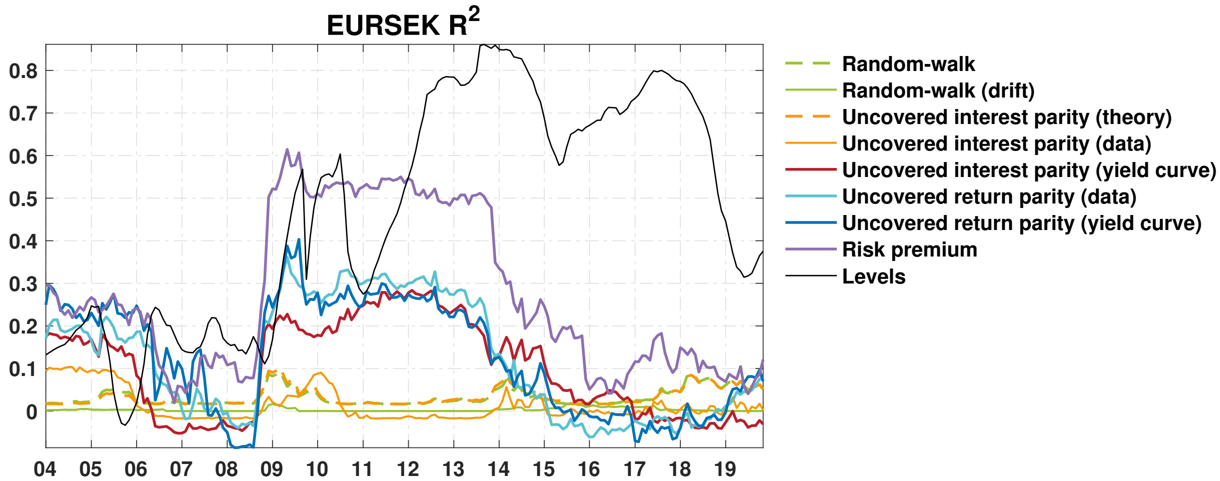
(a) Uncovered return parity. (b) Risk premium model.



Note: This figure plots 12-months-ahead forecasts from set of different models. All models are estimated on a $1999:1-2003:12$ sample and subsequently re-estimated by sequentially adding one new (end-of-month) observation until $2019:12$. Each model estimate is used to produce forecasts up to 12 months ahead in time.

In-Sample Forecasting Performance

Figure C.7: **Krona variation explained in-sample**. The percentage of EURSEK variation explained by the model determinants.



Note: The figure shows recursive window adjusted R^2 estimates across different models for EURSEK exchange rate. All models are estimated on a $1999:1-2003:12$ sample and subsequently re-estimated by sequentially adding one new (end-of-month) observation until $2019:12$. Each model estimate is used to produce a time-varying R^2 coefficient.

Out-Of-Sample Forecasting Performance

Before sample: start evaluation process before the Global financial crisis

Table C.1: **Out-Of-Sample Forecasting Performance** of various EURSEK exchange rate models.

	RMSE ratio			
	1 month	3 months	6 months	12 months
	Benchmark			
Random-walk	1.00	1.00	1.00	1.00
	Simple models			
Random-walk (drift)	1.01	1.00	1.00	1.00
Uncovered interest parity (theory)	1.00	1.00	1.00	1.00
Uncovered interest parity (data)	1.03	1.03	1.03	1.04
Levels	4.06	2.39	1.80	1.33
Forwards	305.36	0.99	0.89	0.71
	Complex models			
Uncovered interest parity (yield curve)	0.97	0.98	1.00	1.02
Uncovered return parity (data)	0.96	0.97	1.00	1.01
Uncovered return parity (yield curve)	0.99	0.95	0.97	0.98
Risk premium	0.88	0.89	0.93	0.96

Note: This table reports the ratio of models' root mean squared error (RMSE) relative to a random-walk, over different forecasting horizons.

	Direction accuracy			
	1 month	3 months	6 months	12 months
	Benchmark			
Random-walk	0.00	0.00	0.00	0.00
	Simple models			
Random-walk (drift)	0.47	0.49	0.50	0.53
Uncovered interest parity (theory)	0.50	0.48	0.49	0.44
Uncovered interest parity (data)	0.48	0.46	0.46	0.48
Levels	0.45	0.50	0.47	0.40
Forwards	0.50	0.57	0.45	0.29
	Complex models			
Uncovered interest parity (yield curve)	0.53	0.51	0.49	0.47
Uncovered return parity (data)	0.52	0.53	0.52	0.45
Uncovered return parity (yield curve)	0.55	0.56	0.56	0.50
Risk premium	0.58	0.61	0.63	0.61

Note: This table reports the percentage of model's forecasts that correctly projected the direction of change (direction accuracy), over different forecasting horizons.

	p-value (Diebold-Mariano)			
	1 month	3 months	6 months	12 months
	Benchmark			
Random-walk	0.00	0.00	0.00	0.00
	Simple models			
Random-walk (drift)	0.86	0.72	0.75	0.58
Uncovered interest parity (theory)	0.28	0.65	0.20	0.07
Uncovered interest parity (data)	0.72	0.46	0.36	0.32
Levels	0.00	0.04	0.06	0.08
Forwards	0.00	0.26	0.32	0.46
	Complex models			
Uncovered interest parity (yield curve)	0.12	0.27	0.55	0.72
Uncovered return parity (data)	0.17	0.17	0.45	0.99
Uncovered return parity (yield curve)	0.30	0.23	0.28	0.08
Risk premium	0.10	0.19	0.27	0.10

Note: This table reports the p-values of a Diebold-Mariano test for the model's performance relative to random-walk, over different forecasting horizons.

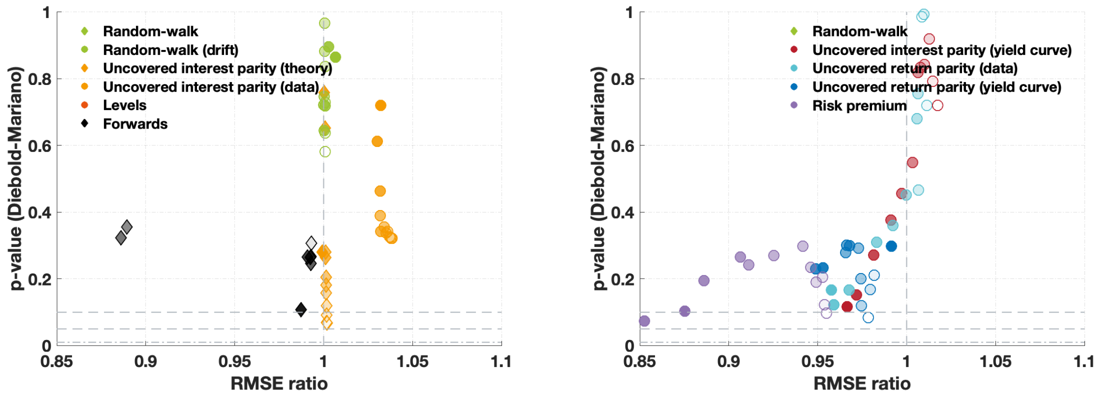
	Bias			
	1 month	3 months	6 months	12 months
	Benchmark			
Random-walk	−0.01	−0.02	−0.05	−0.10
	Simple models			
Random-walk (drift)	0.00	−0.02	−0.04	−0.09
Uncovered interest parity (theory)	−0.01	−0.02	−0.05	−0.10
Uncovered interest parity (data)	0.01	−0.01	−0.03	−0.08
Levels	−0.11	−0.11	−0.11	−0.11
Forwards	−10.66	−0.02	0.02	−0.14
	Complex models			
Uncovered interest parity (yield curve)	0.00	−0.02	−0.05	−0.09
Uncovered return parity (data)	−0.01	−0.03	−0.05	−0.10
Uncovered return parity (yield curve)	0.00	−0.02	−0.05	−0.10
Risk premium	0.00	−0.02	−0.05	−0.10

Note: This table reports the model's mean error (bias), over different forecasting horizons.

Note: The table shows various performance criteria and statistics at different forecasting horizons for the EURSEK exchange rate. All models are estimated on $1999:1 - 2003:12$ sample and subsequently re-estimated by sequentially adding one new (end-of-month) observation until $2019:12$. Each model estimate is used to produce up to 12-month ahead forecast and the forecasting errors across different horizons are used to produce the relevant statistics.

Figure C.8: **Beating the benchmark vs. error magnitude accuracy.**

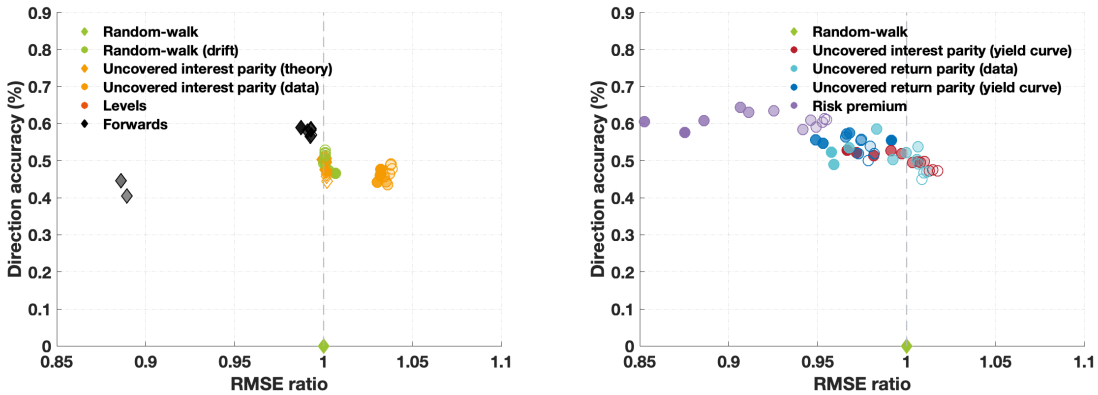
(a) Simple models. (b) Complex models.



