Journal of International Money and Finance 143 (2024) 103067

Contents lists available at ScienceDirect

Journal of International Money and Finance

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

Exchange rates and fundamentals: Forecasting with long maturity forward rates

Zsolt Darvas [a,*], Zoltán Schepp [b]

[a] Bruegel, Belgium, and Corvinus University of Budapest, Hungary
[b] University of Pécs, Hungary

ARTICLE INFO

JEL classification:
F31
F37
F41
G15

Keywords:
Exchange rates
Error correction
Forecasting performance
Monetary model
Out-of-sample
Random walk

ABSTRACT

We show that in a popular model of exchange rate determination, the unobserved expected future exchange rate can be substituted with the observed forward exchange rate. This allows the derivation of a new error-correction forecasting model, which approximates the gap between the fundamental equilibrium exchange rate and the actual exchange rate with the long-maturity forward exchange rate. Our out-of-sample forecasting results for major currencies are unprecedented. The forecasting model is simple, easy to replicate, and the data we use are available in real time and not subject to revisions.

1. Introduction

Forecasting foreign exchange rates is a central issue in international economics and financial market research. Since the seminal work of Meese and Rogoff (1983), hundreds of studies have attempted to outperform the random walk in out-of-sample forecasting with models based on macroeconomic fundamentals. These attempts have been either unsuccessful or if successful, subsequent work has disproved their results. A powerful formulation of the skeptical consensus was presented by Sarno and Taylor (2002) who, after conducting an extensive review of the literature, concluded that "*a model that forecasts well for one exchange rate and time period will tend to perform badly when applied to another exchange rate and/or time period*" (page 137).

Engel and West (2005) offered a theoretical explanation for this empirical forecasting failure: the exchange rate could be arbitrarily close to a random walk if the fundamentals have a unit root and the factor for discounting future fundamentals is close to one. This result implies that the out-of-sample forecasting power relative to the random walk is an unreliable gauge for evaluating exchange-rate models.

While an increasing number of studies have reported successful forecasting results over the past two decades, more recent surveys continue to be cautious about the predictability of exchange rates (Rossi, 2013; Engel, 2014; Cheung et al., 2019).

This article challenges the skeptical conclusion about exchange rate predictability by contributing to the literature in three ways. First, we show that in the general class of rational expectations present value exchange rate models analyzed by Engel and West (2005),

* Corresponding author.
E-mail addresses: zsolt.darvas@uni-corvinus.hu (Z. Darvas), schepp.zoltan@ktk.pte.hu (Z. Schepp).

https://doi.org/10.1016/j.jimonfin.2024.103067

Available online 21 March 2024

the unobserved expected future exchange rate can be substituted out with the observed synthetic forward exchange rate (which is calculated using the spot exchange rate and the interest rates of the two countries). This substitution simplifies the model solution and raises the possibility of forecasting with the help of the forward exchange rate.

Second, using a novel combination of the exchange rate model of Engel and West (2005) and the error-correction forecasting equation of Mark (1995) and Chinn and Meese (1995), we show that the forward exchange rate can be taken as a proxy for the difference between the fundamental equilibrium exchange rate and the current exchange rate when the maturity of the forward rate is long. This result allows the derivation of an error correction model in which the change of exchange rate is regressed on the previous period's long-maturity forward exchange rate. An error correction model implies the possibility of forecasting. While the empirical literature on uncovered interest rate parity (UIP) concludes that forward exchange rates are not unbiased predictors of future spot exchange rates, they can be used efficiently, in our error correction framework, to forecast future exchange rate changes for both short and long-horizon forecast horizons.

Third, we report unprecedented out-of-sample forecasting results. Our forecasting model leads to forecasts more accurate than the random walk in the January 1990 – January 2023 out-of-sample forecasting evaluation period, for major currencies, for all forecasting horizons between 1 month and 5 years, using four different forecast evaluation criteria. While past works have shown better than random-walk forecasts in some cases, our results show a Pareto-improvement relative to these. That is, our results are improved in at least one, and in most cases several, important aspects without sacrificing any other aspect. For example, some works report superior one-period-ahead forecasts, but not longer-horizon forecasts, and others the reverse. Our forecasts beat the random walk both in short and long-horizon forecasts, implying an internal consistency of our forecasts. We use longer out-of-sample forecasting periods and more forecast evaluation criteria than relevant previous works. Furthermore, we test the robustness of our results using various subperiods between 1990 and 2023 and find superior forecasting results with the exception of a few years around the collapse of Lehman Brothers in September 2008 in the case of the German mark(euro)/dollar rate,[1] a period when a sudden increase in demand for safe assets and the introduction of unconventional monetary policies might have altered the information content of long-maturity interest rates. After financial markets stabilized, our forecasting results were strong again.

Our outstanding forecasting results are not a spurious consequence of a possible non-stationarity of the forward exchange rate as a predictor. For the three bilateral pairs of the US dollar, German mark/euro, and British pound sterling, eight alternative tests soundly reject the null hypothesis of a unit root in long-maturity forward exchange rates. We also estimate the confidence interval for the error-correction parameter using the method of Rossi (2007), which is robust to whether the regressor has a unit root or not, and find that zero is not included in the interval. Moreover, in a similar context, forecasting the nominal exchange rate with the real exchange rate, Eichenbaum et al (2021) argue that an out-of-sample forecasting success is extremely unlikely if the nominal exchange rate is not predictable, regardless of whether the real exchange rate is stationary. To test the statistical significance of our forecasting results, we apply the same bootstrap test assuming non-predictability that Eichenbaum et al (2021) applied. We also use the Clark and West (2006, 2007) test statistics to evaluate the statistical significance of our forecasts, which is used by Clark and West (2006, 2007) in a setup with a highly persistent predictor. While our out-of-sample forecasting results are excellent for twelve major currency pairs, in this article, we report forecasting results only for the three bilateral pairs of the US dollar, German mark/euro, and British pound sterling, because there is larger uncertainty about unit roots for the other nine currency pairs. We leave a better understanding of the excellent forecasting results for other currency pairs for further research. The three currency pairs we study accounted for 37 % of global foreign exchange market turnover over the past three decades.

We apply several robustness tests, which all support our strong forecasting results. Our model outperforms various alternative models, such as the simple autoregression, the interest rate differential as the predictor, the real exchange rate as the predictor, and various versions of these models.

Our out-of-sample forecasts have outstanding properties when applying simple ordinary least squares (OLS) separately for each currency pair and also in panel models. The number of parameters to estimate varies from four to eight, no specification search is needed, and there is no need to calibrate any parameter. The simplicity of our models makes the replication of our results easy, in contrast to several works, which require the estimation of large numbers of parameters and/or a time-consuming process of model selection and estimation.

We do not use long-horizon regressions (in which the multi-period ahead change in exchange rate is regressed on explanatory variables) and therefore our forecasting model is not subject to "*overlapping observation*" issues, as discussed by Berkowitz and Giorgianni (2001), Rossi (2007), Darvas (2008) and Engel and Wu (2023). Instead, our longer-horizon forecasts are based on the iteration of one-period forecasts.

Also, data revision is not an issue for our model. For example, Faust et al (2003) found that the findings of Mark (1995) depend on the data vintage used to construct explanatory variables. In contrast, the only explanatory variable included in our model is the synthetic forward exchange rate, which we calculate from the spot exchange rate and the interest rates of the two countries. These data are available in real time and are not revised.

We use the synthetic forward exchange rate for forecasting, not a market-based forward exchange rate, because the synthetic forward exchange rate appears in our theoretical model. Thus, it is irrelevant for both our theoretical framework and our forecasting model whether covered interest rate parity (CIP) holds or not.

The rest of the article is organized as follows. Section 2 briefly describes how our article relates to the literature. Section 3 presents

[1] We use the German mark/US dollar rate up to December 1998 and extend it with the euro/dollar rate multiplied by the German mark/euro conversion rate from January 1999.

the theoretical framework and derives our main forecasting equation. Our benchmark forecasting model, along with alternative models, are described in section 4. Section 5 introduces the data and reports results from some preliminary data analysis. Section 6 presents our out-of-sample forecasting results, and section 7 concludes.

2. Related literature

Our work is related to four main strands of the literature. The first strand is the theoretical characterization of exchange rate behavior in a rational expectations, present value framework. Engel and West (2005) made a seminal contribution to this literature, who solve the model by an infinite forward substitution of unobserved expected future exchange rates. Our work is the first to notice that when the model is solved for a finite horizon, the unobserved expected future exchange rate can be substituted with the observed (synthetic) forward exchange rate. In case the forward exchange rate has a stationary component, exchange rate forecasting is conceivable even if both the fundamentals and foreign exchange risk premium have a unit root, while the Engel and West theorem suggests that the exchange rate is indistinguishable from a random walk in such a case.

The second relevant strand of the literature uses an error correction framework for exchange rate forecasting, which was pioneered by Mark (1995) and Chinn and Meese (1995). The existing literature tried to measure the fundamental equilibrium exchange rate, but such attempts were often burdened with difficulties. Our approach completely differs from the existing literature. By combining the present value model of Engel and West (2005) with the error correction model of Mark (1995) and Chinn and Meese (1995), we demonstrate that the gap between the fundamental equilibrium exchange rate and the actual spot exchange rate can be well approximated by the forward exchange rate when the maturity of the forward rate is long. The approximation does not work when the maturity is short. In the next section, we provide numerical examples of the goodness of this approximation depending on the maturity of the forward rate and other parameters. Thus, instead of estimating the unobserved fundamental equilibrium exchange rate, our framework allows using the observed synthetic long-maturity forward exchange rate as the error correction term.

Third, our work is related to the large body of empirical literature claiming success in forecasting exchange rates. In our reading, these studies can be divided into two groups. One group is theory-oriented and uses fundamental variables, sometimes in a new macroeconomic context (e.g. Engel et al., 2007; Gourinchas and Rey, 2007; Cerra and Saxena, 2010; Ince et al., 2016; Ca'Zorzi and Rubaszek, 2020). The other group is empirically oriented and sometimes relies on *ad-hoc* model specifications (e.g. Clarida and Taylor, 1997; Wright, 2008; Chinn and Moore, 2011; Dal Bianco et al., 2012; Park and Park, 2013). Our work is related to theory-oriented approaches. Nevertheless, as noted in the introduction, our empirical results present a Pareto-improvement to all existing works. Our model beats the random walk in all forecasting horizons between one month and five years. Existing works report either long-run predictability but no short-run predictability, or the reverse, which was a difficult puzzle according to Engel (2014). Our 33-year-long out-of-sample forecasting period is also longer than in any other paper, and we have adopted several robustness checks.

The fourth branch of the relevant literature seeks to explore the relationship between the yield curve, especially its long end, and the exchange rate. Chinn and Meredith (2005), Chinn and Quayyum (2012) and Valchev (2020) reported more favorable results for UIP at long horizons. Yet we find that even long-horizon UIP does poorly in out-of-sample exchange rate forecasting. Chen and Tsang (2013) extracted relative Nelson-Siegel (1987) factors from cross-country yield curve differences and use these for in-sample forecasting of nominal exchange rates. They did not find out-of-sample predictability as we do. A simplified version of the Chen and Tsang (2013) model is studied by Cheung et al (2019), who confirmed its inability to forecast out-of-sample. Lustig et al (2019) studied the term structure of carry trade returns in a model different from ours. They presented in-sample predictive regressions for currency excess returns, but not for exchange rates, and they did not provide any hint of the applicability of their model in out-of-sample forecasting either for excess returns or for exchange rates.

3. Theoretical framework

In this section we demonstrate the usefulness of synthetic long-maturity forward exchange rates in predicting future exchange rate changes in a novel framework different from uncovered interest rate parity. To this end, we point to a hitherto unrecognized substitution option that allows to combine the present value models analyzed by Engel and West (2005) and Engel (2014) with Mark's (1995) and Chinn and Meese's (1995) exchange rate forecasting error correction model.

Mark (1995) and Chinn and Meese (1995) considered the following general error correction model for exchange rate forecasting, based on theoretical models involving fundamental determinants of exchange rates, such as the monetary model:

$$s_{t+k} - s_t = \alpha_k + \beta_k \bullet (f_t - s_t) + \varepsilon_{t+k,k} \tag{1}$$

where s_t is the logarithm of the spot exchange rate (measured as the price of one unit of foreign currency in terms of domestic currency), f_t is the logarithm of the fundamental equilibrium value of the exchange rate, α_k and β_k are model parameters, and $\varepsilon_{t+k,k}$ is the k-period ahead forecast error. $f_t - s_t$ is the error correction term, as it measures the deviation of s_t from its fundamental equilibrium value.

According to this approach, exchange rate changes could be forecast using the difference between the fundamental and actual values of the exchange rate, thereby assuming an error correction mechanism. Papers using this approach typically estimate f_t from a theoretical exchange rate model. We follow a different approach by approximating the difference between the fundamental equilibrium and actual exchange rates $(f_t - s_t)$, which is the long-maturity synthetic forward exchange rate multiplied by a scalar, as we demonstrate below.

3.1. The Engel and West (2005) present value models

We start with the key equation of Engel and West (2005), who analyzed a general class of theoretical exchange rate models in a rational expectation, present-value framework (see Eq. (2) in Engel and West, 2005). Engel (2014) presented the simplified version of this key equation (see Eq. (45) in Engel, 2014) as:

$$s_t = (1-b) \bullet f_{1t} + b \bullet f_{2t} + b \bullet E_t[s_{t+1}] \tag{2}$$

where f_{1t} and f_{2t} are the convex combinations of exchange-rate fundamentals, parameter b is the discount factor, which falls in the range $0 < b < 1$, and $E_t[.]$ denotes the expectations operator based on information available at time t.