Note: This figure reports the performance of various exchange rate models relative to a random-walk over different forecasting horizons. Each dot reports the p-value of the Diebold-Mariano test statistic (y-axis) and the ratio of the model's root mean squared forecast error relative to a random-walk (x-axis). The solid dots represent the statistic values at a one-month horizon. Lighter shades represent the statistic pair value further ahead in time. Clear dots represent the statistic pair at a one-year horizon.

Figure C.9: **Direction vs. error magnitude accuracy.**

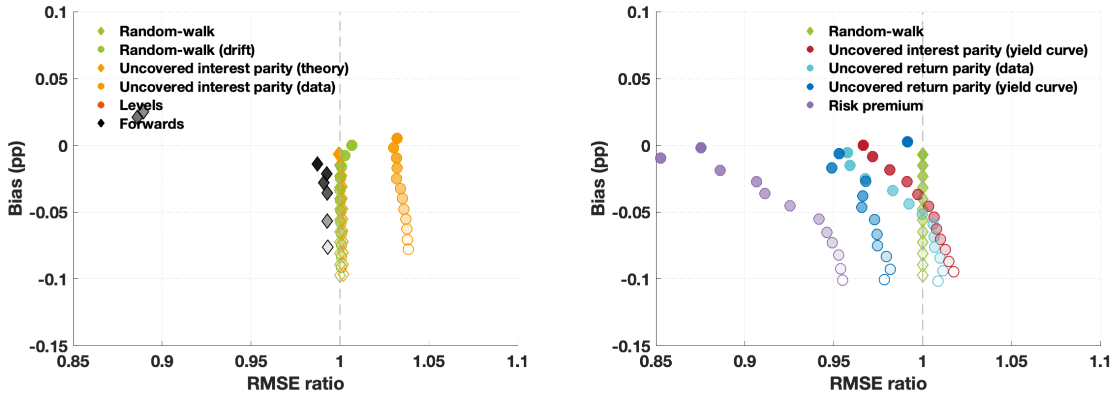
(a) Simple models. (b) Complex models.



Note: This figure reports the performance of various exchange rate models relative to a random-walk over different forecasting horizons. Each dot reports the direction accuracy statistic (y-axis) and the ratio of the model's root mean squared forecast error relative to a random-walk (x-axis). The solid dots represent the statistic values at a one-month horizon. Lighter shades represent the statistic pair value further ahead in time. Clear dots represent the statistic pair at a one-year horizon.

Figure C.10: **Forecasting bias vs. error magnitude accuracy.**

(a) Simple models. (b) Complex models.



Note: This figure reports the performance of various exchange rate models relative to a random-walk over different forecasting horizons. Each dot reports the mean error (bias) statistic (y-axis) and the ratio of the model's root mean squared forecast error relative to a random-walk (x-axis). The solid dots represent the statistic values at a one-month horizon. Lighter shades represent the statistic pair value further ahead in time. Clear dots represent the statistic pair at a one-year horizon.

Out-Of-Sample Forecasting Performance

After sample: start evaluation process after the European sovereign debt crisis

Table C.2: **Out-Of-Sample Forecasting Performance** of various EURSEK exchange rate models.

	RMSE ratio			
	1 month	3 months	6 months	12 months
	Benchmark			
Random-walk	1.00	1.00	1.00	1.00
	Simple models			
Random-walk (drift)	0.99	0.99	0.99	0.99
Uncovered interest parity (theory)	1.00	1.00	1.00	1.00
Uncovered interest parity (data)	1.00	0.99	0.99	0.99
Levels	5.14	3.10	2.47	1.87
Forwards	310.79	0.99	0.99	0.99
	Complex models			
Uncovered interest parity (yield curve)	1.02	1.02	1.00	0.99
Uncovered return parity (data)	1.03	1.04	1.01	1.01
Uncovered return parity (yield curve)	1.05	1.03	0.99	0.98
Risk premium	1.01	0.96	0.93	0.93

Note: This table reports the ratio of models' root mean squared error (RMSE) relative to a random-walk, over different forecasting horizons.

	Direction accuracy			
	1 month	3 months	6 months	12 months
	Benchmark			
Random-walk	0.00	0.00	0.00	0.00
	Simple models			
Random-walk (drift)	0.52	0.65	0.74	0.83
Uncovered interest parity (theory)	0.55	0.51	0.47	0.47
Uncovered interest parity (data)	0.51	0.61	0.74	0.85
Levels	0.48	0.42	0.33	0.25
Forwards	0.58	0.55	0.71	0.53
	Complex models			
Uncovered interest parity (yield curve)	0.54	0.46	0.58	0.60
Uncovered return parity (data)	0.51	0.48	0.48	0.50
Uncovered return parity (yield curve)	0.52	0.49	0.56	0.63
Risk premium	0.66	0.70	0.71	0.77

Note: This table reports the percentage of model's forecasts that correctly projected the direction of change (direction accuracy), over different forecasting horizons.

	p-value (Diebold-Mariano)			
	1 month	3 months	6 months	12 months
	Benchmark			
Random-walk	0.00	0.00	0.00	0.00
	Simple models			
Random-walk (drift)	0.26	0.17	0.13	0.14
Uncovered interest parity (theory)	0.45	0.66	0.57	0.50
Uncovered interest parity (data)	0.41	0.20	0.11	0.09
Levels	0.00	0.03	0.11	0.22
Forwards	0.00	0.07	0.06	0.48
	Complex models			
Uncovered interest parity (yield curve)	0.92	0.84	0.22	0.80
Uncovered return parity (data)	0.80	0.06	0.18	0.61
Uncovered return parity (yield curve)	0.21	0.18	0.99	0.17
Risk premium	0.53	0.19	0.13	0.13

Note: This table reports the p-values of a Diebold-Mariano test for the model's performance relative to random-walk, over different forecasting horizons.

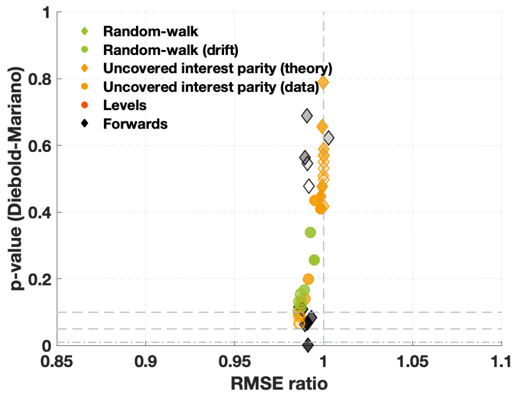
	Bias			
	1 month	3 months	6 months	12 months
	Benchmark			
Random-walk	−0.02	−0.08	−0.15	−0.30
	Simple models			
Random-walk (drift)	−0.02	−0.07	−0.15	−0.29
Uncovered interest parity (theory)	−0.02	−0.08	−0.15	−0.30
Uncovered interest parity (data)	−0.02	−0.07	−0.15	−0.29
Levels	−0.45	−0.48	−0.53	−0.62
Forwards	−14.47	−0.07	−0.15	−0.29
	Complex models			
Uncovered interest parity (yield curve)	−0.01	−0.06	−0.14	−0.28
Uncovered return parity (data)	−0.02	−0.07	−0.15	−0.30
Uncovered return parity (yield curve)	0.00	−0.05	−0.12	−0.27
Risk premium	0.02	−0.03	−0.10	−0.25

Note: This table reports the model's mean error (bias), over different forecasting horizons.

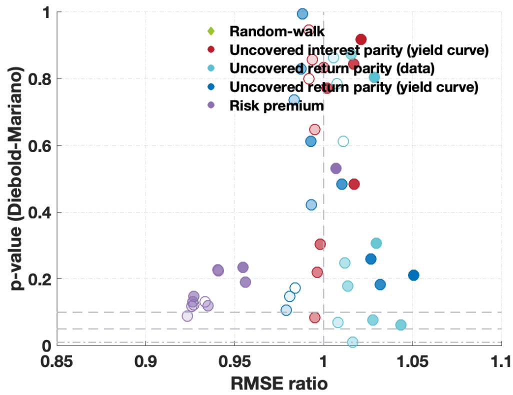
Note: The table shows various performance criteria and statistics at different forecasting horizons for the EURSEK exchange rate. All models are estimated on $1999:1 - 2013:12$ sample and subsequently re-estimated by sequentially adding one new (end-of-month) observation until $2019:12$. Each model estimate is used to produce up to 12-month ahead forecast and the forecasting errors across different horizons are used to produce the relevant statistics.

Figure C.11: **Beating the benchmark vs. error magnitude accuracy.**

(a) Simple models.



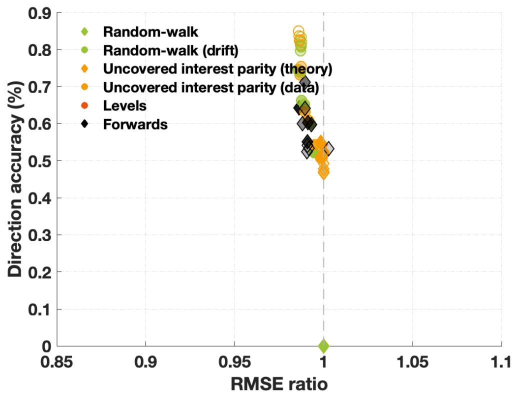
(b) Complex models.