Engel and West (2005) demonstrated that under certain assumptions,[2] the monetary model of exchange rate can be written in this form,[3] in which f_{2t} equals the ex-ante deviation from uncovered interest rate parity (which is sometimes called 'expected excess return' or 'foreign exchange risk premium') multiplied by minus one:

$$\rho_t \equiv (E_t[s_{t+1}] - s_t) - \widetilde{i}_t^{(1)} \tag{3}$$

$$f_{2t} = -\rho_t \tag{4}$$

where $\widetilde{i}_t^{(1)}$ is the one-period interest rate differential between domestic and foreign interest rates applicable for one period. For example, if the monthly interest rate is 1 percent per year, then the actual interest earned in one month is (approximately) 1/12 percent. Since $E_t[s_{t+1}]$ is unobservable, so does ρ_t.

f_{1t} could include observed and unobserved fundamental determinants of the exchange rate depending on the exchange rate theory considered. For example, in the case of the monetary model:

$$f_{1t} = \left[m_t - m_t^* - \gamma\left(y_t - y_t^*\right) + q_t - \left(v_t - v_t^*\right)\right] \tag{5}$$

where m_t denotes the logarithm of domestic money supply, y_t the logarithm of domestic income, q_t the real exchange rate, v_t the home shocks to money demand.[4] Foreign variables are denoted with *. f_{1t} as defined by (5) is essentially what Mark (1995) denoted by f_t. Hence, for simplicity, we denote:

$$f_{1t} = f_t \tag{6}$$

Using this notation, (2) can be rewritten as:

$$s_t = (1-b) \bullet f_t - b \bullet \rho_t + b \bullet E_t[s_{t+1}] \tag{2.b}$$

Engel and West (2005) show that the exchange rate follows a random walk for a discount factor b that is near 1 if $f_{2t}(=\rho_t)$ has a unit root, or $f_{2t} = 0$ and $f_{1t}(=f_t)$ has a unit root. To reach this result, Engel and West (2005) solve the expectational difference Eq. (2.b) by an infinite forward substitution of expected future exchange rates. In this case, the expected long-run exchange rate disappears from the present-value relationship for the spot exchange rate because its parameter, b^j, approaches zero as $j \to \infty$, provided that the expected exchange rate does not explode even faster (no bubble solution).

3.2. Finite horizon solution and forward exchange rate substitution

Infinite forward substitution is the usual way to solve forward-looking asset pricing models. In the Engel and West framework, it

[2] Assumptions for the monetary model: the parameters of the home and foreign money demand relationships are identical, and PPP holds.

[3] The Taylor-rule model can also be written in a similar form, whereby only ρ_t has a parameter which is minus one times the parameter of $E_t[s_{t+1}]$. The other fundamental drivers have different parameters. To reach such a form, the following assumptions should be made for the Taylor-rule model: at least the foreign country follows a monetary policy rule that explicitly includes exchange rates (besides the output gap and the inflation target, which are also included in the home country policy rule) and the target exchange rate is the level of PPP.

[4] Engel and West (2005, page 492) interpreted money demand shocks in the following way: "*Our "shocks" potentially include constant and trend terms, may be serially correlated, and may include omitted variables that in principle could be measured.*".

leads to an exchange rate determination equation in which the current and expected values of f_t and ρ_t appear as determinants. ρ_t is a forward-looking variable, which cannot be observed even after observing the actual future exchange rate.

In contrast, we solve the model for a finite horizon and introduce a novelty. Solving the model for a finite horizon (say for *n-1* periods) will leave the unobserved $E_t[s_{t+n}]$ in the exchange rate equation. The novelty is that we realize that the unobserved $E_t[s_{t+n}]$ can be substituted out with the use of the *n*-period synthetic[5] forward exchange rate, $d_t^{(n)}$, which is an observed variable at time *t*. The inclusion of the forward exchange rate in the spot exchange rate determination equation raises the possibility of exchange rate forecasting.

The synthetic forward exchange rate for one period, $d_t^{(1)}$, is defined as[6]:

$$d_t^{(1)} \equiv s_t + \tilde{i}_t^{(1)}. \tag{7}$$

Notice that (7) and (3) implies:

$$d_t^{(1)} = -\rho_t + E_t[s_{t+1}] \tag{8}$$

By substituting (8) into (2.b), we have eliminated two unobserved variables, ρ_t and $E_t[s_{t+1}]$, from (2.b), and instead included $d_t^{(1)}$, an observed variable:

$$s_t = (1-b)\bullet f_t + b\bullet d_t^{(1)} \tag{2.c}$$

A similar expression can be obtained by recursively substituting the expected future exchange rates up to a finite horizon: instead of the unobserved expected long-run exchange rate, the long-maturity synthetic forward rate will appear on the right-hand side of the exchange rate determination equation, which is observable at time *t*.

Shift (2.b) *j* periods ahead:

$$E_t\left[s_{t+j}\right] = (1-b)\bullet E_t\left[f_{t+j}\right] - b\bullet E_t\left[\rho_{t+j}\right] + b\bullet E_t\left[s_{t+j+1}\right] \tag{9}$$

The synthetic *n*-period forward exchange rate is[7]:

$$d_t^{(n)} \equiv s_t + \tilde{i}_t^{(n)} \tag{10}$$

where $d_t^{(n)}$ is the logarithm of the *n*-period synthetic forward exchange rate and $\tilde{i}_t^{(n)} = ln\left((1+i_t^{(n)})/(1+i_t^{(n)*})\right)$ is the *n*-period interest rate differential between domestic and foreign interest rates, measured as the total interest over period *n*. For example, if the annual interest rate is 2 percent per year and the expected annual interest rate in one year from now is also 2 percent per year, then the current 2-year interest rate is (approximately) 4 percent cumulatively for the two-year period.

The long-maturity domestic interest rate, $i_t^{(n)}$, is the sum of the current and expected future one-period interest rates when the expectation hypothesis of the terms structure (EHTS) holds. When the EHTS does not hold, a term premium, τ_t, should be included as well. Similarly, the foreign long-maturity interest rate might contain a term premium (τ_t^*). In our derivation, we do not assume that EHTS holds and hence do not assume that interest rates are predictable. We allow an interest rate term premium in the home and foreign countries and only assume that the term premium is the same ($\tau_t = \tau_t^*$). In this case, the term premium cancels out when we calculate the long-maturity interest rate differentials, which can be written as the sum of current and expected future one-period interest rate differentials:

[5] We use the synthetic forward exchange rate calculated from the spot exchange rate and the interest rates of the two countries rather than the actual forward exchange rate quoted on markets. The synthetic forward exchange rate is equal to the actual forward exchange rate if covered interest party (CIP) holds. However, since the synthetic forward exchange rate is part of our derivation and the synthetic forward exchange rate is used in our empirical analysis, it is not necessary for CIP to hold, nor is a liquid market required, for example, for the 10-year maturity actual forward exchange rate.

[6] Eq. (7) is the logarithm of $D_t^{(1)} = S_t\bullet\left(1+i_t^{(1)}\right)/\left(1+i_t^{(1)*}\right)$, where $D_t^{(1)}$ is the level of the 1-period forward rate, S_t is the level of the spot exchange rate, $i_t^{(1)}$ and $i_t^{(1)*}$ are the domestic and foreign 1-period interest rates measured at the frequency of the data (e.g. a 1 percent annualized interest rates corresponds to 1/12=0.0833 percent at the monthly frequency). Thereby, $\tilde{i}_t^{(1)} = ln\left((1+i_t^{(1)})/(1+i_t^{(1)*})\right)$.

[7] The definition in Eq. (10) is the logarithm of $D_t^{(n)} \equiv S_t\bullet\left(1+i_t^{(n)}\right)/\left(1+i_t^{(n)*}\right)$, where $D_t^{(n)}$ is the level of the *n*-period forward rate, S_t is the level of the spot exchange rate, $i_t^{(n)}$ and $i_t^{(n)*}$ are the domestic and foreign *n*-period interest rates measured as the total interest over period *n*. For example, if the 5-year interest rate is 2% per year, then the total interest over the 5-year period is approximately 10%. However, since all datasets we use include interest rates at the annualized level, in our empirical calculations we use the expression $D_t^{(n)} \equiv S_t\bullet\left(\left(1+\bar{i}_t^{(n)}\right)/\left(1+\bar{i}_t^{(n)*}\right)\right)^n$ to calculate the synthetic forward rate, in which $\bar{i}_t^{(n)}$ and $\bar{i}_t^{(n)*}$ are the domestic and foreign *n*-period annualized interest rates measured in percent and *n* indicates the maturity in years.

$$\tilde{i}_t^{(n)} = \tilde{i}_t^{(1)} + E_t\left[\tilde{i}_{t+1}^{(1)}\right] + \cdots + E_t\left[\tilde{i}_{t+n-1}^{(1)}\right] \tag{11}$$

This assumption is in line with the simplifying assumptions of the Engel and West (2005) model, which include the fulfilment of the purchasing power parity (PPP) hypothesis and identical specification and parameters of the money demand equations in both countries.

Analogously to (8), it is easy to derive:

$$d_t^{(n)} = -\rho_t - E_t[\rho_{t+1}] - E_t[\rho_{t+2}] - \cdots - E_t[\rho_{t+n-1}] + E_t[s_{t+n}] \tag{12}$$

By recursively substituting expected future exchange rates for $j = 1$ to n-1 as indicated in (9) into (2.b), instead of the unobserved $E_t[s_{t+n}]$, the currently observed $d_t^{(n)}$ will appear on the right-hand side of the exchange rate determination equation:

$$\begin{aligned} s_t &= (1-b)\bullet[f_t + b\bullet E_t[f_{t+1}] + b^2 \bullet E_t[f_{t+2}] + \cdots + b^{n-1} \bullet E_t[f_{t+n-1}]] \\ &- \left[b \bullet \rho_t + b^2 \bullet E_t[\rho_{t+1}] + \cdots + b^{n-1} \bullet E_t[\rho_{t+n-2}]\right] \\ &+ b^n \bullet [\rho_t + E_t[\rho_{t+1}] + \cdots + E_t[\rho_{t+n-2}]] \\ &- b^n \bullet [\rho_t + E_t[\rho_{t+1}] + \cdots + E_t[\rho_{t+n-1}]] + b^n \bullet E_t[s_{t+n}] = \\ &(1-b) \bullet \sum_{j=0}^{n-1} b^j E_t\left[f_{t+j}\right] - b \bullet \sum_{j=0}^{n-2} b^j E_t\left[\rho_{t+j}\right] + b^n \bullet \sum\nolimits_{j=0}^{n-2} E_t\left[\rho_{t+j}\right] + b^n \bullet d_t^{(n)}. \end{aligned} \tag{2.d}$$

A simple closed form solution can be obtained for s_t if we assumed that f_t and ρ_t follow autoregressive processes with iid errors, for which we denote the autoregressive parameters by β_f and β_ρ, respectively.[8] In this case (2.d) can be written in the following form:

$$s_t = g_f(n, b, \beta_f) \bullet f_t + g_\rho(n, b, \beta_\rho) \bullet \rho_t + g_d(n, b) \bullet d_t^{(n)} \tag{2.e}$$

That is, the current exchange rate can be written as a weighted sum of the time t values of f_t, ρ_t and $d_t^{(n)}$. Table 1 shows the analytical expressions for the weights, by discriminating the random walk case from the more general autoregressive case.

The weight of long-maturity forward rate is higher with higher values of the discount factor, but at the same time it decreases with forward maturity.

According to (2.e), the forward exchange rate can play a meaningful role in exchange rate determination within the Engel and West asset pricing model framework. Thus, in case the long-maturity forward exchange rate has a stationary component, exchange rate predictability is conceivable even if both f_t and ρ_t have a unit root, although according to the Engel and West theorem, the exchange rate is indistinguishable from random walk in such a case.

In order to derive a forecasting model, we reorganize Eq. (2.e):

$$g_f \bullet f_t - s_t = -g_\rho \bullet \rho_t - g_d \bullet d_t^{(n)} \tag{13}$$

The left side of (13) is closer to the gap between the fundamentals and the spot exchange rate, the closer g_f falls to 1. Table 2 calibrates this parameter under alternative values of the discount factor, maturity of forward exchange rate and the autoregressive parameter in the data generating process of f_t.[9]

Table 2 illustrates well that g_f is close to 1 for more persistent fundamentals, longer maturity forward rates, and a lower discount factor. It is possible to get g_f to be near 1 with reasonable assumptions on the coefficient values. When g_f is close to 1, then the right side of (13) well approximates the difference between the fundamentals and the spot exchange rate, which is the error correction term in (1), the forecasting equation of Mark (1995) and Chinn and Meese (1995). Even if the unobservable ρ_t is not equal to zero, the observable $d_t^{(n)}$ correlates with the error correction term. When $\rho_t = 0$, that is, when UIP holds, $d_t^{(n)}$ is the error correction term (multiplied with a scalar). More favorable results for long-run UIP (Chinn and Meredith, 2005; Chinn and Quayyum, 2012) imply that the long-run risk premium is zero. This assertion is also supported by Valchev (2020) and Berg and Mark (2019), who showed that over long horizons, the exchange rate tends to adjust relative to interest differentials, towards establishing UIP.

Even if g_f is somewhat less than 1, the synthetic long-maturity forward exchange rate correlates with the gap between the fundamentals and the spot exchange rate, while its negative coefficient on the right-hand side of (13) fits well with the logic of Mark's and Chinn and Meese's error correction prediction model.

On the contrary, for short maturity forward exchange rates, such as the three-month maturity rate considered in the first data block of Table 2, g_f is close to zero irrespective of other parameters. Hence, short-maturity forward exchange rates cannot be considered as proxies for the gap between the fundamentals and the spot exchange rate.

[8] The assumption of autoregressive processes for f_t and ρ_t follows Engel et al. (2007) and Engel (2014).

[9] Engel et al. (2007, page 425) used $\beta_\rho = 0.95$ at the quarterly frequency and assumed that β_f is even more persistent. At the monthly frequency, the cubic root of 0.95 should be used, which is 0.983.

Table 1
Weights of f_t, ρ_t and $d_t^{(n)}$ in Eq. (2.e).

	f_t and ρ_t are random walks	f_t and ρ_t are AR(1)
g_f	$1-b^{n-1}$	$\frac{b-1}{b\beta_f-1}\bullet\left[1-\left(b\beta_f\right)^{n-1}\right]$
g_ρ	$\frac{b}{1-b}\bullet\left(b^{n-2}-1\right)+b^n\bullet(n-2)$	$\frac{b}{b\beta_\rho-1}\bullet\left[1-\left(b\beta_\rho\right)^{n-2}\right]+b^n\bullet\frac{\beta_\rho^{n-2}-1}{\beta_\rho-1}$
g_d	b^n	b^n

Source: authors' calculations.

Table 2
The weight g_f as the function of alternative values of the discount factor, maturity of forward exchange rate and the autoregressive parameter in the data generating process of f_t.

Forward maturity	3-month forward rate			3-year forward rate			5-year forward rate			10-year forward rate		
DGP of f_t	I (1)	AR1 $\beta_f = 0.99$	AR1 $\beta_f = 0.98$	I (1)	AR1 $\beta_f = 0.99$	AR1 $\beta_f = 0.98$	I (1)	AR1 $\beta_f = 0.99$	AR1 $\beta_f = 0.98$	I (1)	AR1 $\beta_f = 0.99$	AR1 $\beta_f = 0.98$
b = 0.99	0.02	0.02	0.02	0.30	0.25	0.22	0.45	0.35	0.28	0.70	0.46	0.33
b = 0.98	0.04	0.04	0.04	0.51	0.44	0.38	0.70	0.56	0.46	0.91	0.65	0.50
b = 0.97	0.06	0.06	0.06	0.66	0.57	0.50	0.83	0.69	0.58	0.97	0.75	0.61
b = 0.96	0.08	0.08	0.08	0.76	0.67	0.60	0.91	0.77	0.66	0.99	0.80	0.68
b = 0.95	0.10	0.10	0.10	0.83	0.74	0.67	0.95	0.82	0.71	1.00	0.84	0.72

Source: authors' calculations.

When approximating $f_t - s_t$ with $d_t^{(n)}$ with large n in Eq. (1), the one-period ahead forecasting equation becomes:

$$s_{t+1} - s_t = \alpha_1 + \beta_1 d_t^{(n)} + \varepsilon_{t+1,1} \tag{14}$$

Since $d_t^{(n)}$ has a negative parameter in Eq. (13), we expect a negative parameter for β_1.

In our empirical work, we find a statistically significant negative parameter when we regress the one-period ahead change of the exchange rate on the level of the forward rate. We find that the estimated parameter (in absolute value) and the R2 of the regression is larger for longer-maturity forward rates (Table 3), thereby empirically supporting our analytical derivation.

4. Forecasting models

4.1. Our baseline specification

Eq. (1) with $k > 1$ belongs to the family of long-horizon regressions, leading to "*overlapping observations*", which poses immense econometric problems, as discussed by Berkowitz and Giorgianni (2001), Rossi (2007), Darvas (2008) and Engel and Wu (2023). In addition, such long-horizon regressions lead to information losses for two reasons. First, when forming out-of-sample forecasts from period t to period $t+k$, the estimation sample takes into account information contained in the explanatory variable only up to period t-k. Thereby, potentially important information between t-k and t-1 is not taken into account in the forecast. Second, when the explanatory variable, that is, $f_t - s_t$ in Eq. (1) and $d_t^{(n)}$ in Eq. (14), is stationary, the explanatory variable is expected to converge to its stationary mean in the forecast period and the speed of this convergence will also influence the forecast of s_{t+k}. The long-horizon regression does not take this information into account.

We therefore restrict our attention to the case of $k = 1$ when equations such as (1) and (14) can produce only one-period ahead forecasts on their own. But when combined with an autoregressive model for the predictor, multi-period ahead forecasts can be obtained by iterating one-period ahead forecasts. We use a first order autoregression for $d_t^{(n)}$ and thus our baseline forecasting specification is a very simple two-equation model[10]:

$$\begin{aligned} s_{t+1} - s_t &= \theta_1 + \theta_2 \bullet d_t^{(n)} + \varepsilon_{t+1,1}^{(1)} \\ d_{t+1}^{(n)} &= \theta_3 + \theta_4 \bullet d_t^{(n)} + \varepsilon_{t+1,1}^{(2)} \end{aligned} \tag{15}$$

where θ_1, θ_2, θ_3 and θ_4 are parameters and $\varepsilon_{t+1,1}^{(1)}$ and $\varepsilon_{t+1,1}^{(2)}$ are the error terms.

We use the standard iterative forecast procedure, which is used, for example, for non-overlapping autoregressions and vector-

[10] This two-equation model was initially proposed in the working paper Darvas and Schepp (2007), which reported outstanding forecasting results in the 1990–2006 out-of-sample forecast evaluation period, but provided only econometric justification for the model, not theoretical justification.

Table 3
Regression statistics of the one period change in the exchange rate.

Regressor		Monthly data			Annual data, January of each year		
		DEM/ USD	GBP/ USD	GBP/ DEM	DEM/ USD	GBP/ USD	GBP/ DEM
Spot exchange rate	δ_1	−0.013	−0.019	−0.008	−0.195	−0.268	−0.122
	t	−1.87	−2.17	−1.48	−2.13	−2.51	−1.71
	R2	0.007	0.009	0.004	0.098	0.130	0.065
	DW	1.91	1.85	1.90	1.65	1.71	2.10
	N	528	528	528	44	44	44
1-month synthetic forward exchange rate	δ_1	−0.013	−0.019	−0.008	−0.197	−0.272	−0.124
	t	−1.89	−2.21	−1.50	−2.16	−2.54	−1.73
	R2	0.007	0.009	0.004	0.100	0.133	0.066
	DW	1.91	1.85	1.90	1.65	1.71	2.09
	N	528	528	528	44	44	44
3-year synthetic forward exchange rate	δ_1	−0.021	−0.028	−0.011	−0.256	−0.366	−0.176
	t	−2.68	−2.99	−1.97	−2.52	−3.29	−2.17
	R2	0.013	0.017	0.007	0.132	0.205	0.101
	DW	1.91	1.85	1.89	1.65	1.68	2.06
	N	528	528	528	44	44	44
5-year synthetic forward exchange rate	δ_1	−0.027	−0.032	−0.013	−0.297	−0.419	−0.209
	t	−3.17	−3.39	−2.17	−2.78	−3.73	−2.40
	R2	0.019	0.021	0.009	0.156	0.249	0.120
	DW	1.91	1.86	1.89	1.64	1.66	2.04
	N	528	528	528	44	44	44
10-year synthetic forward exchange rate	δ_1	−0.037	−0.035	−0.015	−0.367	−0.438	−0.231
	t	−3.99	−3.75	−2.18	−3.19	−4.28	−2.55
	R2	0.029	0.026	0.009	0.195	0.303	0.134
	DW	1.91	1.86	1.89	1.66	1.63	2.00
	N	528	528	528	44	44	44
Real exchange rate	δ_1	−0.018	−0.019	−0.029	−0.283	−0.300	−0.432
	t	−2.16	−2.09	−2.97	−2.62	−2.57	−3.65
	R2	0.009	0.008	0.016	0.140	0.136	0.241
	DW	1.90	1.85	1.88	1.63	1.68	1.94
	N	528	528	528	44	44	44

Notes: Equation estimated: $s_{t+1} - s_t = \delta_0 + \delta_1 x_t + \varepsilon_{t+1}$, where x_t is either the logarithm of the spot exchange rate (first block), the logarithm of the *n*-period maturity synthetic forward rate (blocks 2–5; the maturity is showed in the first column), or the logarithm of the real exchange rate (last block). t: OLS t-statistics. R2: coefficient of determination; DW: Durbin-Watson. N: number of observations. The sample includes monthly data from January 1979 – January 2023. The annual data includes non-overlapping observations from January each year.
Source: authors' calculations.

autoregressions. Our model structure is like the model studied by Pincheira and West (2016). Let $\widehat{\theta}_{i|t}$ denote the estimate of θ_i based on information available at time *t*. The one-period ahead forecasts from the two equations are: $\widehat{s}_{t+1|t} = s_t + \widehat{\theta}_{1|t} + \widehat{\theta}_{2|t} \bullet d_t^{(n)}$ and $\widehat{d}_{t+1|t}^{(n)} = \widehat{\theta}_{3|t} + \widehat{\theta}_{4|t} \bullet d_t^{(n)}$, where $\widehat{s}_{t+1|t}$ and $\widehat{d}_{t+1|t}^{(n)}$ denote the one-period ahead forecasts based on time *t* information. Note that s_t and $d_t^{(n)}$ are observed at time *t*. The two-period ahead forecasts are obtained as: $\widehat{s}_{t+2|t} = \widehat{s}_{t+1|t} + \widehat{\theta}_{1|t} + \widehat{\theta}_{2|t} \bullet \widehat{d}_{t+1|t}^{(n)}$ and $\widehat{d}_{t+2|t}^{(n)} = \widehat{\theta}_{3|t} + \widehat{\theta}_{4|t} \bullet \widehat{d}_{t+1|t}^{(n)}$. Thus, the two-period ahead forecasts use information available only up to time *t*. And so on, we iterate the two equations forward based on information available only up to time *t*. Note that when $\widehat{\theta}_{4|t}$ is less than one (in absolute value), mean reversion of $d_t^{(n)}$ is also a driver of $\widehat{s}_{t+k|t}$ $(k > 1)$ forecasts.

This model is not estimated on overlapping samples and thereby avoids all problems associated with long-horizon regressions. It includes only four parameters that we estimate with OLS. There is no need to calibrate any parameter. We do not perform any specification search, though one might check if longer lags for $d_t^{(n)}$ help forecasting. We consider alternative values for *n*: 3 years, 5 years and 10 years. Thus, our model is simple, and replication of our results is very easy.

4.2. Alternative specifications using the synthetic forward rate

VECM: The first equation of model (15) is an error-correction equation, while the second equation is an autoregression for the error-correction term (the synthetic forward exchange rate). An alternative version is a standard vector error correction model (VECM):

$$\begin{aligned} s_{t+1} - s_t &= \varphi_1 + \varphi_2(s_t - s_{t-1}) + \varphi_3\left(\widetilde{i}_t^{(n)} - \widetilde{i}_{t-1}^{(n)}\right) + \varphi_4 \bullet d_t^{(n)} + \varepsilon_{t+1,1}^{(1)} \\ \widetilde{i}_{t+1}^{(n)} - \widetilde{i}_t^{(n)} &= \varphi_5 + \varphi_6(s_t - s_{t-1}) + \varphi_7\left(\widetilde{i}_t^{(n)} - \widetilde{i}_{t-1}^{(n)}\right) + \varphi_8 \bullet d_t^{(n)} + \varepsilon_{t+1,1}^{(2)} \end{aligned} \tag{16}$$

along with the definition of $d_t^{(n)} = s_t + \widetilde{i}_t^{(n)}$ in Eq. (10); φ_i are parameters to be estimated.

VAR in levels: A vector autoregression model (VAR) in levels can be consistently estimated irrespective of whether the variables have a unit root or not. An unconstrained VAR model does not specifically capture an error-correction mechanism. We employ a simple VAR(1):

$$\begin{bmatrix} s_{t+1} \\ d_{t+1}^{(n)} \end{bmatrix} = \begin{bmatrix} \omega_{0,1} \\ \omega_{0,2} \end{bmatrix} + \begin{bmatrix} \omega_{1,1} & \omega_{1,2} \\ \omega_{2,1} & \omega_{2,2} \end{bmatrix} \begin{bmatrix} s_t \\ d_t^{(\mathrm{n})} \end{bmatrix} + \begin{bmatrix} \varepsilon_{t+1,1}^{(1)} \\ \varepsilon_{t+1,1}^{(2)} \end{bmatrix}, \tag{17}$$

where $\omega_{i,j}$ are parameters to be estimated.

Both the VECM in Eq. (16) and the VAR in Eq. (17) share some of the advantageous features of model (15), since these are not estimated on overlapping samples, not subject to the various information losses we described earlier, and multi-step ahead out-of-sample forecasts are calculated with a dynamic iteration method. Model (16) includes eight parameters to be estimated, while model (17) includes six parameters, and both can easily be estimated with OLS.

4.3. Alternative predictors

For comparison we also report the forecasting performance of alternative predictions and predictors.

Forward rate: we consider the forward rate itself as the prediction, that is, $\widehat{s}_{t+1|t} = d_t^{(1)}$, $\widehat{s}_{t+2|t} = d_t^{(2)}$, …, $\widehat{s}_{t+60|t} = d_t^{(60)}$. While the empirical literature on UIP concludes that forward exchange rates are not unbiased predictors of future spot exchange rates at short horizons, recent research found more support for the long-horizon UIP (Chinn and Meredith, 2005; Chinn and Quayyum, 2012; Valchev, 2020).

We highlight that our forecasts are based on the error correction model (15), and thereby our forecasts are different the forward rate and so the prediction of UIP.

Simple autoregression: the predictor is s_t only, which implies a stationary process for the spot exchange rate when $|\theta_2| < 1$:

$$s_{t+1} - s_t = \theta_1 + \theta_2 \bullet s_t + \varepsilon_{t+1,1} \tag{18}$$

Price-level targeting central banks and the fulfillment of PPP would imply a stationary spot exchange rate. If central banks were inflation targeting, the exchange rate would have a unit root, but if the volatility of inflation were low, the exchange rate could be similar to a stationary process.

The spot exchange rate and the interest rate differential as separate predictors: the predictor in our baseline model (15) is the synthetic forward rate, $d_t^{(n)} = s_t + \widetilde{i}_t^{(n)}$ (see Eq. (10)). As a robustness check, we allow s_t and $\widetilde{i}_t^{(n)}$ to have different parameters:

$$\begin{aligned} s_{t+1} - s_t &= \theta_1 + \theta_2 \bullet s_t + \theta_3 \bullet \widetilde{i}_t^{(n)} + \varepsilon_{t+1,1}^{(1)} \\ \widetilde{i}_{t+1}^{(n)} &= \theta_4 + \theta_5 \bullet \widetilde{i}_t^{(n)} + \varepsilon_{t+1,1}^{(2)} \end{aligned} \tag{19}$$

Interest rate differential model: we also check $\widetilde{i}_t^{(n)}$ alone as a predictor, which is the other component of the synthetic forward rate:

$$\begin{aligned} s_{t+1} - s_t &= \theta_1 + \theta_2 \bullet \widetilde{i}_t^{(n)} + \varepsilon_{t+1,1}^{(1)} \\ \widetilde{i}_{t+1}^{(\mathrm{n})} &= \theta_3 + \theta_4 \bullet \widetilde{i}_t^{(n)} + \varepsilon_{t+1,1}^{(2)} \end{aligned} \tag{20}$$

Real exchange rate model: The real exchange rate is defined as $q_t = s_t + p_t^* - p_t$, where q_t is the logarithm of the real exchange rate, p_t is the logarithm of domestic consumer price level and p_t^* is the logarithm of foreign consumer price level:

$$\begin{aligned} s_{t+1} - s_t &= \theta_1 + \theta_2 \bullet q_t + \varepsilon_{t+1,1}^{(1)} \\ q_{t+1} &= \theta_3 + \theta_4 \bullet q_t + \varepsilon_{t+1,1}^{(2)} \end{aligned} \tag{21}$$

Several recent papers found that the real exchange rate is stationary, and that it is the nominal exchange rate that does most of the adjustment. Among the eight models studied by Cheung et al (2019), the real exchange rate model led to the most promising exchange rate forecasts. Ca'Zorzi and Rubaszek (2020) and Eichenbaum et al (2021) use the real exchange rate in long-horizon regressions to predict the nominal exchange rate and find that long-horizon (2 to 5 years in Ca'Zorzi and Rubaszek (2020) and 3 to 6 years in Eichenbaum et al (2021)) forecasts are better than that of the random walk, while forecasts for shorter horizons are not.