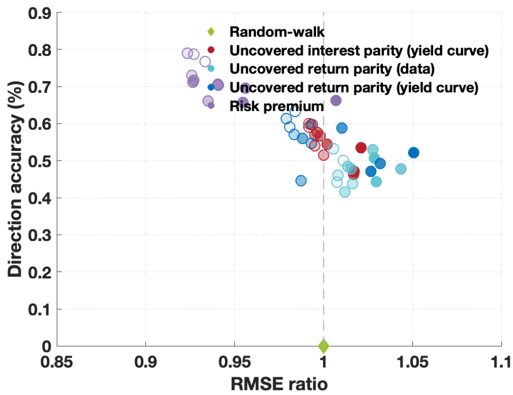
Note: This figure reports the performance of various exchange rate models relative to a random-walk over different forecasting horizons. Each dot reports the p-value of the Diebold-Mariano test statistic (y-axis) and the ratio of the model's root mean squared forecast error relative to a random-walk (x-axis). The solid dots represent the statistic values at a one-month horizon. Lighter shades represent the statistic pair value further ahead in time. Clear dots represent the statistic pair at a one-year horizon.

Figure C.12: **Direction vs. error magnitude accuracy.**

(a) Simple models.



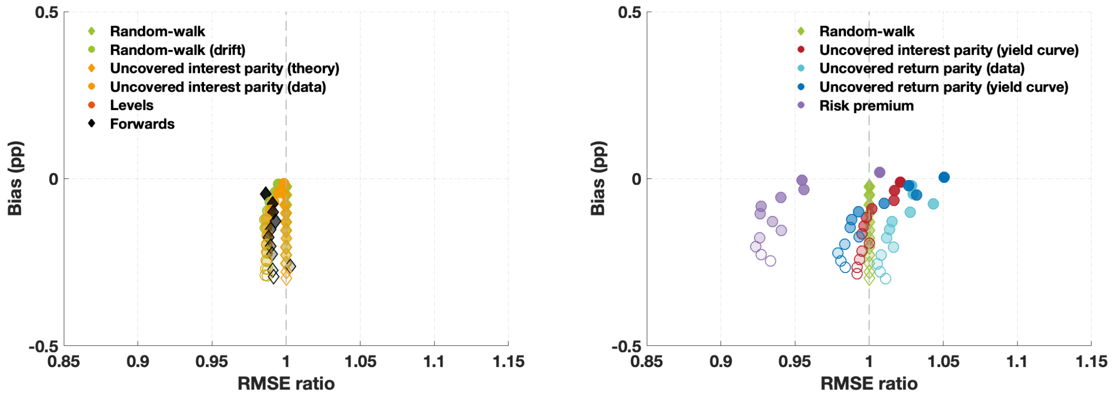
(b) Complex models.



Note: This figure reports the performance of various exchange rate models relative to a random-walk over different forecasting horizons. Each dot reports the direction accuracy statistic (y-axis) and the ratio of the model's root mean squared forecast error relative to a random-walk (x-axis). The solid dots represent the statistic values at a one-month horizon. Lighter shades represent the statistic pair value further ahead in time. Clear dots represent the statistic pair at a one-year horizon.

Figure C.13: **Forecasting bias vs. error magnitude accuracy.**

(a) Simple models. (b) Complex models.

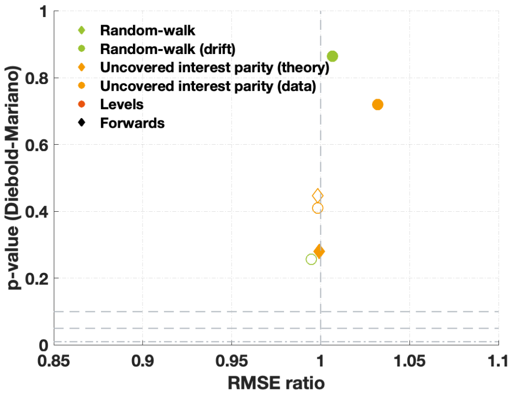


Note: This figure reports the performance of various exchange rate models relative to a random-walk over different forecasting horizons. Each dot reports the mean error (bias) statistic (y-axis) and the ratio of the model's root mean squared forecast error relative to a random-walk (x-axis). The solid dots represent the statistic values at a one-month horizon. Lighter shades represent the statistic pair value further ahead in time. Clear dots represent the statistic pair at a one-year horizon.

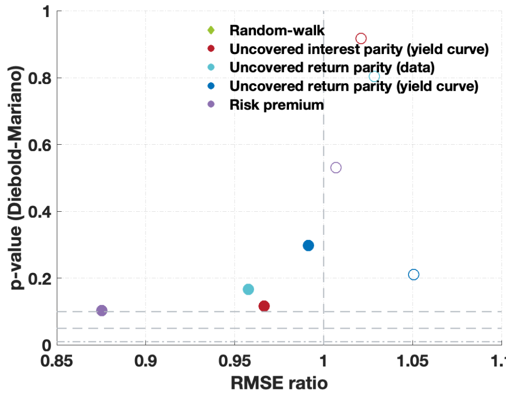
Out-Of-Sample Forecasting Performance Comparison: sample matters

Figure C.14: **Beating the benchmark vs. error magnitude accuracy (sample comparison).**

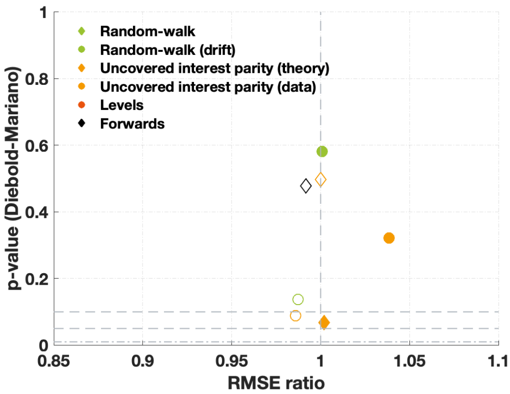
(a) Simple models, one-month forecasting horizon.



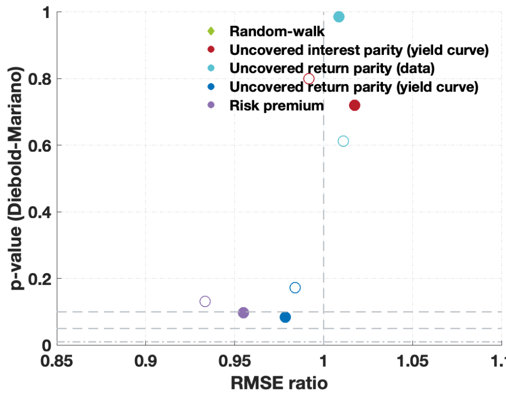
(b) Complex models, one-month forecasting horizon.



(c) Simple models, one-year forecasting horizon.



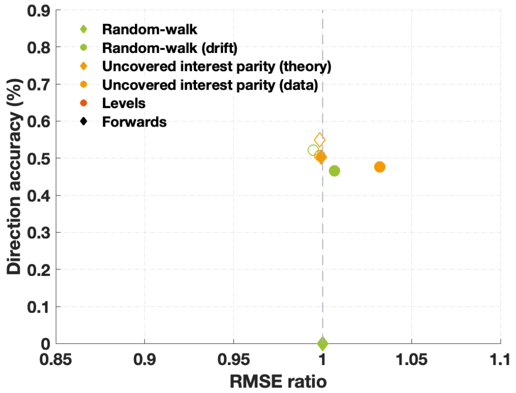
(d) Complex models, one-year forecasting horizon.



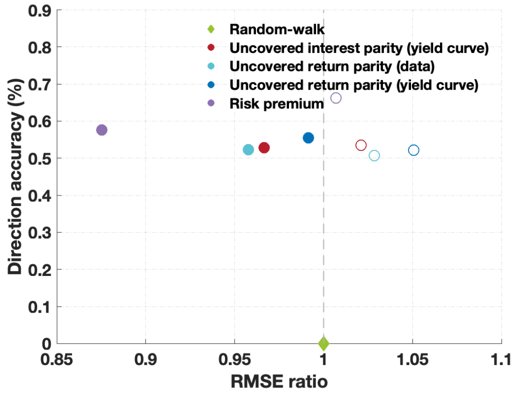
Note: This figure reports the performance of various exchange rate models relative to a random-walk over different sample periods. Each dot reports the p-value of the Diebold-Mariano test statistic (y-axis) and the ratio of the model's root mean squared forecast error relative to a random-walk (x-axis). The solid dots represent the statistic values computed over the *before sample* (models forecasting performance evaluation starting in 2004, before the beginning of the Global financial crisis). Clear dots represent the statistic pair computrd over the *after sample* (models forecasting performance evaluation starting in 2014, after the end of the European sovereign debt crisis).

Figure C.15: **Direction vs. error magnitude accuracy (sample comparison).**

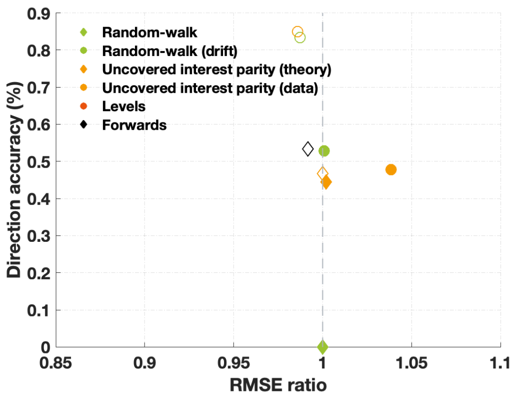
(a) Simple models, one-month forecasting horizon.



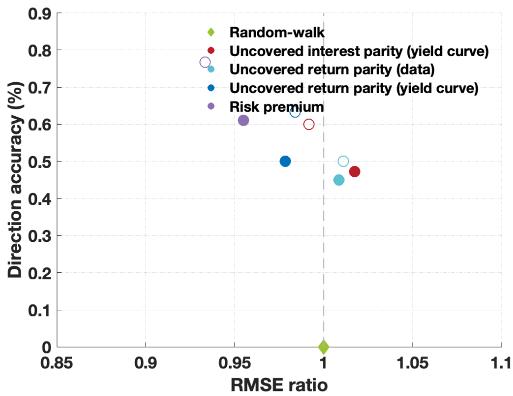
(b) Complex models, one-month forecasting horizon.



(c) Simple models, one-year forecasting horizon.



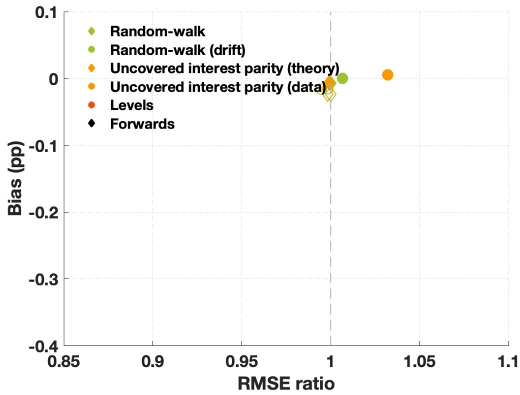
(d) Complex models, one-year forecasting horizon.



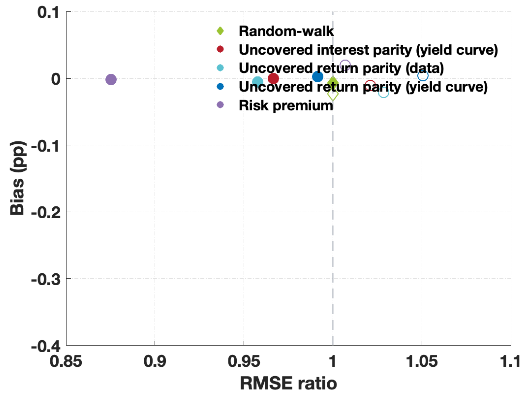
Note: This figure reports the performance of various exchange rate models relative to a random-walk over different sample periods. Each dot reports the direction accuracy statistic (y-axis) and the ratio of the model's root mean squared forecast error relative to a random-walk (x-axis). The solid dots represent the statistic values computed over the *before sample* (models forecasting performance evaluation starting in 2004, before the beginning of the Global financial crisis). Clear dots represent the statistic pair computrd over the *after sample* (models forecasting performance evaluation starting in 2014, after the end of the European sovereign debt crisis).

Figure C.16: **Forecasting bias vs. error magnitude accuracy (sample comparison).**

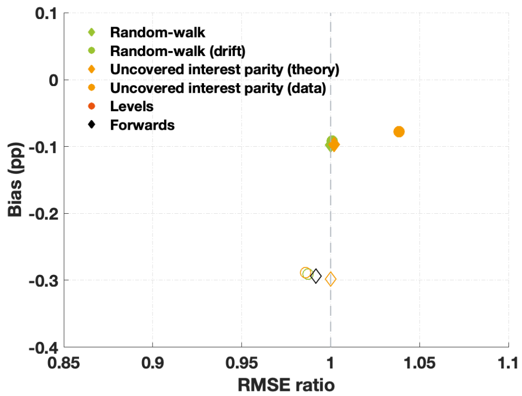
(a) Simple models, one-month forecasting horizon.



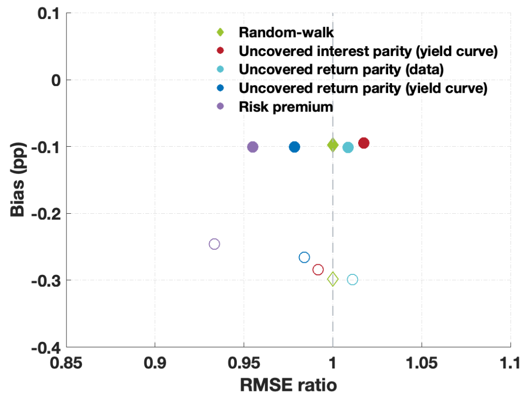
(b) Complex models, one-month forecasting horizon.



(c) Simple models, one-year forecasting horizon.



(d) Complex models, one-year forecasting horizon.

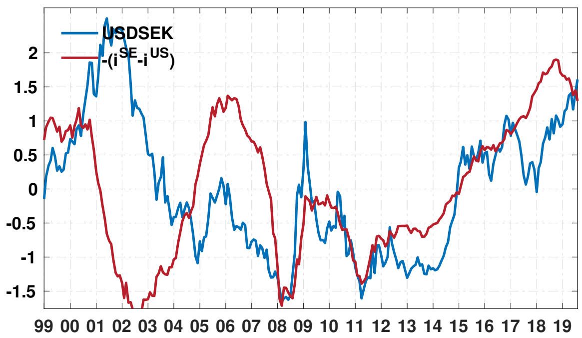


Note: This figure reports the performance of various exchange rate models relative to a random-walk over different sample periods. Each dot reports the mean error (bias) statistic (y-axis) and the ratio of the model's root mean squared forecast error relative to a random-walk (x-axis). The solid dots represent the statistic values computed over the *before sample* (models forecasting performance evaluation starting in 2004, before the beginning of the Global financial crisis). Clear dots represent the statistic pair computrd over the *after sample* (models forecasting performance evaluation starting in 2014, after the end of the European sovereign debt crisis).

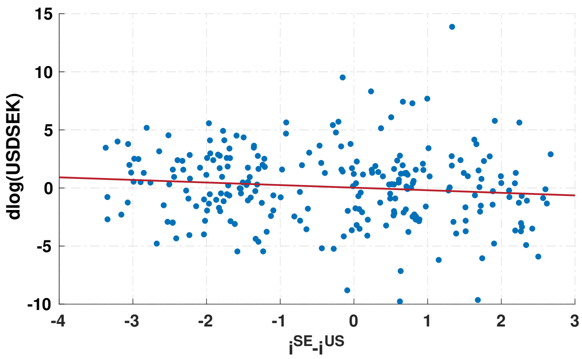
Appendix D – USDSEK

Figure D.1: **Krona exchange rate.**

(a) **Krona level and interest differential.**

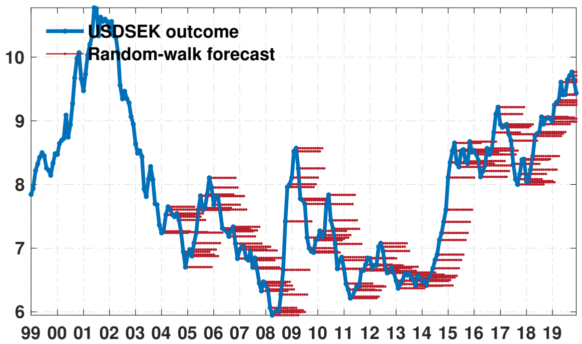


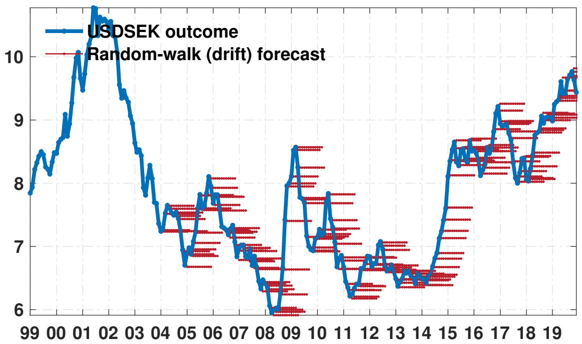
(b) **Uncovered Interest Parity.**



Note: The LHS figure plots USDSEK exchange rate level next to interest differential as standardized timeseries. The RHS figure shows scatter plots of the USDSEK exchange rate depreciation versus the corresponding interest differentials.

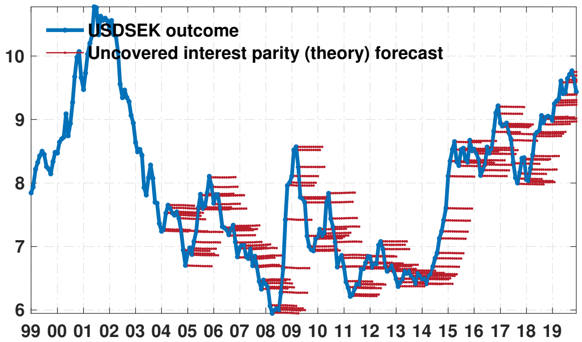
Figure D.2: **USDSEK random-walk forecasts.**

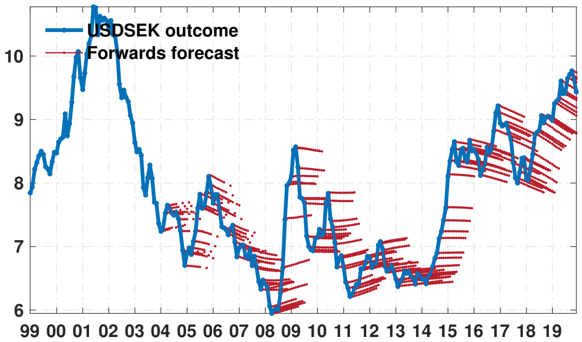




Note: This figure plots 12-months-ahead forecast for USDSEK coming from the random-walk model without drift (left) and the random walk model with drift (right).