Two components of the real exchange rate have different parameters: similarly to checking the two components of the synthetic forward rate separately, we also allow the components of the real exchange rate to have different parameters[11]:

$$\begin{aligned} s_{t+1} - s_t &= \theta_1 + \theta_2 \bullet s_t + \theta_3 \bullet \left(p_t^* - p_t\right) + \varepsilon_{t+1,1}^{(1)} \\ \left(p_{t+1}^* - p_{t+1}\right) &= \theta_4 + \theta_5 \bullet \left(p_t^* - p_t\right) + \varepsilon_{t+1,1}^{(2)} \end{aligned} \quad (22)$$

5. Data and some empirical preliminaries

5.1. Data

We test the forecasting performance of our model for the three bilateral pairs of the US dollar (USD), German mark/euro (DEM), and British pound sterling (GBP). For the German mark, we rescale the euro exchange rate from 1999 using the fixed conversion exchange rate between the mark and the euro. We continue to use German interest rates from 1999 rather than an average euro interest rate, because we also use the German interest rate in the pre-1999 period; thus, we consider the German interest rate and Germany's currency in the fully sample period. Moreover, the euro interest rate has been influenced by default risk and euro-exit risk after 2008.

For these three currency pairs, the underlying assumptions of the monetary model could be valid, eight tests reject the null hypothesis of unit root in long-maturity forward rates, and the unit root-robust methodology of Rossi (2007) confirms the adequacy of our main forecasting equation. Our model also forecasts well for nine other currency pairs, but since there is a larger uncertainty about unit roots in forward exchange rates for these currency pairs, we leave a better understanding of those forecasting results for further research. The three currency pairs we study accounted for 37 % of global foreign exchange market turnover on average in 1992–2019,[12] with the DEM/USD rate having the largest share at 25.1 %, followed by the GBP/USD rate 9.7 %, and the GBP/DEM rate at 2.2 %.

The sample includes end-of-month data from January 1979 to January 2023.

We use constant maturity zero coupon yields for interest rates. We calculate synthetic forward rates from the spot exchange rate and the interest rates of the two countries (see footnotes 6 and 7). All exchange and interest rate data are available in real-time and not revised. Exchange rate and US interest rate data are from the Federal Reserve,[13] the German mark/euro conversion rate is from the European Central Bank,[14] German interest rates are from the Bundesbank,[15] and UK interest rates are from the Bank of England.[16]

The source of consumer price index (which is needed for the real exchange rate model) is the International Monetary Fund's International Financial Statistics (IFS) dataset, which includes seasonally unadjusted series measured as the average of 2010 = 100. Since consumer prices have seasonal components and seasonality can differ across countries, we seasonally adjust the series using the X12 method and calculate the real exchange rate using seasonally adjusted domestic and foreign consumer price indices and the (unadjusted) nominal exchange rate. When using the real exchange rate for our out-of-sample forecasting exercise, we abstract from two issues, which probably help the forecasting performance of the real exchange rate model. First, we disregard the problem of the delayed publication of price indices, but assume that at the end of each month, the consumer price for that month is known. Second, we seasonally adjusted the full-sample latest available vintage of consumer prices. For a truly out-of-sample exercise, one needs to adjust data that was available in real time. For example, for a forecast made in December 1989, the CPI data up to November 1989 was available and one need to seasonally adjust that series and either make a forecast for the price level in December 1989, or use the one-period delayed value of the price level as it is done in the supplementary material of Eichenbaum et al (2021).

5.2. The one-period regression

The one-period regression, Eq. (1) with $k = 1$ or Eq. (14), is important for forecasting, as Berkowitz and Giorgianni (2001) demonstrate. We compare the full-sample one-period regression results for our model using long-maturity forward rates with the simple autoregression, an equation which uses the one-month forward rate as the explanatory variable, and an equation in which the real exchange rate is the explanatory variable (Table 3). We present results for our baseline monthly frequency, and for the annual frequency by using (non-overlapping) data from January each year.

Even the simple autoregression has some explanatory power. The model with the one-month synthetic forward rate leads to almost the same result as the autoregression, because the interest-differential component of the one-month forward rate is minor. The variance decomposition presented in the supplementary material confirms that the interest rate differential component has a negligible contribution to the variance of the one-month synthetic forward exchange rate. The model with the real exchange rate is somewhat

[11] We have also checked a version in which all three components of q_t (s_t, p_t^* and p_t) have different parameters, which specification led to worse forecasting result than the preceding two specifications and thereby we do not report it.

[12] The BIS triennial surveys (https://www.bis.org/statistics/rpfx22.htm) measure foreign exchange turnover in April of every third year between 1992 and 2019: we calculate the average of percentage shares reported by the surveys.

[13] https://www.federalreserve.gov/datadownload/Choose.aspx?rel=H10 and https://www.federalreserve.gov/datadownload/Choose.aspx?rel=H15.

[14] https://www.ecb.europa.eu/euro/exchange/de/html/index.en.html.

[15] https://www.bundesbank.de/en/statistics/money-and-capital-markets/interest-rates-and-yields/term-structure-of-interest-rates.

[16] https://www.bankofengland.co.uk/statistics/yield-curves.

better than the autoregression. For the two USD rates, models using long-maturity forward exchange rates as the explanatory variable are even better. The explanatory power increases with the maturity of the forward exchange rate. For example, for the DEM/USD rate at the monthly frequency, the R2 is 0.007 for the autoregression and the model with the 1-month forward rate, 0.013 when the 3-year maturity forward rate is used, 0.019 when the 5-year maturity forward rate is used, and 0.029 when the 10-year maturity forward exchange rate is used. A similar increase is observed for annual non-overlapping data, with the R2 reaching 0.2 for the DEM/USD rate and 0.3 for the GBP/USD rate. The point estimate for the slope coefficient of the explanatory variable is also higher (in absolute value) when long-maturity forward exchange rates are used. These results highlight that long-maturity forward exchange rates have additional explanatory power beyond the autoregression and the real exchange rate. The variance decomposition presented in the supplementary material confirms that the interest rate differential component has a large contribution to the variance of the ten-year maturity synthetic forward rate: its contribution is about three-quarters of the contribution of the spot exchange rate for the DEM/USD rate, while for the GBP/USD and GBP/DEM rates, the interest differential has an even larger contribution than the spot exchange rate. The real exchange rate is an inferior predictor compared to the long-maturity synthetic forward rate for the two USD-rates, but it does a better job than the synthetic forward rate for the GBP/DEM rate.

5.3. Robust confidence interval for the regression parameter

The issue of persistent or non-stationary predictors is hardly considered in the exchange-rate forecasting literature. This issue is not even mentioned in the literature survey of Rossi (2013) and in the comparative analysis of Cheung et al (2019). Nevertheless, we pay attention to this issue. Table 4 shows that for eight alternative unit root tests, the test statistics decline with an increase in maturity of the forward rate, suggesting lower persistence for longer maturity forward rates.[17] The null hypothesis of a unit root in long-maturity synthetic forward rates is soundly rejected by all eight tests employed for three currency pairs we study.[18] Interestingly, five of the eight tests reject the null hypothesis of unit root in the spot DEM/USD rate at the 10 % significance level, while the GBP/USD and GBP/DEM spot rates seem to contain unit roots. The rejection of the unit root null hypothesis in real exchange rates is less clear-cut than in the case of long-maturity forward rates: only one of the eight tests rejects the null hypothesis in the case of the DEM/USD rate, six of the eight tests in the case of GBP/USD rate, and two of the eight tests in the case of the GBP/DEM rate.

To further demonstrate that our results are robust to whether long maturity forward exchange rates include unit roots or not, we use the method proposed by Rossi (2007) to calculate confidence intervals of regression parameter of the one-period regression we studied in Table 3. This method is robust whether the regressor has a unit root or not and has good coverage properties. We use the Matlab code available at the website of Barbara Rossi. While the title of Rossi's article refers to long-horizon regressions, it can be applied for the one-period regression as well. Rossi (2007) herself studied the properties of her method in the case of the one-period regression and presented empirical applications, which included the one-period regression.

When the 1-month forward rate is used as the regressor, zero is always within the 95 percent confidence interval (Table 5). When the 10-year forward rate is used as the regressor, zero is not within the confidence interval. We therefore conclude that the synthetic long-maturity forward rates of the three currency pairs we study do not include unit roots according to eight alternative tests and can be useful in forecasting the spot exchange rate according to the method of Rossi (2007).

6. Out-of-sample forecasting

6.1. Forecast evaluation sample

We use the 1979–1989 sample period to form an initial estimate and evaluate our out-of-sample forecasts for the January 1990 – January 2023 period using forecasting horizons between 1 month and 5 years. We use a recursive estimation scheme in our baseline forecasting exercise, but as a robustness check we also use a rolling estimation scheme with windows of various lengths.

6.2. Forecast evaluation criteria

Rossi (2013) highlighted that "*The toughest benchmark is the random walk without drift*" (p. 1063), which we use as the benchmark for our models. We use four out-of-sample forecasting evaluation criteria.

The first indicator, which is the most widely used in the literature, is the mean squared forecast error (MSFE) relative to the driftless random walk ($MSFER_k$):

[17] We use the standard tests for unit root of Dickey and Fuller (1979) and Phillips and Perron (1988) and six other unit root tests. Elliott et al. (1996) proposed a family of test statistics that are invariant to the trend parameters and suggested two particular tests: a modified version of the Dickey-Fuller *t*-test, which is based on a local GLS detrending, and another feasible point optimal test, both having substantially improved power when an unknown mean or trend is present. Ng and Perron (2001) exploited the findings of Elliott et al. (1996) and applied the idea of GLS detrending to modify existing tests and showed non-negligible size and power gains can be made when used in conjunction with an autoregressive spectral density estimator at frequency zero. They suggested modifications of three test statistics studied by Perron and Ng (1996) and the feasible point optimal test statistics of Elliott et al. (1996).

[18] Darvas and Schepp (2009) were the first to notice that some long-maturity synthetic forward exchange rates of major currencies are stationary.

Table 4
Unit root tests.

	s_t	$d_t^{(1)}$	$d_t^{(12)}$	$d_t^{(36)}$	$d_t^{(60)}$	$d_t^{(120)}$	q_t
DEM/USD							
ADF	−1.87	−1.87	−1.91	−2.08	−2.31	−2.98**	−2.49
PP	−2.05	−2.06	−2.11	−2.27	−2.41	−2.95**	−2.67*
DFGLS	−1.78*	−1.79*	−1.90*	−2.08**	−2.31**	−2.98***	−0.76
ERS	3.95*	3.88*	3.41*	2.86**	2.32**	1.42***	12.17
NP MZa	−6.26*	−6.35*	−7.17*	−8.54**	−10.51**	−17.19***	−1.81
NP MZt	−1.77*	−1.78*	−1.89*	−2.07**	−2.29**	−2.93***	−0.76
NP MSB	0.28	0.28	0.26*	0.24*	0.22**	0.17***	0.42
NP MPT	3.92*	3.86*	3.42*	2.87**	2.33**	1.43***	11.17

	s_t	$d_t^{(1)}$	$d_t^{(12)}$	$d_t^{(36)}$	$d_t^{(60)}$	$d_t^{(120)}$	q_t
GBP/USD							
ADF	−2.17	−2.18	−2.25	−2.35	−3.18**	−3.97***	−2.16
PP	−2.42	−2.43	−2.50	−2.63*	−2.91**	−3.47***	−2.44
DFGLS	−0.73	−0.73	−0.81	−1.41	−2.84***	−2.92***	−1.87*
ERS	10.69	10.64	9.97	5.57	2.00**	1.55***	3.77*
NP MZa	−1.96	−1.98	−2.26	−5.24	−16.92***	−16.38***	−8.55**
NP MZt	−0.72	−0.73	−0.81	−1.40	−2.77***	−2.85***	−1.85*
NP MSB	0.37	0.37	0.36	0.27*	0.16***	0.17**	0.22**
NP MPT	9.92	9.87	9.24	5.26	1.96**	1.52***	3.68*

	s_t	$d_t^{(1)}$	$d_t^{(12)}$	$d_t^{(36)}$	$d_t^{(60)}$	$d_t^{(120)}$	q_t
GBP/DEM							
ADF	−1.48	−1.48	−1.47	−1.59	−1.83	−2.70*	−3.04**
PP	−1.53	−1.53	−1.54	−1.67	−1.87	−2.60*	−3.11**
DFGLS	−0.19	−0.20	−0.42	−0.96	−1.64*	−2.07**	−1.20
ERS	24.00	23.37	17.50	9.00	4.43	3.02**	8.88
NP MZa	−0.31	−0.35	−0.82	−2.63	−6.24*	−8.48**	−3.00
NP MZt	−0.18	−0.20	−0.41	−0.96	−1.64*	−2.05**	−1.20
NP MSB	0.59	0.58	0.5	0.36	0.26*	0.24*	0.40
NP MPT	22.35	21.78	16.42	8.61	4.36*	2.92**	8.11

Notes: The sample includes monthly data between January 1979 and January 2023. s_t: spot exchange rate; $d_t^{(1)}$: 1-month maturity synthetic forward exchange rate, $d_t^{(12)}$: 12-month maturity synthetic forward exchange rate, and so on; q_t: real exchange rate. ADF: augmented test of Dickey-Fuller (1979); PP: test of Phillips-Perron (1988); ERS DF: DF test with GLS detrending suggested by Elliott-Rothenberg-Stock (1996); ERS FPO: feasible point-optimal test of Elliott-Rothenberg-Stock (1996), NP MZa & MZt & MSB & MPT: four tests suggested by Ng-Perron (2001). Null hypothesis is unit root for all tests. The 1%, 5%, and 10% critical values are the following. ADF and PP: –3.45, –2.87, –2.57. ERS DF: –2.57, –1.94, –1.62. ERS FPO: 1.96, 3.23, 4.42. NP MZa −13.8, −8.1, −5.7. NP MZt: −2.58, −1.98, −1.62. NP MSB: 0.174, 0.233, 0.275. NP MPT: 1.78, 3.17, 4.45. ***, **, and * indicate rejection of the null hypothesis at 1%, 5%, and 10% significance level, respectively. *Source:* Authors' calculations.

$$MSFER_k = 100 \bullet \frac{P^{-1}\sum\left(s_{t+k} - \widehat{s}_{t+k|t}^{(M)}\right)^2}{P^{-1}\sum\left(s_{t+k} - \widehat{s}_{t+k|t}^{(RW)}\right)^2} \tag{23}$$

where k is the forecast horizon, s_{t+k} is the log exchange rate at period $t+k$, $\widehat{s}_{t+k|t}^{(M)}$ is the log of the forecast made at time t for $t+k$ by a model, $\widehat{s}_{t+k|t}^{(RW)} \equiv s_t$ is the log of the random walk forecast made at time t for $t+k$, and P is the number of forecasts made. Therefore, this measure is calculated as a percentage where a value below 100 indicates that the model outperforms the driftless random walk.

We use two methods test the null hypothesis of $MSFER_k < 100$: the Clark and West (2006, 2007) statistics and a bootstrap test.