Figure D.3: **USDSEK interest parity forecasts.**

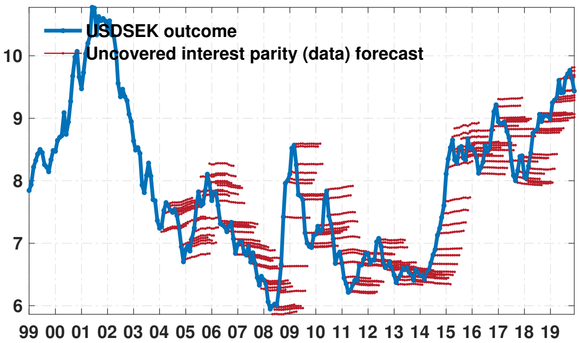




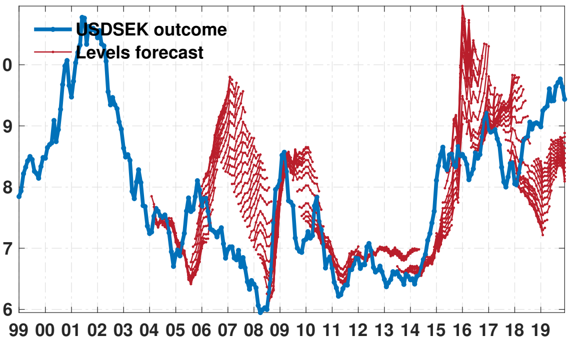
Note: This figure plots 12-months-ahead forecasts for USDSEK coming from the uncovered interest rate parity condition (left) and implied USDSEK forwards (right) corresponding to the covered interest parity condition.

Figure D.4: **Out-of-sample forecasts coming from empirically estimated interest parity models.**

(a) Uncovered interest parity.



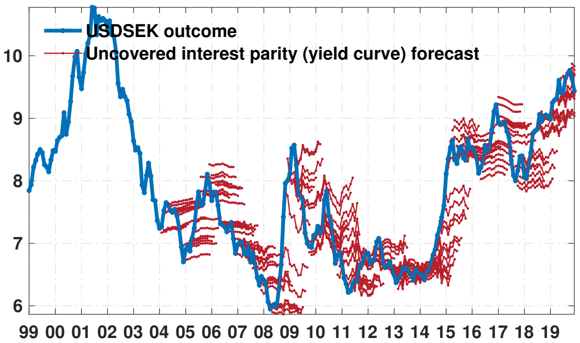
(b) Krona level.



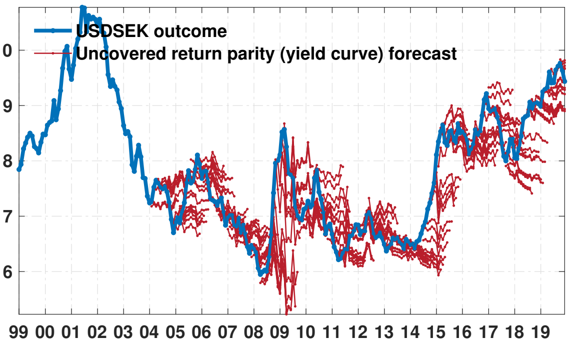
Note: This figure plots 12-months-ahead forecasts from set of different models. All models are estimated on a $1999:1-2003:12$ sample and subsequently re-estimated by sequentially adding one new (end-of-month) observation until $2019:12$. Each model estimate is used to produce forecasts up to 12 months ahead in time.

Figure D.5: **Out-of-sample forecasts coming from empirically estimated parity models augmented by yield curve factors.**

(a) Yield curve augmented uncovered interest parity.



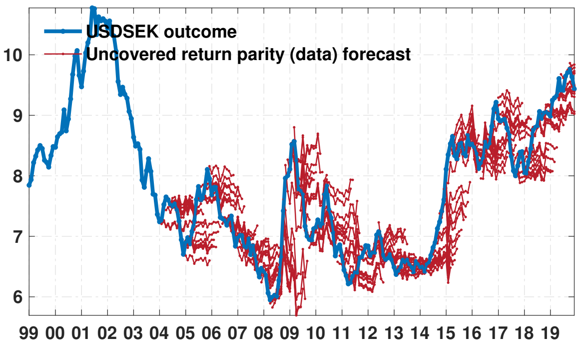
(b) Yield curve augmented uncovered return parity.



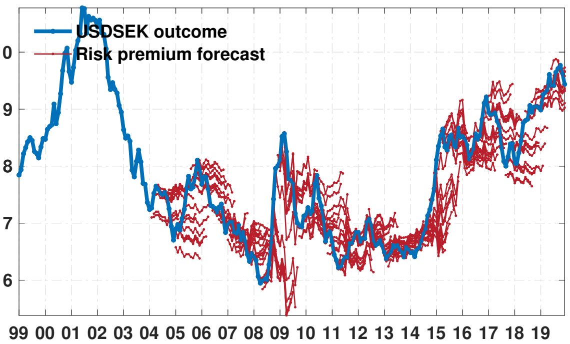
Note: This figure plots 12-months-ahead forecasts from set of different models. All models are estimated on a $1999:1-2003:12$ sample and subsequently re-estimated by sequentially adding one new (end-of-month) observation until $2019:12$. Each model estimate is used to produce forecasts up to 12 months ahead in time.

Figure D.6: **Out-of-sample forecasts coming from empirically estimated return parity models.**

(a) Uncovered return parity.



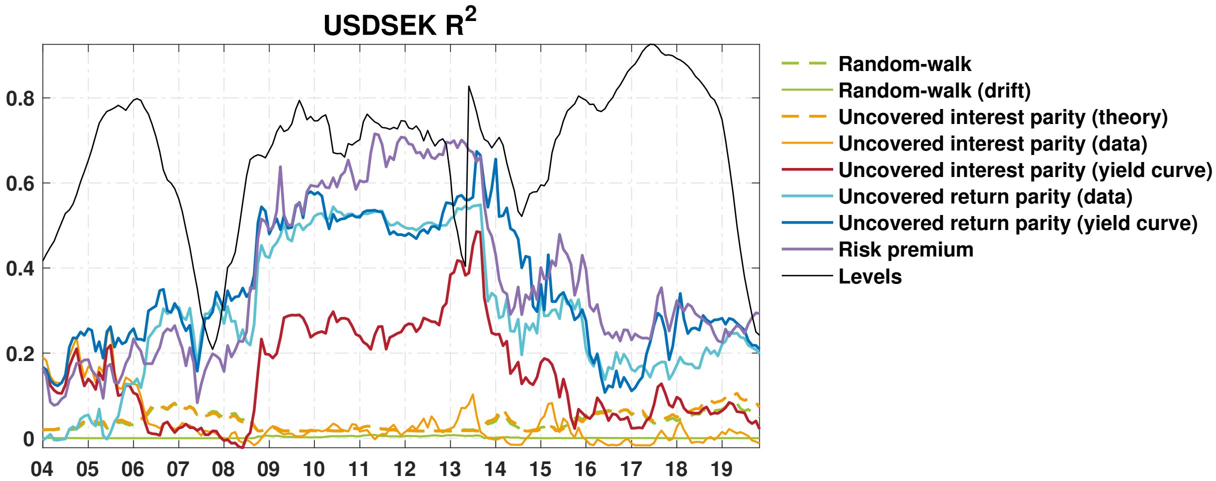
(b) Risk premium model.



Note: This figure plots 12-months-ahead forecasts from set of different models. All models are estimated on a $1999:1-2003:12$ sample and subsequently re-estimated by sequentially adding one new (end-of-month) observation until $2019:12$. Each model estimate is used to produce forecasts up to 12 months ahead in time.

In-Sample Forecasting Performance

Figure D.7: **Krona variation explained in-sample**. The percentage of USDSEK variation explained by the model determinants.



Note: The figure shows recursive window adjusted R^2 estimates across different models for USDSEK exchange rate. All models are estimated on a $1999:1-2003:12$ sample and subsequently re-estimated by sequentially adding one new (end-of-month) observation until $2019:12$. Each model estimate is used to produce a time-varying R^2 coefficient.

Out-Of-Sample Forecasting Performance

Before sample: start evaluation process before the Global financial crisis

Table D.1: **Out-Of-Sample Forecasting Performance** of various USDSEK exchange rate models.

	RMSE ratio			
	1 month	3 months	6 months	12 months
	Benchmark			
Random-walk	1.00	1.00	1.00	1.00
	Simple models			
Random-walk (drift)	1.00	1.00	1.00	1.00
Uncovered interest parity (theory)	1.00	1.00	1.00	1.00
Uncovered interest parity (data)	1.05	1.02	1.02	1.02
Levels	3.10	1.84	1.51	1.48
Forwards	11.78	0.99	1.01	1.06
	Complex models			
Uncovered interest parity (yield curve)	1.01	0.95	0.97	0.99
Uncovered return parity (data)	0.91	0.86	0.92	0.95
Uncovered return parity (yield curve)	1.00	0.97	1.03	1.09
Risk premium	0.98	0.91	0.94	0.96

Note: This table reports the ratio of models' root mean squared error (RMSE) relative to a random-walk, over different forecasting horizons.

	Direction accuracy			
	1 month	3 months	6 months	12 months
	Benchmark			
Random-walk	0.00	0.00	0.00	0.00
	Simple models			
Random-walk (drift)	0.54	0.54	0.62	0.62
Uncovered interest parity (theory)	0.49	0.49	0.44	0.37
Uncovered interest parity (data)	0.51	0.46	0.48	0.43
Levels	0.42	0.52	0.46	0.47
Forwards	0.38	0.57	0.49	0.40
	Complex models			
Uncovered interest parity (yield curve)	0.59	0.58	0.60	0.57
Uncovered return parity (data)	0.61	0.61	0.61	0.64
Uncovered return parity (yield curve)	0.57	0.55	0.53	0.52
Risk premium	0.61	0.66	0.61	0.61

Note: This table reports the percentage of model's forecasts that correctly projected the direction of change (direction accuracy), over different forecasting horizons.

	p-value (Diebold-Mariano)			
	1 month	3 months	6 months	12 months
	Benchmark			
Random-walk	0.00	0.00	0.00	0.00
	Simple models			
Random-walk (drift)	0.33	0.43	0.35	0.25
Uncovered interest parity (theory)	0.31	0.49	0.63	0.18
Uncovered interest parity (data)	0.04	0.35	0.28	0.04
Levels	0.00	0.01	0.17	0.25
Forwards	0.00	0.21	0.29	0.40
	Complex models			
Uncovered interest parity (yield curve)	0.65	0.13	0.29	0.83
Uncovered return parity (data)	0.11	0.26	0.25	0.14
Uncovered return parity (yield curve)	0.42	0.43	0.35	0.07
Risk premium	0.57	0.30	0.14	0.65

Note: This table reports the p-values of a Diebold-Mariano test for the model's performance relative to random-walk, over different forecasting horizons.

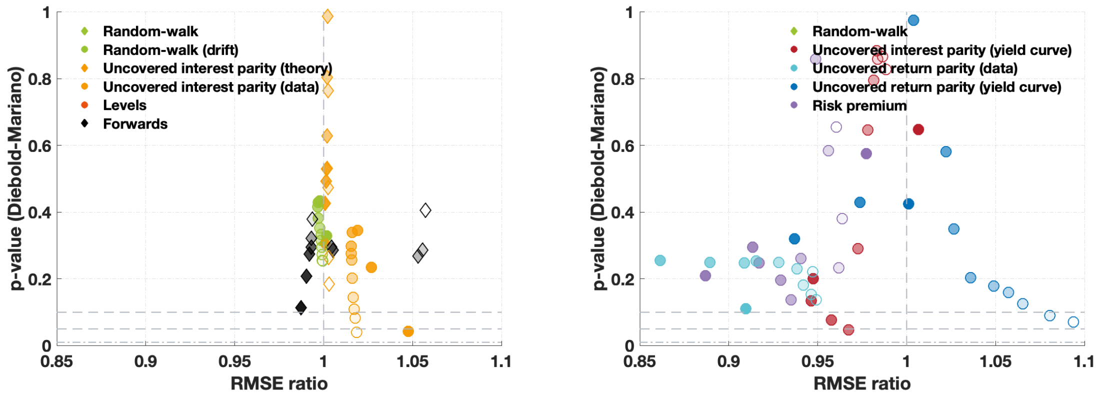
	Bias			
	1 month	3 months	6 months	12 months
	Benchmark			
Random-walk	−0.01	−0.04	−0.07	−0.14
	Simple models			
Random-walk (drift)	−0.01	−0.04	−0.07	−0.15
Uncovered interest parity (theory)	−0.02	−0.04	−0.07	−0.14
Uncovered interest parity (data)	0.01	−0.01	−0.05	−0.11
Levels	0.10	0.15	0.24	0.43
Forwards	2.13	−0.04	−0.11	−0.27
	Complex models			
Uncovered interest parity (yield curve)	0.01	−0.02	−0.05	−0.11
Uncovered return parity (data)	−0.01	−0.04	−0.08	−0.15
Uncovered return parity (yield curve)	−0.04	−0.09	−0.16	−0.29
Risk premium	−0.03	−0.06	−0.10	−0.17

Note: This table reports the model's mean error (bias), over different forecasting horizons.

Note: The table shows various performance criteria and statistics at different forecasting horizons for the USDSEK exchange rate. All models are estimated on $1999:1 - 2003:12$ sample and subsequently re-estimated by sequentially adding one new (end-of-month) observation until $2019:12$. Each model estimate is used to produce up to 12-month ahead forecast and the forecasting errors across different horizons are used to produce the relevant statistics.

Figure D.8: **Beating the benchmark vs. error magnitude accuracy.**

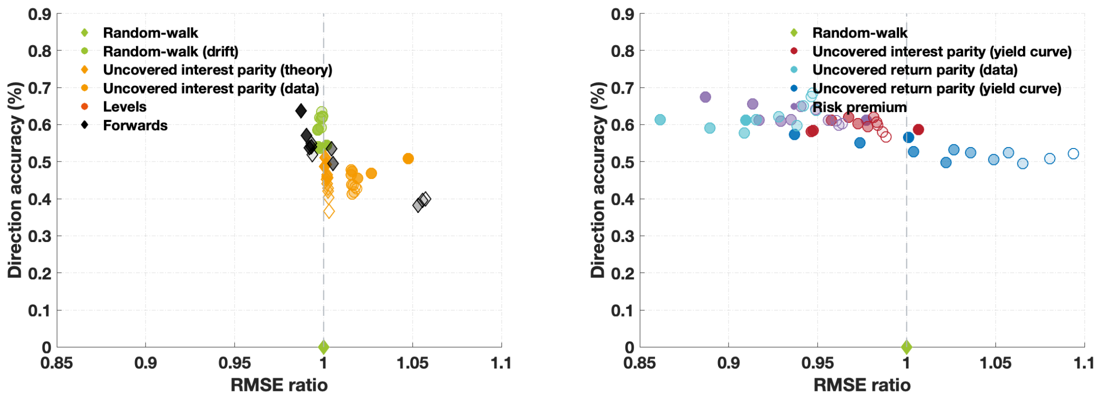
(a) Simple models. (b) Complex models.



Note: This figure reports the performance of various exchange rate models relative to a random-walk over different forecasting horizons. Each dot reports the p-value of the statistic pair(y-axis) and the ratio of the model's root mean squared forecast error relative to a random-walk (x-axis). The solid dots represent the statistic values at a one-month horizon. Lighter shades represent the statistic pair value further ahead in time. Clear dots represent the statistic pair at a one-year horizon.

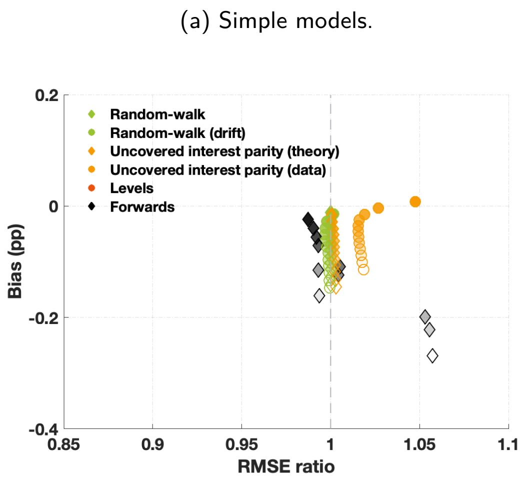
Figure D.9: **Direction vs. error magnitude accuracy.**

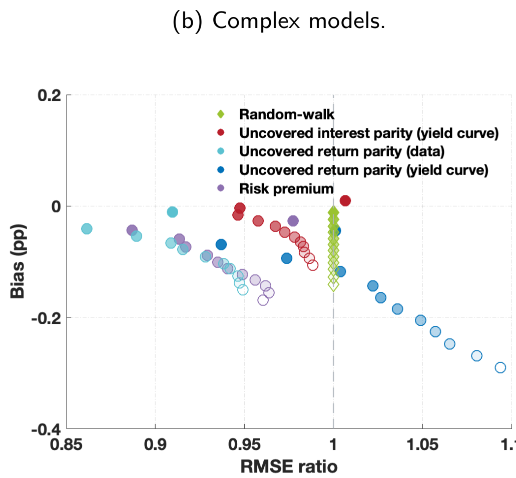
(a) Simple models. (b) Complex models.



Note: This figure reports the performance of various exchange rate models relative to a random-walk over different forecasting horizons. Each dot reports the direction accuracy statistic (y-axis) and the ratio of the model's root mean squared forecast error relative to a random-walk (x-axis). The solid dots represent the statistic values at a one-month horizon. Lighter shades represent the statistic pair value further ahead in time. Clear dots represent the statistic pair at a one-year horizon.

Figure D.10: **Forecasting bias vs. error magnitude accuracy.**





Note: This figure reports the performance of various exchange rate models relative to a random-walk over different forecasting horizons. Each dot reports the mean error (bias) statistic (y-axis) and the ratio of the model's root mean squared forecast error relative to a random-walk (x-axis). The solid dots represent the statistic values at a one-month horizon. Lighter shades represent the statistic pair value further ahead in time. Clear dots represent the statistic pair at a one-year horizon.

Out-Of-Sample Forecasting Performance

After sample: start evaluation process after the European sovereign debt crisis

Table D.2: **Out-Of-Sample Forecasting Performance** of various USDSEK exchange rate models.

	RMSE ratio			
	1 month	3 months	6 months	12 months
	Benchmark			
Random-walk	1.00	1.00	1.00	1.00
	Simple models			
Random-walk (drift)	1.01	1.01	1.01	1.01
Uncovered interest parity (theory)	1.01	1.01	1.01	1.01
Uncovered interest parity (data)	1.01	1.01	1.01	1.00
Levels	5.93	3.11	2.20	1.68
Forwards	9.06	1.00	1.02	1.05
	Complex models			
Uncovered interest parity (yield curve)	1.08	0.95	0.93	0.95
Uncovered return parity (data)	1.04	0.98	0.97	0.98
Uncovered return parity (yield curve)	1.05	0.96	0.95	0.95
Risk premium	1.09	0.92	0.90	0.90

Note: This table reports the ratio of models' root mean squared error (RMSE) relative to a random-walk, over different forecasting horizons.