The driftless random walk benchmark is nested in all models. When comparing nested models, standard asymptotic tests do not apply when testing the null hypothesis of equal forecast accuracy. Clark and West (2006, 2007) showed that under the null hypothesis that the data generating process is the random walk (or any parsimonious model), estimation of parameters of a larger model introduces noise into the forecasting process that will, in finite samples, inflate its MSFE. Clark and West (2006, 2007) also suggested an adjustment of mean squared prediction error statistics, which leads to approximately normal tests. Clark and West (2007) used this test statistic in a setup with highly persistent predictors. Clark and West (2006, 2007) derived their results for models estimated in direct form, i.e., in the form of long-horizon regressions, and when the forecasts are evaluated using a rolling-window estimation technique. However, Pincheira and West (2016) found that the Clark and West (2006, 2007) statistics also worked reasonably well when the iterated method is used to obtain multi-step forecasts and the recursive estimation scheme is used, which is our baseline setup. For the iterated method they considered a simple first order autoregression for the predictor, in the same way as in our forecasting model (15).

Table 5
95 percent confidence interval of the slope parameter of the one-period regression for alternative maturities of the forward rate, using the local-to-unity robust estimator of Rossi (2007).

	DEM/USD	GBP/USD	GBP/DEM
Regressor: 1-month synthetic forward rate			
Upper	0.0008	0.0062	0.0068
Lower	−0.0248	−0.0085	−0.0049
Regressor: 3-year synthetic forward rate			
Upper	−0.0047	0.0030	0.0052
Lower	−0.0391	−0.0240	−0.0134
Regressor: 5-year synthetic forward rate			
Upper	−0.0095	0.0005	0.0039
Lower	−0.0469	−0.0325	−0.0190
Regressor: 10-year synthetic forward rate			
Upper	−0.0176	−0.0090	−0.0003
Lower	−0.0593	−0.0488	−0.0306

Notes: Equation estimated: $s_{t+1} - s_t = \delta_0 + \delta_1 d_t^{(n)} + \varepsilon_{t+1}$. where s_t denotes the spot exchange rate;$d_t^{(n)}$ denotes the n-period maturity synthetic forward rate. The sample includes monthly data from January 1979 – January 2023.
Source: authors' calculations using the Matlab code of Barbara Rossi available at: https://sites.google.com/site/barbararossiwebsite/Barbara-Rossi-research.

We estimate the long-run variance using the method of Newey and West (1987).

We also use a non-parametric bootstrap test like those used in related papers, such as Mark (1995), Kilian (1999), McCracken and Sapp (2005), Engel and Wu (2023) and Eichenbaum et al (2021). We impose the null hypothesis of no predictability in the bootstrap data generating process corresponding to our baseline model specified in Eq. (15):

$$\begin{aligned} s_{t+1} - s_t &= \varepsilon_{t+1,1}^{(1)} \\ d_{t+1}^{(n)} &= \theta_3 + \theta_4 \bullet d_t^{(n)} + \varepsilon_{t+1,1}^{(2)} \end{aligned} \tag{24}$$

We estimate the bootstrap model (24) and store the estimated residuals. We sample with replacement the residuals for a sample of 2000 plus the actual length of the time series. Like Eichenbaum et al (2021), we jointly sample the two residuals, $\hat{\varepsilon}_{t+1,1}^{(1)}$ and $\hat{\varepsilon}_{t+1,1}^{(2)}$, to preserve the contemporaneous correlations between them. Using the first actual observation of the data and the bootstrapped residuals, we recursively create bootstrapped time series for s_t and $d_t^{(n)}$ and then discard the first 2000 observations. We then compute forecasts based on model (15) and the random walk without drift using the bootstrapped sample in the same way as with our observed sample and calculate forecast statistics. We repeat this 5000 times to arrive at an empirical bootstrapped distribution of test statistics and calculate the p-value as the lower tail of this distribution from the test statistics for observed data. Eichenbaum et al (2021) specified two versions of the bootstrap data generating process: one in which the predictor (the real exchange rate in their case) is assumed to be stationary, and another in which it is non-stationary. Since we found that the predictor in our case, the long-maturity forward exchange rate, is stationary, we use only the stationary version of the bootstrap data generating process.

We find that the Clark and West p-values and the bootstrap p-values are rather similar, which we demonstrate for the detailed DEM/USD forecasting results (Table 6). Since calculating bootstrapped p-values is time consuming, but the calculation of the Clark and West statistics is instantaneous, in most of this article we derive the p-values from the Clark and West statistics.

The second indicator is the share of correct sign (i.e. direction of change) predictions relative to the spot exchange rate ($DOCS_k$):

$$DOCS_k = 100 \bullet P^{-1} \sum I\left(s_{t+k} - s_t,\ \hat{s}_{t+k|t}^{(M)} - s_t\right) \tag{25}$$

where the $I(.,.)$ is an indicator function having the value 1 if its two arguments have the same sign and zero otherwise. Therefore, this measure is calculated as a percentage where a value above 50 indicates our model predicts the direction of change well more than half of the time. We use the test developed by Pesaran and Timmermann (1992) to test the null hypothesis that a model has no power in predicting the direction of change in the exchange rate.

The third indicator is the share of correct sign (i.e. direction of change) predictions relative to the forward exchange rate ($DOCF_k$):

$$DOCF_k = 100 \bullet P^{-1} \sum I\left(s_{t+k} - d_t^{(k)},\ \hat{s}_{t+k|t}^{(M)} - d_t^{(k)}\right) \tag{26}$$

From a currency trading perspective, the share of correct sign predictions relative to the forward exchange rate is more relevant than predicting the direction of change relative to the spot exchange rate, because a forward transaction is settled at the forward rate. The deviation of the future spot rate from the forward rate (and not from the current spot rate) determines whether there is a profit or

Table 6
Out-of-sample forecast evaluation. baseline results for DEM/USD.

	Full sample, different forecast horizons								Different samples, 36 M forecast horizon			
	1 M	3 M	6 M	12 M	24 M	36 M	48 M	60 M	1993–99	2000–06	2007–10	2011–23
	Mean squared forecast error (MSFE), random walk without drift = 100											
Model 3Y	99.54	98.71	97.41	92.33	84.58	74.14	67.44	62.15	57.15	56.51	279.59	49.81
p (BS)	(0.039)	(0.044)	(0.046)	(0.031)	(0.032)	(0.022)	(0.020)	(0.021)	(0.072)	(0.061)	(0.943)	(0.008)
p (CW)	(0.094)	(0.079)	(0.060)	(0.009)	(0.001)	(0.000)	(0.000)	(0.000)	(0.002)	(0.000)	(0.999)	(0.000)
Model 5Y	99.32	98.17	96.16	89.91	81.27	70.29	62.86	56.31	42.39	44.47	336.63	54.29
p (BS)	(0.023)	(0.028)	(0.026)	(0.017)	(0.019)	(0.013)	(0.014)	(0.015)	(0.026)	(0.023)	(0.968)	(0.016)
p (CW)	(0.045)	(0.037)	(0.024)	(0.003)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.999)	(0.000)
Model 10Y	99.78	99.39	97.51	90.93	83.13	77.79	72.55	66.47	60.73	45.50	319.30	81.25
p (BS)	(0.058)	(0.062)	(0.040)	(0.018)	(0.021)	(0.025)	(0.029)	(0.028)	(0.069)	(0.022)	(0.968)	(0.106)
p (CW)	(0.013)	(0.011)	(0.007)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)	(0.998)	(0.000)
Combined	99.14	97.75	95.32	88.38	79.31	69.69	63.35	57.58	44.10	46.19	310.51	56.24
p (BS)	(0.010)	(0.014)	(0.013)	(0.007)	(0.009)	(0.010)	(0.010)	(0.012)	(0.026)	(0.024)	(0.960)	(0.013)
p (CW)	(0.027)	(0.023)	(0.015)	(0.002)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.999)	(0.000)
	Correct sign prediction compared to the spot rate, %											
Model 3Y	53.4	53.4	55.1	58.8	59.6	66.6	69.7	68.0	53.6	88.1	18.8	77.9
p	(0.029)	(0.018)	(0.003)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	n.a.	(0.000)	n.a.	(0.000)
Model 5Y	52.6	52.7	55.6	59.3	58.0	66.0	68.9	67.2	63.1	86.9	18.8	71.7
p	(0.060)	(0.039)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	n.a.	(0.000)
Model 10Y	52.1	54.2	57.7	60.1	59.9	66.9	65.7	64.8	85.7	67.9	18.8	71.7
p	(0.166)	(0.024)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	n.a.	(0.000)
Combined	52.6	53.2	56.6	61.9	58.0	65.7	68.6	65.4	73.8	77.4	18.8	70.3
p	(0.089)	(0.039)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	n.a.	(0.000)
	Correct sign prediction compared to the forward rate, %											
Model 3Y	52.9	55.4	58.2	60.1	66.6	75.1	75.4	72.8	85.7	88.1	20.8	80.0
p	(0.038)	(0.005)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	n.a.	n.a.
Model 5Y	51.9	54.2	57.9	60.1	66.8	76.5	76.6	73.7	88.1	91.7	20.8	80.0
p	(0.110)	(0.028)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	n.a.	n.a.
Model 10Y	52.6	57.2	60.2	61.4	67.6	74.0	78.3	76.6	84.5	82.1	20.8	81.4
p	(0.088)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	n.a.	(0.000)
Combined	51.4	54.9	58.7	61.7	69.0	79.6	78.6	76.3	94.0	97.6	20.8	80.7
p	(0.176)	(0.013)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	n.a.	(0.000)
	Mean annualized profit, % per year											
Carry trade	3.2	3.3	3.8	3.1	3.5	2.8	1.8	1.3	4.6	5.1	−1.6	2.0
p (=0)	(0.037)	(0.023)	(0.004)	(0.002)	(0.000)	(0.000)	(0.000)	(0.002)	(0.000)	(0.000)	(0.992)	(0.000)
Model 3Y	3.7	3.4	3.3	3.3	3.4	3.4	2.9	2.3	4.4	7.2	−2.0	2.6
p (=0)	(0.022)	(0.019)	(0.012)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.998)	(0.000)

(*continued on next page*)

Table 6 (*continued*)

	Full sample, different forecast horizons								Different samples, 36 M forecast horizon			
	1 M	3 M	6 M	12 M	24 M	36 M	48 M	60 M	1993–99	2000–06	2007–10	2011–23
p (=CT)	(0.404)	(0.472)	(0.623)	(0.453)	(0.532)	(0.028)	(0.008)	(0.005)	(0.843)	(0.001)	(0.842)	(0.004)
Model 5Y	2.8	2.9	3.2	3.3	3.5	3.5	2.9	2.4	4.6	7.4	−2.0	2.6
p (=0)	(0.054)	(0.037)	(0.014)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.998)	(0.000)
p (=CT)	(0.567)	(0.575)	(0.640)	(0.421)	(0.497)	(0.012)	(0.004)	(0.001)	(0.603)	(0.000)	(0.842)	(0.004)
Model 10Y	3.9	3.5	4.0	3.5	3.5	3.1	2.8	2.4	4.5	5.7	−2.0	2.7
p (=0)	(0.022)	(0.014)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.998)	(0.000)
p (=CT)	(0.385)	(0.447)	(0.458)	(0.337)	(0.505)	(0.186)	(0.012)	(0.002)	(0.981)	(0.279)	(0.842)	(0.001)
Combined	3.3	3.1	3.3	3.6	3.8	3.7	3.1	2.6	4.9	7.8	−2.0	2.6
p (=0)	(0.041)	(0.028)	(0.010)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.998)	(0.000)
p (=CT)	(0.488)	(0.538)	(0.622)	(0.295)	(0.160)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.842)	(0.002)

Notes: The forecasting model is described in Eq. (15). “Combined” refers to a combined forecast, whereby three forecasts are combined with equal weights from three models using alternative maturity synthetic forward rates, 3 years, 5 years and 10 years. The sample period includes monthly data from January 1979 to January 2023. Using the recursive estimation window, out-of-sample evaluation of forecasts was performed in the 1990–2023 period except in the last four data columns, for which the evaluation period is indicated in the heading. For MSFE. p (BS) is the bootstrap p value of testing the null hypothesis that the model MSFE is the same as that of the random walk against the one-sided alternative hypothesis that the model is better, based on 5000 bootstrap draws. p (CW) is the p-value of the same null hypothesis based on the test of Clark and West (2006, 2007). The p values for the sign predictions are based on the test of Pesaran and Timmermann (1992). This test assumes that both the predictor and the outcome change sign in the forecast evaluation period, which assumption is not satisfied for some of the sub-periods we consider (“n.a.” is indicated in these cases). For the mean annualized profit, p (=0) is the p value of the null hypothesis that the Sharpe-ratio is zero against the one-sided alterative that it is positive, while p (=CT) the p value of the null hypothesis that the Sharpe-ratio of our model-based forecast is the same as the Sharpe-ratio of the carry trade strategy, against the one-sided alterative that the Sharpe-ratio based on our model is larger. See the results for other currency pairs in Table A1 of the supplementary material.
Source: authors’ calculations.

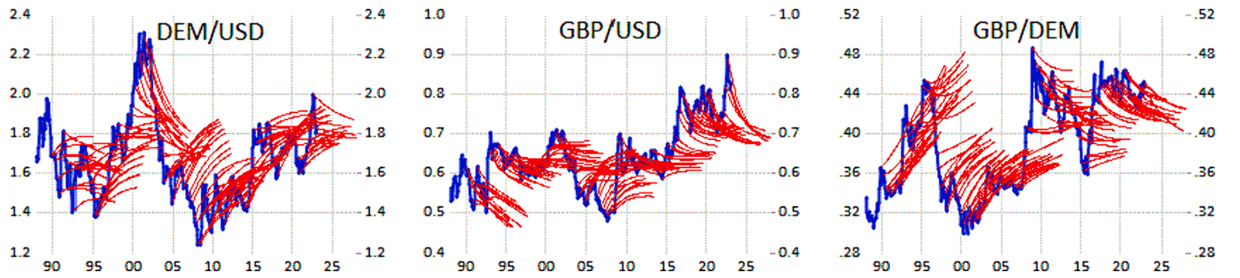

Fig. 1. Out-of-sample forecasts and actual exchange rates, 1988–2028 *Note:* The panels show actual exchange rate movements (blue line in 1988–2023) and out-of-sample forecasts for five years ahead (light red lines), starting, at each date, from the actual exchange rate. The latest forecast was made in January 2023 for the period from February 2023 to January 2028. Forecasts made in March, June, September and December of each year are shown. Although data was used in logarithmic form for estimation and forecasting, panels of this figure show data in their natural units (the price of one US dollar in terms of the other currencies in the first two panels and the price of one German mark in terms of the British pound sterling on the last panel). The forecasting model is defined in Eq. (15). The results of combined forecasts are reported, whereby three forecasts are combined with equal weights from three models using alternative maturity synthetic forward exchange rates, 3 years, 5 years, and 10 years. The sample period includes monthly data from January 1979 to January 2023. Using the recursive estimation window, out-of-sample evaluation of forecasts was performed in the 1990–2023 period, with estimates starting in 1979 (corresponding to our baseline results). *Source:* authors' calculations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

loss. We are not aware of papers calculating the share of correct sign predictions relative to the forward exchange rate. We again use the test of Pesaran and Timmermann (1992).