	Direction accuracy			
	1 month	3 months	6 months	12 months
	Benchmark			
Random-walk	0.00	0.00	0.00	0.00
	Simple models			
Random-walk (drift)	0.44	0.46	0.42	0.33
Uncovered interest parity (theory)	0.44	0.43	0.30	0.25
Uncovered interest parity (data)	0.41	0.39	0.47	0.48
Levels	0.42	0.43	0.41	0.42
Forwards	0.65	0.49	0.44	0.43
	Complex models			
Uncovered interest parity (yield curve)	0.65	0.65	0.74	0.73
Uncovered return parity (data)	0.59	0.61	0.64	0.65
Uncovered return parity (yield curve)	0.55	0.61	0.65	0.63
Risk premium	0.66	0.67	0.62	0.65

Note: This table reports the percentage of model's forecasts that correctly projected the direction of change (direction accuracy), over different forecasting horizons.

	p-value (Diebold-Mariano)			
	1 month	3 months	6 months	12 months
	Benchmark			
Random-walk	0.00	0.00	0.00	0.00
	Simple models			
Random-walk (drift)	0.19	0.25	0.25	0.29
Uncovered interest parity (theory)	0.27	0.42	0.25	0.08
Uncovered interest parity (data)	0.08	0.16	0.08	0.03
Levels	0.00	0.01	0.05	0.12
Forwards	0.00	0.37	0.64	0.68
	Complex models			
Uncovered interest parity (yield curve)	0.31	0.34	0.27	0.31
Uncovered return parity (data)	0.92	0.49	0.39	0.43
Uncovered return parity (yield curve)	0.81	0.28	0.26	0.32
Risk premium	0.30	0.29	0.28	0.35

Note: This table reports the p-values of a Diebold-Mariano test for the model's performance relative to random-walk, over different forecasting horizons.

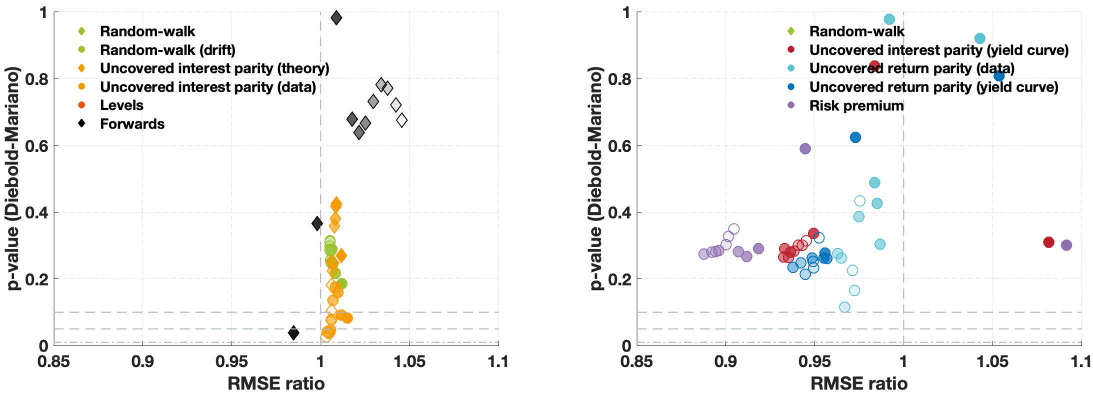
	Bias			
	1 month	3 months	6 months	12 months
	Benchmark			
Random-walk	−0.04	−0.14	−0.28	−0.52
	Simple models			
Random-walk (drift)	−0.05	−0.14	−0.29	−0.53
Uncovered interest parity (theory)	−0.05	−0.15	−0.29	−0.53
Uncovered interest parity (data)	−0.04	−0.13	−0.27	−0.51
Levels	−0.78	−0.88	−1.03	−1.32
Forwards	1.55	−0.15	−0.32	−0.62
	Complex models			
Random-walk	−0.04	−0.14	−0.28	−0.52
Uncovered interest parity (yield curve)	0.05	−0.04	−0.18	−0.42
Uncovered return parity (data)	−0.01	−0.11	−0.25	−0.50
Uncovered return parity (yield curve)	0.02	−0.08	−0.22	−0.46
Risk premium	0.06	−0.03	−0.17	−0.40

Note: This table reports the model's mean error (bias), over different forecasting horizons.

Note: The table shows various performance criteria and statistics at different forecasting horizons for the USDSEK exchange rate. All models are estimated on $1999:1-2013:12$ sample and subsequently re-estimated by sequentially adding one new (end-of-month) observation until $2019:12$. Each model estimate is used to produce up to 12-month ahead forecast and the forecasting errors across different horizons are used to produce the relevant statistics.

Figure D.11: **Beating the benchmark vs. error magnitude accuracy.**

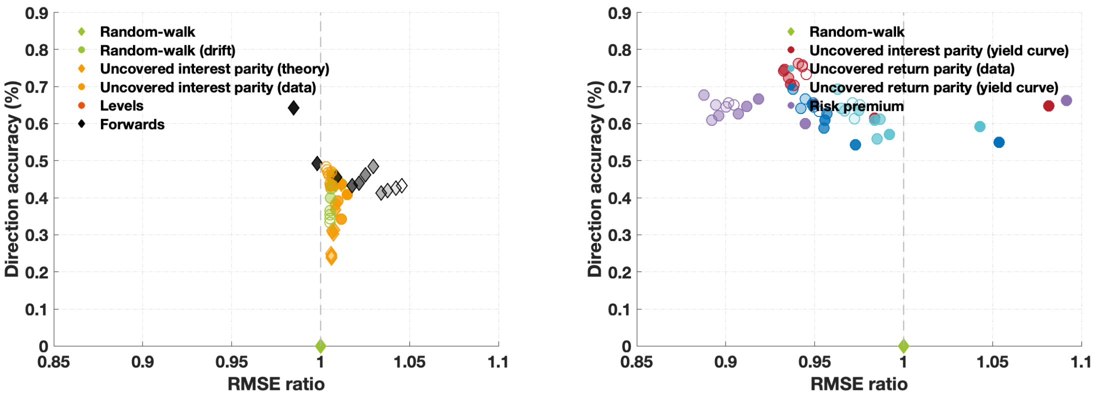
(a) Simple models. (b) Complex models.



Note: This figure reports the performance of various exchange rate models relative to a random-walk over different forecasting horizons. Each dot reports the p-value of the Diebold-Mariano test statistic (y-axis) and the ratio of the model's root mean squared forecast error relative to a random-walk (x-axis). The solid dots represent the statistic values at a one-month horizon. Lighter shades represent the statistic pair value further ahead in time. Clear dots represent the statistic pair at a one-year horizon.

Figure D.12: **Direction vs. error magnitude accuracy.**

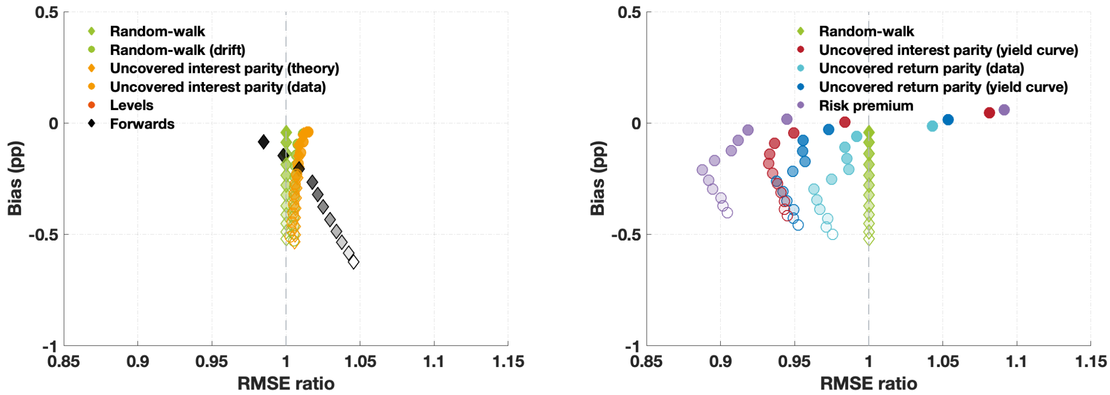
(a) Simple models. (b) Complex models.



Note: This figure reports the performance of various exchange rate models relative to a random-walk over different forecasting horizons. Each dot reports the direction accuracy statistic (y-axis) and the ratio of the model's root mean squared forecast error relative to a random-walk (x-axis). The solid dots represent the statistic values at a one-month horizon. Lighter shades represent the statistic pair value further ahead in time. Clear dots represent the statistic pair at a one-year horizon.

Figure D.13: **Forecasting bias vs. error magnitude accuracy.**

(a) Simple models. (b) Complex models.

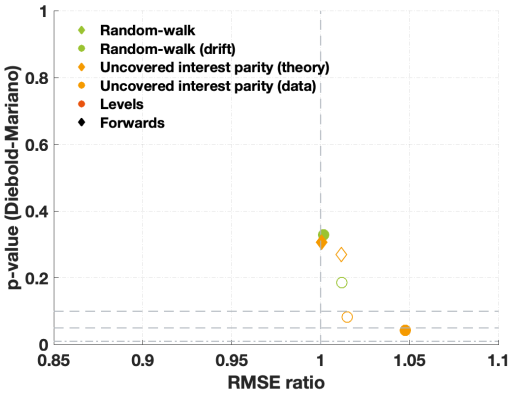


Note: This figure reports the performance of various exchange rate models relative to a random-walk over different forecasting horizons. Each dot reports the mean error (bias) statistic (y-axis) and the ratio of the model's root mean squared forecast error relative to a random-walk (x-axis). The solid dots represent the statistic values at a one-month horizon. Lighter shades represent the statistic pair value further ahead in time. Clear dots represent the statistic pair at a one-year horizon.