The fourth indicator is the excess return on a trading strategy based on our forecasting model where a positive value indicates excess profit. For comparison, we also report the return on a simple carry trade investment strategy.

The carry trade strategy on currency markets postulates that the currency with the higher interest rate is purchased by borrowing in a currency which has a lower interest rate. An equivalent carry trade transaction can be conducted in forward or futures markets, by buying a high-yield currency forward against a low-yield currency. The excess return, ignoring transaction costs and leverage, to the simple carry trade strategy is:

$$r^{(CT)}_{t,t+k} = \begin{cases} d^{(k)}_t - s_{t+k} & \text{if } d^{(k)}_t > s_t \\ 0 & \text{if } d^{(k)}_t = s_t \\ s_{t+k} - d^{(k)}_t & \text{if } d^{(k)}_t < s_t \end{cases} \tag{27}$$

where $r^{(CT)}_{t,t+k}$ is the excess return of the carry trade strategy realized in time $t+k$ for a forward transaction opened in time t for k-period ahead. For example, consider the UK pound sterling/USD dollar pair. When the UK interest rate is higher than the US interest rate ($d^{(k)}_t > s_t$), the UK pound sterling is purchased against the US dollar and the return after k periods is $d^{(k)}_t - s_{t+k}$. That is, if the UK pound sterling appreciates ($s_{t+k} < s_t$), remains unchanged ($s_{t+k} = s_t$), or depreciates less than the forward premium ($s_{t+k} < d^{(k)}_t$), a profit is realized. This return is in excess of the risk-free rate. To enter a forward or futures transaction, the investor must post collateral in the form of cash or appropriate high-quality marketable securities, such as a government bond. The investor earns interest income from the collateral. Therefore, the payoff for the forward transaction can be regarded as the return in excess of the risk-free interest rate, such as a government bond yield. Forward currency market transactions typically involve the use of leverage, as only a small percentage of the notional amount of the transaction (for example 4 %) is required by the financial intermediator as collateral. In our return calculations, we do not consider leveraged positions and transition costs.[19]

The trading strategy return implied by our forecasting model is defined as follows:

$$r^{(M)}_{t,t+k} = \begin{cases} d^{(k)}_t - s_{t+k} & \text{if } d^{(k)}_t > \widehat{s}^{(M)}_{t+k|t} \\ 0 & \text{if } d^{(k)}_t = \widehat{s}^{(M)}_{t+k|t} \\ s_{t+k} - d^{(k)}_t & \text{if } d^{(k)}_t < \widehat{s}^{(M)}_{t+k|t} \end{cases} \tag{28}$$

where $r^{(M)}_{t,t+k}$ is the excess return of an investment strategy based on model M realized in time $t+k$ for a transaction opened in time t for k-period ahead. Continuing the example above, if our forecast suggests the UK pound sterling will appreciate ($\widehat{s}^{(M)}_{t+k|t} < s_t$), will remain unchanged ($\widehat{s}^{(M)}_{t+k|t} = s_t$), or will depreciate less than implied by the forward rate ($\widehat{s}^{(M)}_{t+k|t} < d^{(k)}_t$), the UK pound sterling is purchased against the US dollar and the return for period $t+k$ is $d^{(k)}_t - s_{t+k}$.

[19] See Darvas (2009) for the quantification of transaction costs and the impact of leverage on carry trade returns.

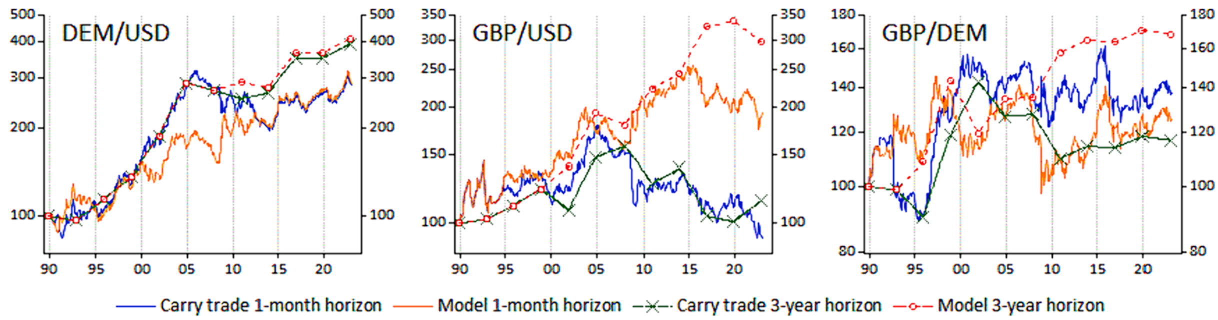

Fig. 2. Cumulative excess return to trading strategies based on our model and the carry trade, 1990–2023 *Note:* The values show the cumulative excess return to investment strategies based on our model (Eq. (28)) and the carry trade (Eq. (27)) for an initial 100 investment in December 1989, for monthly reinvestments based on monthly forecasts and three-year reinvestments based non-overlapping three-year forecasts. For our model defined in Eq. (15), the trading strategy based on combined forecasts is used, whereby three forecasts are combined with equal weights from the three models using alternative maturity synthetic forward exchange rates, 3 years, 5 years or 10 years. The sample period includes monthly data from January 1979 to January 2023. Using the recursive estimation window, out-of-sample evaluation of forecasts was performed in the 1990–2023 period, with estimates starting in 1979 (corresponding to our baseline results). Source: authors' calculations.

We report the mean annualized profit over our out-of-sample evaluation period. We test whether the mean annualized profit based on our model is larger than zero and whether it is larger than the profit of carry trade. We test these hypotheses by *t*-tests based on the Sharpe ratio (profit divided by its standard deviation), for which we estimate the long-run variance using the method of Newey and West (1987).

6.3. Baseline results for the German mark / US dollar rate

Our full-sample baseline results for the USD/DEM exchange rate indicate better-than-random walk forecasts for both short and long forecasting horizons, using three alternative maturity forward rates and four forecast evaluation criteria (first eight data columns of Table 6). The point estimates of the mean squared forecast error (MSFE) of our models are lower than that of the driftless random walk for forecasting horizons between 1 month and 5 years and for all alternative versions of our baseline model using different maturity forward rates. For all models, longer forecasting horizons are associated with stronger results relative to the driftless random walk. Table 6 also presents results from forecasts using the simple equally-weighted combination of the three models with alternative-maturity forward rates. Our findings corroborate findings from forecast combination literature (see e.g. Timmermann, 2006) showing a simple equal-weight combination performs well. Interestingly, the combined forecast outperforms the best individual model for six of eight alternative forecast horizons reported based on MSFE statistics. The improvement in forecast accuracy as measured by the MSFE of the combined model over the driftless random walk is 0.9 % for 1-month ahead forecasts, 11.6 % for 1-year ahead forecasts, 30.3 % for 3-year ahead forecasts, and 42.4 % for 5-year ahead forecasts. These are rather large improvements relative to models presented in past works. For example, Rossi (2013) analyzed the predictive ability of six single equation and two multiple-equation models for different sample periods staring between the 1960s and 1990s and ending in mid-2011 and found the bulk of the MSFE ratios are over 1 (or 100 if expressed it in percent) for both short-horizon (1 month or 1 quarter) and long horizon (4 years) forecasts. Among the 111 ratios reported for the 4-year forecasting horizon, only two were lower than 1, and both by only 1 percent.

The two alternative ways for testing the null hypothesis that the MSFE of our model is the same as that of the random walk led to rather similar results, in line with the findings of Pincheira and West (2016). The exceptions are few and do not change the big picture.

A visual representation highlights some great features of our forecasts. Fig. 1 indicates that our model was able to indicate both upward and downward turning points rather well, although many of the large excessive swings were forecasted to turn around earlier.

Turning to direction of change predictions, our models predict the correct signs in more than half of the cases for all three alternative models, as well as their combination, and for all forecasting horizons between one month and five years. At the one-month horizon correct sign predictions relative to the spot rate were made in 52–53 % of cases, while at the five-year horizon in 65–68 % of cases, which is large. All but one of these sign predictions are statistically significant at least at the 10 % level, and most of them are significant at the 1 % level.

Correct sign prediction relative to the forward rate is even more impressive, with 73 %-80 % correct predictions at 3–5-year horizons by the combined forecast.

It is therefore not surprising that a trading strategy based on our model leads to profit. The annualized excess return amounts to about 3 percent per year, which is economically significant, given that the average annualized US dollar interest rate was 2.9 % for the one-month interbank rate and 4.2 % for the 10-year government bond yield from 1990 to 2023. The annualized excess returns are significantly larger than zero, though generally not significantly larger than the return based on carry trade on shorter forecasting horizons, but significantly larger on longer forecasting horizons.

By assuming an initial investment value of 100, Fig. 2 visualizes the trading profit by considering one-month and three-year reinvestment decisions. That is, in the former case, the cumulative value of investment is reinvested according to our one-month ahead forecast for one-month horizon, while for the latter case, the investment decision is made in every third year based on our

Table 7
Sensitivity analysis and alternative models. Mean squared forecast error (random walk = 100).

DEM/USD	Full sample, different forecast horizons								Different samples, 36 M forecast horizon			
	1 M	3 M	6 M	12 M	24 M	36 M	48 M	60 M	1993–99	2000–06	2007–10	2011–23
A) Baseline (15)	99.1	97.7	95.3	88.4	79.3	69.7	63.3	57.6	44.1	46.2	310.5	56.2
	(0.027)	(0.023)	(0.015)	(0.002)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.999)	(0.000)
B) Estimation sample starts in 1985 (15)	100.0	99.8	98.3	92.3	81.6	73.9	67.8	63.9	76.1	52.0	266.6	47.9
	(0.019)	(0.015)	(0.008)	(0.002)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.002)	(0.995)	(0.000)
C) Panel (15)	98.9	97.1	94.3	87.7	79.4	70.5	65.1	60.9	50.6	53.5	265.5	52.9
	(0.025)	(0.021)	(0.013)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.998)	(0.000)
D) VECM (16)	97.9	97.4	95.2	88.8	80.6	72.1	66.3	61.7	51.8	51.6	288.7	55.3
	(0.007)	(0.017)	(0.016)	(0.003)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.999)	(0.000)
E) VAR in levels (17)	102.9	107.9	109.6	104.7	89.8	73.8	64.5	62.3	59.4	37.1	353.3	69.7
	(0.062)	(0.091)	(0.051)	(0.008)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.999)	(0.000)
F) Autoregression (18)	100.1	100.5	100.8	98.5	95.2	91.3	88.0	85.1	102.1	84.4	174.4	60.9
	(0.428)	(0.460)	(0.443)	(0.133)	(0.018)	(0.003)	(0.000)	(0.000)	(0.214)	(0.000)	(0.999)	(0.000)
G) Spot e.r. & interest-dif. separately (19)	102.9	107.8	109.2	103.1	85.9	67.6	60.0	58.8	47.6	39.3	302.3	64.8
	(0.062)	(0.092)	(0.051)	(0.007)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.998)	(0.000)
H) Interest rate model (20)	103.7	111.8	121.6	140.4	169.2	201.3	218.3	244.5	541.8	92.3	91.3	202.9
	(0.425)	(0.545)	(0.558)	(0.655)	(0.654)	(0.695)	(0.729)	(0.900)	(0.734)	(0.040)	(0.035)	(1.000)
I) Real exchange rate model (21)	99.9	99.6	98.8	95.0	88.2	82.4	75.4	69.6	78.8	71.2	99.6	111.2
	(0.203)	(0.160)	(0.109)	(0.018)	(0.003)	(0.002)	(0.000)	(0.000)	(0.010)	(0.002)	(0.309)	(0.036)
J) Spot e.r. & price level-dif. separately (22)	100.9	102.4	104.3	105.0	103.4	106.3	99.2	92.1	132.3	87.9	71.3	148.4
	(0.590)	(0.622)	(0.610)	(0.415)	(0.145)	(0.161)	(0.030)	(0.003)	(0.702)	(0.052)	(0.000)	(0.932)
K) Forward rate	100.7	101.9	103.7	107.5	113.8	120.7	116.6	107.8	148.7	117.1	70.0	124.2
	(0.844)	(0.872)	(0.932)	(0.984)	(0.999)	(1.000)	(0.997)	(0.940)	(1.000)	(1.000)	(0.002)	(0.996)

GBP/USD	Full sample, different forecast horizons								Different samples, 36M forecast horizon			
	1M	3M	6M	12M	24M	36M	48M	60M	1993–99	2000–06	2007–10	2011–23
A) Baseline (15)	99.6	98.0	94.8	92.6	82.7	78.4	72.4	73.7	91.9	71.1	67.4	85.6
	(0.036)	(0.016)	(0.009)	(0.010)	(0.001)	(0.000)	(0.000)	(0.000)	(0.020)	(0.000)	(0.008)	(0.014)
B) Estimation sample starts in 1985 (15)	100.9	100.7	98.2	100.1	93.9	97.2	93.6	98.8	170.5	60.6	68.3	111.6
	(0.035)	(0.016)	(0.008)	(0.011)	(0.003)	(0.005)	(0.002)	(0.004)	(0.814)	(0.000)	(0.008)	(0.116)
C) Panel (15)	99.3	97.5	94.6	91.6	82.7	78.7	74.2	75.2	90.1	81.5	63.0	80.5
	(0.030)	(0.013)	(0.008)	(0.009)	(0.002)	(0.000)	(0.000)	(0.000)	(0.028)	(0.000)	(0.008)	(0.006)
D) VECM (16)	99.3	97.6	95.0	92.9	83.0	77.8	72.3	73.9	86.2	74.9	65.4	84.8
	(0.036)	(0.015)	(0.009)	(0.010)	(0.001)	(0.000)	(0.000)	(0.000)	(0.009)	(0.000)	(0.008)	(0.012)
E) VAR in levels (17)	101.5	102.7	99.8	98.5	82.8	77.4	72.4	70.9	48.8	58.1	67.2	121.4
	(0.263)	(0.236)	(0.089)	(0.040)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.013)	(0.154)
F) Autoregression (18)	99.8	99.1	97.7	95.0	87.8	84.9	80.1	76.1	49.6	71.6	67.6	132.2
	(0.129)	(0.103)	(0.076)	(0.037)	(0.003)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)	(0.019)	(0.366)
G) Spot e.r. & interest-dif. separately (19)	101.5	102.9	100.8	101.3	88.1	82.5	75.7	73.0	88.2	67.2	64.2	108.2
	(0.263)	(0.255)	(0.126)	(0.088)	(0.005)	(0.001)	(0.000)	(0.000)	(0.012)	(0.000)	(0.011)	(0.113)
H) Interest rate model (20)	103.1	108.5	112.4	124.9	133.3	139.0	140.1	144.3	243.9	154.4	79.8	105.1
	(0.589)	(0.658)	(0.640)	(0.827)	(0.847)	(0.778)	(0.726)	(0.746)	(1.000)	(0.876)	(0.008)	(0.005)
I) Real exchange rate model (21)	100.2	100.0	98.8	94.5	84.7	81.4	81.0	82.5	59.1	72.2	73.2	110.2
	(0.208)	(0.168)	(0.111)	(0.026)	(0.001)	(0.000)	(0.000)	(0.000)	(0.006)	(0.000)	(0.029)	(0.085)

(*continued on next page*)

Table 7 (*continued*)

J) Spot e.r. & price level-dif. separately (22)	100.8	102.2	103.1	101.8	95.8	93.9	89.1	84.9	110.0	69.4	69.3	128.2
	(0.411)	(0.457)	(0.396)	(0.174)	(0.022)	(0.008)	(0.002)	(0.001)	(0.586)	(0.000)	(0.022)	(0.268)
K) Forward rate	100.1	99.7	99.3	97.2	97.4	99.7	103.5	98.7	62.1	123.1	102.5	95.4
	(0.524)	(0.464)	(0.443)	(0.310)	(0.347)	(0.483)	(0.687)	(0.424)	(0.046)	(0.995)	(0.610)	(0.250)
GBP/DEM	Full sample, different forecast horizons								Different samples, 36M forecast horizon			
	1M	3M	6M	12M	24M	36M	48M	60M	1993–99	2000–06	2007–10	2011–23
A) Baseline (15)	99.9	99.3	98.3	96.5	94.3	93.6	94.6	96.8	97.9	87.7	82.8	102.7
	(0.135)	(0.080)	(0.038)	(0.007)	(0.001)	(0.002)	(0.007)	(0.027)	(0.237)	(0.030)	(0.000)	(0.074)
B) Estimation sample starts in 1985 (15)	100.4	100.0	98.7	95.1	89.2	87.7	91.0	97.2	79.4	153.4	66.8	96.8
	(0.148)	(0.082)	(0.042)	(0.004)	(0.000)	(0.000)	(0.001)	(0.003)	(0.002)	(0.121)	(0.000)	(0.005)
C) Panel (15)	100.6	101.4	101.1	99.1	93.9	91.2	83.0	78.4	75.4	79.9	84.9	140.2
	(0.092)	(0.061)	(0.030)	(0.012)	(0.002)	(0.000)	(0.000)	(0.000)	(0.006)	(0.004)	(0.000)	(0.087)
D) VECM (16)	100.7	99.5	98.5	95.6	92.1	90.1	89.1	90.1	88.4	84.6	83.0	107.1
	(0.240)	(0.071)	(0.042)	(0.006)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.018)	(0.000)	(0.083)
E) VAR in levels (17)	101.3	102.1	102.3	99.4	98.0	103.2	118.6	134.4	119.3	96.0	80.3	104.4
	(0.155)	(0.115)	(0.089)	(0.025)	(0.018)	(0.059)	(0.438)	(0.761)	(0.335)	(0.051)	(0.000)	(0.129)
F) Autoregression (18)	100.4	101.3	102.7	104.7	109.8	113.9	122.8	131.9	135.4	89.9	101.2	98.4
	(0.798)	(0.831)	(0.876)	(0.860)	(0.887)	(0.920)	(0.985)	(0.990)	(0.995)	(0.036)	(0.940)	(0.093)
G) Spot e.r. & interest-dif. separately (19)	101.3	102.0	101.7	96.1	87.3	82.0	81.5	77.7	66.3	85.4	82.8	114.5
	(0.155)	(0.108)	(0.070)	(0.005)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.014)	(0.000)	(0.099)
H) Interest rate model (20)	102.7	107.0	112.5	117.9	127.6	143.7	167.0	191.4	180.6	155.4	60.8	170.8
	(0.327)	(0.299)	(0.314)	(0.281)	(0.281)	(0.367)	(0.610)	(0.708)	(0.549)	(0.603)	(0.000)	(0.777)
I) Real exchange rate model (21)	100.1	99.7	97.7	92.7	79.7	70.6	65.5	69.0	55.8	186.2	54.2	57.7
	(0.071)	(0.040)	(0.016)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.137)	(0.000)	(0.002)
J) Spot e.r. & price level-dif. separately (22)	101.3	102.8	103.2	99.9	88.0	79.5	72.4	70.3	63.2	206.8	63.4	62.7
	(0.254)	(0.178)	(0.107)	(0.019)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.255)	(0.000)	(0.001)
K) Forward rate	101.1	102.6	104.3	106.5	108.9	105.2	100.3	98.6	120.6	149.8	73.9	88.4
	(0.919)	(0.887)	(0.908)	(0.923)	(0.905)	(0.729)	(0.514)	(0.448)	(0.954)	(0.983)	(0.000)	(0.091)

Notes: The number after the model's name in the first column indicates the equation number that defined the model. For models using alternative maturity synthetic forward exchange rates or interest rates, the results of combined forecasts are reported, whereby three forecasts are combined with equal weights from the three models using alternative maturity synthetic forward rates or interest rates (3 years, 5 years or 10 years). The sample period includes monthly data from January 1979 to January 2023. Using the recursive estimation window, out-of-sample evaluation of forecasts was performed in the 1990–2023 period except in the last four data columns, for which the evaluation period is indicated in the heading. p values are reported in parentheses of testing the null hypothesis that the model MSFE is the same as that of the random walk against the one-sided alternative hypothesis that the model is better, based on the test of Clark and West (2006, 2007).
Source: authors' calculations.

Table 8
Recursive vs rolling estimation. Out-of-sample forecast evaluation. Mean squared forecast error (random walk = 100). DEM/USD rate.

		Forecast horizon							
Estimation method	Rolling estimation months	1 M	3 M	6 M	12 M	24 M	36 M	48 M	60 M
rolling	60	102.1	105.6	103.6	98.9	93.3	88.3	99.9	97.7
		(0.102)	(0.161)	(0.056)	(0.014)	(0.001)	(0.000)	(0.004)	(0.003)
	80	101.2	102.8	100.9	94.3	103.7	115.0	118.3	117.0
		(0.056)	(0.060)	(0.021)	(0.001)	(0.025)	(0.111)	(0.133)	(0.086)
	100	101.7	104.4	104.2	100.0	107.8	116.8	118.9	115.1
		(0.144)	(0.171)	(0.091)	(0.017)	(0.074)	(0.170)	(0.175)	(0.081)
	120	101.3	103.6	103.9	100.7	104.8	100.7	94.7	92.7
		(0.117)	(0.140)	(0.087)	(0.025)	(0.068)	(0.037)	(0.012)	(0.006)
	140	101.4	103.4	104.3	100.9	96.0	87.0	82.7	77.6
		(0.188)	(0.188)	(0.134)	(0.031)	(0.018)	(0.014)	(0.006)	(0.001)
	160	100.9	102.3	103.6	100.2	95.9	84.3	77.3	70.4
		(0.157)	(0.145)	(0.125)	(0.033)	(0.033)	(0.015)	(0.003)	(0.001)
	180	101.6	103.6	105.8	102.2	90.8	73.2	64.2	57.4
		(0.186)	(0.139)	(0.110)	(0.022)	(0.011)	(0.003)	(0.000)	(0.000)
	200	99.8	98.9	97.3	88.5	75.7	60.5	54.3	48.8
		(0.038)	(0.029)	(0.020)	(0.006)	(0.004)	(0.001)	(0.000)	(0.000)
	220	99.0	97.5	95.3	86.3	73.9	61.0	54.0	48.3
		(0.026)	(0.024)	(0.018)	(0.003)	(0.001)	(0.000)	(0.000)	(0.000)
	240	99.0	97.4	95.2	87.4	76.9	64.8	57.3	52.3
		(0.026)	(0.023)	(0.017)	(0.002)	(0.001)	(0.000)	(0.000)	(0.000)
recursive		99.1	97.7	95.3	88.4	79.3	69.7	63.3	57.6
		(0.027)	(0.023)	(0.015)	(0.002)	(0.001)	(0.000)	(0.000)	(0.000)

Notes: The forecasting model is defined in Eq. (15). The results of combined forecasts are reported, whereby three forecasts are combined with equal weights from three models using alternative maturity synthetic forward rates, 3 years, 5 years and 10 years. The sample period includes monthly data from January 1979 to January 2023. Using either the recursive or rolling estimation windows, out-of-sample evaluation of forecasts was performed in the 1990–2023 period. p values are reported in parentheses of testing the null hypothesis that the model MSFE is the same as that of the random walk against the one-sided alternative hypothesis that the model is better, based on the test of Clark and West (2006, 2007). See the results for other currency pairs in Table A2 of the supplementary material.
Source: authors' calculations.

three-year ahead forecasts. The same exercise is done for the carry trade too. For the DEM/USD currency pair, the one-month reinvestment horizon led to rather similar total cumulative returns for our model and the carry trade in 1990–2023, though return volatility is lower in the case of our model than for carry trade returns. For the three-year reinvestment horizon, 9 of the 11 decisions were the same for our model and for the carry trade. It is therefore not surprising that the overall performance over the 33-year period is rather similar for our model and the carry trade, even though our model led to much lower MSFE than the random walk.[20] This finding suggests that better forecasting ability of a model than the random walk might not necessarily be associated with significantly better model-based return than the return on carry trade.

The last four columns of Table 6 studies subperiod sensitivity for the 3-year ahead forecasts,[21] by breaking down the pre-global financial crisis period into two subperiods, 1993–99, 2000–2006, a shorter period covering the crisis, 2007–2010, and the post-crisis period of 2011–2023. By all four forecast evaluation criteria, our model outperforms the random walk in the three longer subperiods, but not during the global financial crisis period.

Possible explanations for the weaker forecasting performance of our model in 2007–2010 include the large increase in demand for safe assets, or 'flight to quality', and the initial impacts of unexpected unconventional monetary policy measures following the collapse of Lehman Brothers. Caballero et al. (2017) showed that the supply of safe assets halved relative to world GDP between 2007 and 2011, while the share of U.S. government bonds within safe assets rose dramatically. They argue that a drop in US treasury yields and a rise in the value of the US dollar played significant roles in absorbing the increased demand for safe assets. The unexpected announcement of large-scale asset purchases, or quantitative easing, by the Federal Reserve after the collapse of Lehman Brothers in September 2008 might have altered the information content of the yield curve. Increased demand for government bonds and the introduction of various unconventional monetary policy measures also characterized some other countries. We hypothesize that a few years after the collapse of Lehman Brothers, market participants adjusted their expectations to the flight to quality phenomenon and unconventional monetary policy measures, which might explain why our excellent forecasting results was restored for the 2011–2023 period.

[20] Note that the 3-year trading simulation reported in Fig. 2 is a particular non-overlapping result, showing the cumulative value of an initial investment of 100 made in December 1989. The average annualized 3-year return results reported in Table 6 consider investments made in each month in December 1989 – December 1992, which average is significantly larger in the case of our models than the carry trade returns.
[21] 3-year ahead forecasts shorten the effective out-of-sample evaluation period by three years, so our first subsample starts in 1993.

6.4. Alternative models

Results for the various alternative models listed in Sections 4.2 and 4.3 are summarized in Table 7. For models using either 3-year or 5-year or 10-year maturity forward exchange rates or interest rates, we report the results from the equally weighted average of the three forecasts from the three versions corresponding to the alternative maturities.

Our baseline model (15) outperform the random walk in both short- and long-horizon forecasts for all three currency pairs (Table 7). Results for the GBP/USD rate are almost as good as for the DEM/USD rate, while for the GBP/DEM rate the improvement compared to the random walk is smaller, nevertheless the $MSFER_k$ ratio is below 100 for all forecast horizons.

In our analysis thus far, we have used the 1979–1989 sample period to form an initial estimate and evaluated our forecasts in the 1990–2023 period. The US dollar experienced large price fluctuations in 1980–1984. As a robustness test, we shorten the estimation sample period to start in January 1985, but continue to evaluate the out-of-sample forecasts in 1990–2023. Block B of Table 7 shows that forecasting results are slightly weaker for the DEM/USD and GBP/DEM rates, and considerably weaker for the GBP/USD rate, though they still beat the random walk. Our generally favorable results for the longer sample period are likely related to the ability of our model to capture adjustments to the equilibrium value of the exchange rate, which can be better estimated using longer sample periods. We thus estimate the remaining models in the sample starting in 1979.

The panel version of our baseline model, for which we force parameters θ_2 and θ_4 in model (15) to be common across currency pairs, but allow the intercepts, θ_1 and θ_3, to vary, leads to slightly better short-run forecasts and slightly worse long-run forecasts in the case of the two USD rates (block C of Table 7). For the DEM/GBP rate, short-run forecasts are slightly worse, while long-run forecasts are much better.

The VECM specification (Eq. (16)), block D of Table 7) is slightly better at 1-month and 3-month forecasting horizons than our baseline model, while our baseline specification is somewhat better at longer-horizon forecasts for the two dollar rates. The opposite holds for the GBP/DEM rate.

The VAR model (Eq. (17)), block E of Table 7), which uses the synthetic forward rate as the predictor but not in an error-correction framework, is worse than our baseline models, especially for short-horizon forecasts. Long-horizon forecasts are impressive from the VAR model for the two dollar rates, but poor for the GBP/DEM rate. These findings suggest that the forecasting ability of the synthetic forward rate comes from the error-correction mechanism we established in Section 3, while using this variable in an unconstrained form leads to poorer forecasts.

The simple autoregression (Eq. (18)), block F of Table 7) forecasts well the two dollar rates (but less so than our baseline model), while it leads to poor results for the GBP/DEM rate. For the two dollar rates, the autoregression beats the random walk at least in longer forecasting horizons. This suggests that some spot exchange rates of the major advanced country currencies are not really random walks, confirming our intuition from the one-period regression in Section 5.2.[22]

The model that includes the spot exchange rate and the interest rate differential as separate predictors (Eq. (19)), block G of Table 7) results in rather weak forecasts at forecasting horizons up to one year and good forecasts for longer horizons in the case of the two dollar rates. This finding highlights the importance of imposing the restriction implied by the long maturity synthetic forward rate on the linear combination of the spot exchange rate and the interest rate differential, in line with our theoretical derivation. It seems that allowing unconstrained parameters in the estimation can lead to noise, which harms forecasts. However, this conclusion should be nuanced when considering the GBP/DEM rate, for which short-horizon forecasts are weaker than our baseline, but long-horizon forecasts are excellent.