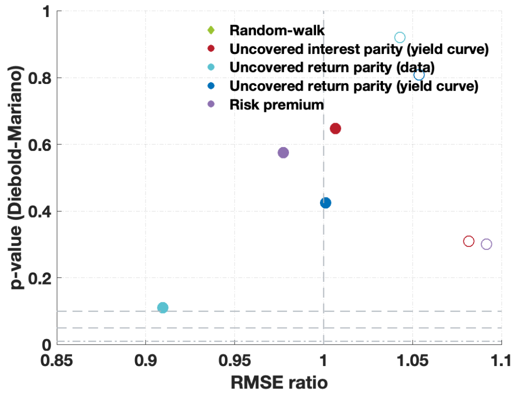
Out-Of-Sample Forecasting Performance Comparison: sample matters

Figure D.14: **Beating the benchmark vs. error magnitude accuracy (sample comparison).**

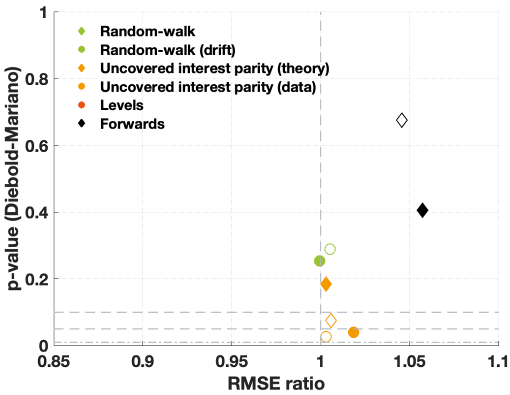
(a) Simple models, one-month forecasting horizon.



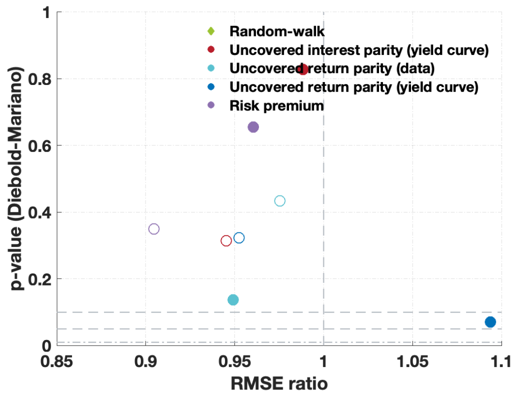
(b) Complex models, one-month forecasting horizon.



(c) Simple models, one-year forecasting horizon.



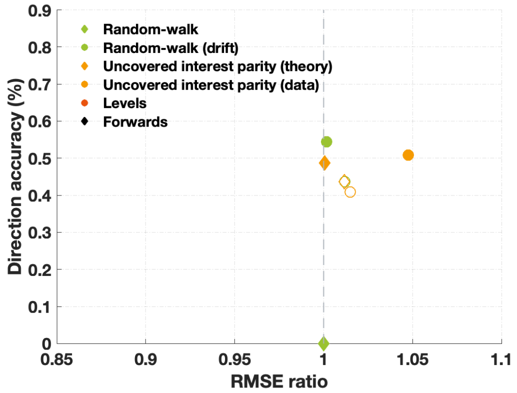
(d) Complex models, one-year forecasting horizon.



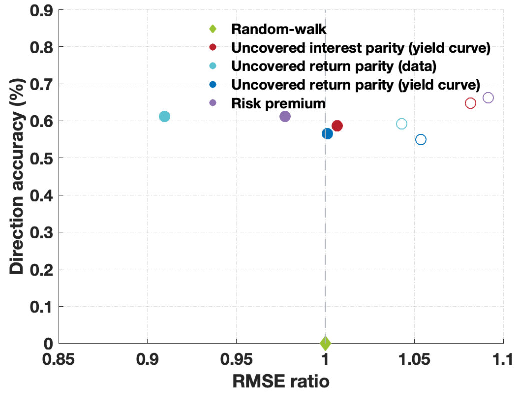
Note: This figure reports the performance of various exchange rate models relative to a random-walk over different sample periods. Each dot reports the p-value of the Diebold-Mariano test statistic (y-axis) and the ratio of the model's root mean squared forecast error relative to a random-walk (x-axis). The solid dots represent the statistic values computed over the *before sample* (models forecasting performance evaluation starting in 2004, before the beginning of the Global financial crisis). Clear dots represent the statistic pair computrd over the *after sample* (models forecasting performance evaluation starting in 2014, after the end of the European sovereign debt crisis).

Figure D.15: **Direction vs. error magnitude accuracy (sample comparison).**

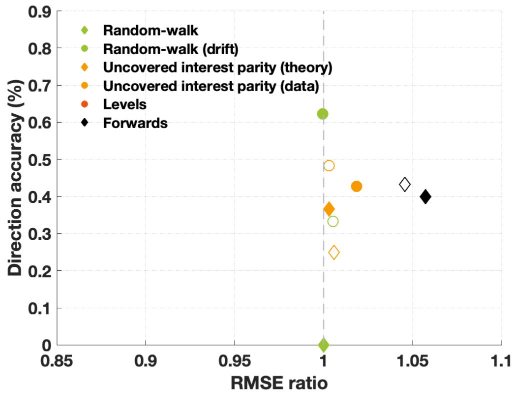
(a) Simple models, one-month forecasting horizon.



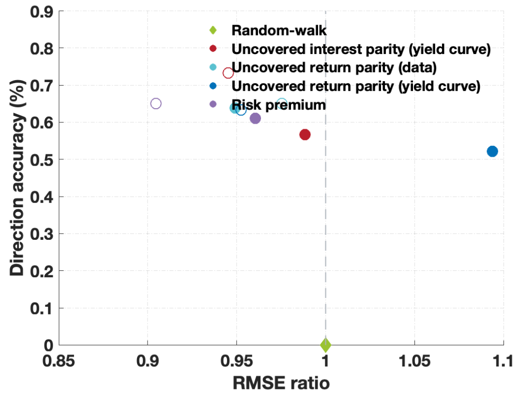
(b) Complex models, one-month forecasting horizon.



(c) Simple models, one-year forecasting horizon.



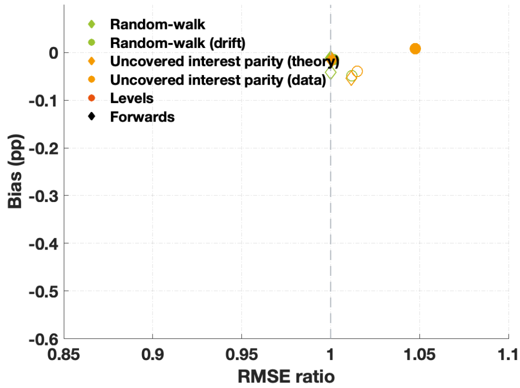
(d) Complex models, one-year forecasting horizon.



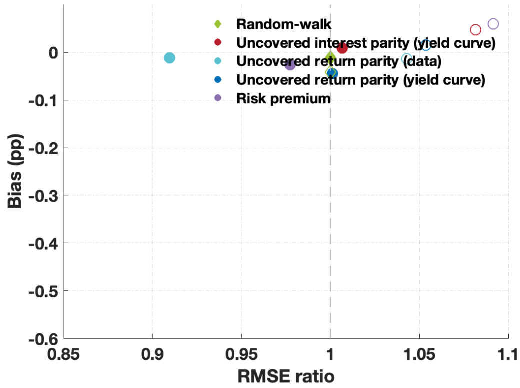
Note: This figure reports the performance of various exchange rate models relative to a random-walk over different sample periods. Each dot reports the direction accuracy statistic (y-axis) and the ratio of the model's root mean squared forecast error relative to a random-walk (x-axis). The solid dots represent the statistic values computed over the *before sample* (models forecasting performance evaluation starting in 2004, before the beginning of the Global financial crisis). Clear dots represent the statistic pair computrd over the *after sample* (models forecasting performance evaluation starting in 2014, after the end of the European sovereign debt crisis).

Figure D.16: **Forecasting bias vs. error magnitude accuracy (sample comparison).**

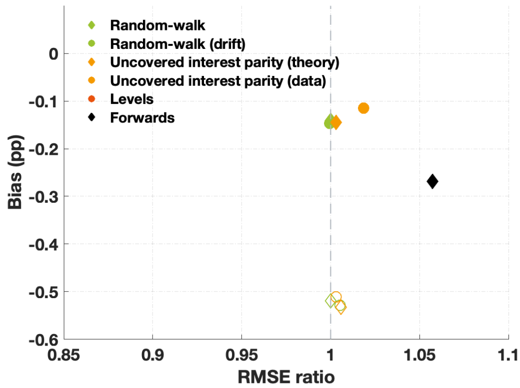
(a) Simple models, one-month forecasting horizon.



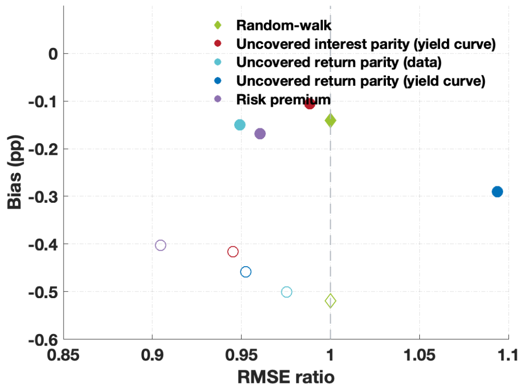
(b) Complex models, one-month forecasting horizon.



(c) Simple models, one-year forecasting horizon.



(d) Complex models, one-year forecasting horizon.



Note: This figure reports the performance of various exchange rate models relative to a random-walk over different sample periods. Each dot reports the mean error (bias) statistic (y-axis) and the ratio of the model's root mean squared forecast error relative to a random-walk (x-axis). The solid dots represent the statistic values computed over the *before sample* (models forecasting performance evaluation starting in 2004, before the beginning of the Global financial crisis). Clear dots represent the statistic pair computrd over the *after sample* (models forecasting performance evaluation starting in 2014, after the end of the European sovereign debt crisis).

SVERIGES RIKSBANK
103 37 Stockholm
(Brunkebergstorg 11)

Tel 08 - 787 00 00
Fax 08 - 21 05 31
registratorn@riksbank.se
www.riksbank.se