The model which uses the interest rate alone as the predictor (Eq. (20)), block H of Table 7) is clearly the worst of all models studied. This finding highlights that our baseline model (15) markedly differs from a model that uses interest rates only.

The real exchange rate model (Eq. (21)), block I of Table 7) does a good forecasting job, but it is inferior to our baseline model at all forecasting horizon for the two dollar rates. For the GBP/DEM rate, the real exchange rate model is also worse than our baseline model for short-horizon forecasts, but much better in long-horizon forecasts. The good forecasting performance of the real exchange rate model at longer forecasting horizons is in line with the findings of Cheung et al (2019), Ca'Zorzi and Rubaszek (2020), and Eichenbaum et al (2021).

The model that includes the spot exchange rate and the price level differential separately (Eq. (22), block J of Table 7) performs worse than the real exchange rate model (Eq. (21), block I of Table 7). This suggests that the theoretical restriction imposed by the real exchange rate is important, while allowing unconstrained parameters in the estimation can lead to noise, which harms forecasts.

Finally, the synthetic forward exchange rate as the prediction leads to MSFE ratios above 100 at all forecasting horizons for the

[22] Engel and Wu (2023) studied the forecasting performance of the autoregression for major currencies. Their work differs from ours in three major aspects. First, for multi-period ahead forecasts, they consider the long-horizon regression, $s_{t+k} - s_t = \theta_1 + \theta_2 \bullet s_t + \varepsilon_{t+k,k}$ with $k>1$, which suffers from estimation problems related to overlapping observations. We consider only the case of $k=1$ (for all our models). Second, they use a rather short, 5-year long, rolling sample size to estimate the model parameters. Our baseline forecasts use a recursive estimation, for which the initial 11-year-long estimation sample is gradually extended to 44 years. For both reasons, our parameter estimates are less subject to biases than their estimates. Nevertheless, we use the same bootstrap technique to test the statistical significance of forecasts as they do. Third, our out-of-sample forecast evaluation period, 1990–2023, is more than twice as long as the 2004–2020 evaluation period of Engel and Wu (2023). Related to the second difference, we emphasize that for our baseline model (Eq. (15)), which is not studied by Engel and Wu (2023)), we find that either the recursive estimation technique or relatively long (e.g. 20 years) rolling windows work the best, but our forecasts are weak when only a 5-year rolling window is used to estimate the parameters; see in Section 6.5.

DEM/USD rate, all but one forecasting horizons for the GBP/DEM rate, while the below-100 ratios for the GBP/USD rate are not statistically significant (block K of Table 7). In other words, UIP forecasts poorly. This finding confirms our intuition suggesting that while forward exchange rates are not unbiased predictors of future spot exchange rates, long-maturity forward rates can be used efficiently, in our error correction framework (Eqs. (15) and (16)), to forecast future exchange rate changes for both short and long forecast horizons.

6.5. Recursive versus rolling estimation

Up to this point, we have used recursive estimation for our forecasts, whereby we extended the last observation of the estimation sample by one month for each forecast round. We test the robustness of the forecasts from our baseline model to rolling estimation samples with varying lengths between 60 and 240 months. Table 8 (for the DEM/USD rates) and Table A2 in the Supplementary material (for the other two currency pairs) show forecasting results are weak when relatively short rolling windows are used for estimation and are conversely stronger for relatively longer rolling windows. In the case of the DEM/USD rate, the 220- and 240-month rolling window estimations led to somewhat better results than our benchmark results based on the recursive estimation technique for all forecasting horizons, while the 200-month rolling window is slightly better than the recursive estimation for longer (but not for shorter) forecast horizons.

For the GBP/USD and GBP/DEM rates, the recursive estimation scheme strongly dominates all rolling estimation schemes in terms of forecast accuracy (Table A2, Supplementary material). The excellent forecasting results with the recursive scheme, and the improved forecast accuracy for longer rolling estimation windows compared to shorter rolling windows, indicate our model incorporates longer-run tendencies in the form of an error-correction relationship, which can better be captured over longer periods.

7. Summary

This article presented a new model based on a novel combination of the general class of theoretical exchange rate models analyzed by Engel and West (2005) and the error-correction forecasting equation of Mark (1995) and Chinn and Meese (1995). We showed, using the example of the general form of money–income models, that in the exchange rate determination equation, the unobserved expected future exchange rate, $E_t[s_{t+n}]$, can be substituted out with the use of the n-period forward exchange rate, $d_t^{(n)}$, which is an observed variable at time t, for any maturity n. Moreover, we demonstrated that the gap between the fundamental equilibrium exchange rate and the actual exchange rate can be approximated with the forward exchange rate when its maturity is relatively long, such as five or ten years (large n), but not when the maturity of the forward rate is short, such as one month (small n). Since an error-correction model has the potential to forecast at both short and long-horizons, our theoretical derivation suggests that the long-maturity forward exchange rate has the potential to forecast the change in the spot exchange rate at any forecast horizon. For example, we insert the 10-year maturity forward exchange rate into the error-correction model to forecast the change in the spot exchange rate at all horizons between one month and five years. This error-correction forecast based on the long-maturity forward rate differs from the prediction of uncovered interest rate parity, which would use the 1-month forward rate to predict the 1-month-ahead exchange rate, the 2-month forward rate to predict the 2-month-ahead exchange rate, and so on.

Using four forecast evaluation criteria in the 1990–2023 period to evaluate out-of-sample forecasts, we find that our new model significantly outperforms the driftless random walk in both short and long-horizon forecasting for the three bilateral pairs of the US dollar, German mark (euro since 1999), and British pound sterling, which three currency pairs account for 37 % of the global foreign exchange market turnover. The forecast accuracy of our models improves with the forecast horizon and is economically significant. Our model leads to impressive direction of change forecasting results, relative to both the spot exchange rate and the forward exchange rate, while a trading simulation based on our model forecasts leads to economically large excess returns.

Our results are statistically significant and robust to alternate sample periods, single currency pairs or panel estimation, specification of the error correction model, and recursive versus rolling sample estimation provided that long rolling windows are used. Our baseline model outperforms several alternative models, such as the autoregression, the interest rate differential as the predictor, the real exchange rate as the predictor, and models in which the components of the forward exchange rate and the real exchange rate are allowed to have different parameters. The comparison of model forecasts highlights the importance of including the synthetic forward exchange rate in an error-correction specification, in line with our theoretical derivation. Our forecasting results are markedly better when using long estimation samples, either the recursive estimation technique or relatively long rolling windows, such as 20 years, than when using shorter rolling estimation windows. The improved forecast accuracy for longer estimation windows likely indicates our model incorporates longer-run tendencies in the form of an error-correction relationship, which can better be captured over longer periods.

Our forecasting results exhibit a marked improvement over past works with a novel model that is based on a novel theoretical justification. Besides, in contrast to various overly complicated models in the literature, our forecasting model is extremely simple, linear, easy to replicate, and the data we use are available in real time and not subject to revisions.

Our forecasting results suggest a close relationship between the expected components of long-term yields and the expected path of the exchange rate, with implications for exchange rate theories. Further research should develop theoretical models incorporating this relationship.

Data availability

Data is downloadable among supplementary materials.

Acknowledgements

We are thankful for comments and suggestions to Charles Engel, Philippe Bacchetta, Menzie D. Chinn, Todd E. Clark, Andrew Fodor, Simon Gilchrist, Pierre-Olivier Gourinchas, James D. Hamilton, Henk Jager, Franc Klaassen, Zalán Kocsis, Gábor Kőrösi, Gábor Rappai, Andrew K. Rose, Pierre L. Siklos, András Simon, Mark P. Taylor, Timo Teräsvirta, Casper de Vries, Ákos Valentinyi, anonymous reviewers, and conference and seminar participants at the 11th International Conference on Macroeconomic Analysis and International Finance (Rethymno, 2007), Central Bank of Hungary (Budapest, 2007), the 5th INFINITI Conference on International Finance (Dublin, 2007), Corvinus University of Budapest (2007 and 2020), University of Pécs (2007 and 2014), Royal Economic Society Annual Conference (Warwick, 2008), Tinbergen Institute Rotterdam (2008), the Bank for International Settlements (Basel, 2008), ECOMOD (Berlin, 2008), NAKE Research Day (Utrecht, 2008), Hungarian Society of Economics (Budapest, 2008 and 2017), the 24th Annual Congress of the European Economic Association (Barcelona, 2009), the 44th Annual Conference of the Canadian Economic Association (Quebec City, 2010), the 34th International Symposium on Forecasting (Rotterdam, 2014), the 12th INFINITI Conference on International Finance (Prato, 2014), the 5th World Finance Conference (Venice, 2014), the 6th Annual Conference of the International Association for Applied Econometrics (Nicosia, 2019), and the 2nd Vienna Workshop on Economic Forecasting (2020). Zsolt Darvas was affiliated with Argenta ZRt when part of this research was conducted.

Funding

Financial support from OTKA (Hungarian Scientific Research Fund) Grant No. K 61221, TÁMOP (Social Renewal Operational Programme) Grant N.o. 4.2.2.C-11/1/KONV-2012-0005 and Project no. TKP2021-NKTA-19 under the TKP2021-NKTA funding scheme of the National Research, Development and Innovation Fund of Hungary are gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jimonfin.2024.103067.

References

Berg, K.A., Mark, N.C., 2019. Where's the risk: the forward premium bias, the carry-trade premium, and risk-reversals in general equilibrium. J. Int. Money Financ. 95, 297–316.
Berkowitz, J., Giorgianni, L., 2001. Long-horizon exchange rate predictability? Rev. Econ. Stat. 83 (1), 81–91.
Bianco, D., Marcos, M.C., Quiros, G.P., 2012. Short-run forecasting of the euro-dollar exchange rate with economic fundamentals. J. Int. Money Financ. 31 (2), 377–396.
Ca'Zorzi, M., Rubaszek, M., 2020. Exchange rate forecasting on a napkin. J. Int. Money Financ. 104.
Caballero, R.J., Fahri, E., Gourinchas, P.-O., 2017. The safe assets shortage conundrum. J. Econ. Perspect. 31 (3), 29–46.
Cerra, V., Saxena, S.C., 2010. The monetary model strikes back: evidence from the world. J. Int. Econ. 81 (2), 184–196.
Chen, Y.-C., Tsang, K.P., 2013. What does the yield curve tell us about exchange rate predictability? Rev. Econ. Stat. 95 (1), 185–205.
Cheung, Y.-W., Chinn, M.D., Pascual, A.G., Zhang, Y., 2019. Exchange rate prediction redux: New models, new data, new currencies. J. Int. Money Financ. 95 (1), 332–362.
Chinn, Menzie D., and Guy Meredith. 2005. "Testing Uncovered Interest Rate Parity at Short and Long Horizons during the Post-Bretton Woods Era". NBER Working Paper no. 11077. https://www.nber.org/papers/w11077.
Chinn, M.D., Meese, R.A., 1995. Banking on currency forecasts: How predictable is change in money? J. Int. Econ. 38, 161–178.
Chinn, M.D., Moore, M.J., 2011. Order flow and the monetary model of exchange rates: evidence from a novel data set. J. Money Credit Bank. 43 (8), 1599–1624.
Chinn, M.D., Quayyum, S.A., 2012. Long Horizon Uncovered Interest Parity Re-Assessed NBER Working Paper no. 18482.
Clarida, R.H., Taylor, M.P., 1997. The term structure of forward exchange rate premiums and the forecastability of spot exchange rates: correcting the errors. Rev. Econ. Stat. 79 (3), 353–361.
Clark, T.E., West, K.D., 2006. Using out-of-sample mean squared prediction errors to test the martingale difference hypothesis. J. Econ. 135 (1), 155–186.
Clark, T.E., West, K.D., 2007. Approximately normal tests for equal predictive accuracy in nested models. J. Econ. 138 (1), 291–311.
Darvas, Z., 2008. Estimation bias and inference in overlapping autoregressions: implications for the target zone literature. Oxf. Bull. Econ. Stat. 70 (1), 1–22.
Darvas, Z., 2009. Leveraged carry trade portfolios. J. Bank. Financ. 33 (5), 944–957.
Darvas, Z., Z. Schepp. 2007. "Forecasting exchange rates of major currencies with long maturity forward rates." Working Paper 2007/5, Department of Mathematical Economics and Economic Analysis, Corvinus University of Budapest. https://ideas.repec.org/p/mkg/wpaper/0705.html.
Darvas, Z., Schepp, Z., 2009. Long maturity forward rates of major currencies are stationary. Appl. Econ. Lett. 16 (11), 1175–1181.
Dickey, D., Fuller, W.A., 1979. Distribution of the estimators for autoregressive time series with a unit root. J. Am. Stat. Assoc. 74 (366), 427–431.
Eichenbaum, M.S., Johannsen, B.K., Rebelo, S.T., 2021. Monetary policy and predictability of nominal exchange rates. Rev. Econ. Stud. 88, 192–228.
Elliott, G., Rothenberg, T.J., Stock, J.H., 1996. Efficient tests for an autoregressive unit root. Econometrica 64 (4), 813–836.
Engel, C., 2014. Exchange rates and interest parity. In: Gopinath, G., Helpman, E., Rogoff, K. (Eds.), Handbook of International Economics, Vol. 4. North Holland, Amsterdam, pp. 453–522.
Engel, C., Wu, S.P.Y., 2023. Forecasting the U.S. Dollar in the 21st Century. J. Int. Econ. https://doi.org/10.1016/j.jinteco.2023.103715.
Engel, C., Mark, N.C., West, K.D., 2007. Exchange Rate Models Are Not as Bad as You Think. In: NBER Macroeconomics Annual, vol. 22, pp. 381–441.
Engel, C., West, K.D., 2005. Exchange rates and fundamentals. J. Polit. Econ. 113 (3), 485–517.
Faust, J., Rogers, J.H., Wright, J.H., 2003. Exchange rate forecasting: the errors we've really made. J. Int. Econ. 60 (1), 35–59.
Gourinchas, P.-O., Rey, H., 2007. International financial adjustment. J. Polit. Econ. 115 (4), 665–713.

Ince, O., Molodtsova, T., Papell, D.H., 2016. Taylor rule deviations and out-of-sample exchange rate predictability. J. Int. Money Financ. 69 (1), 22–44.
Kilian, L., 1999. Exchange rates and monetary fundamentals: what do we learn from long-horizon regressions? J. Appl. Economet. 14 (5), 491–510.
Lustig, H., Stathopoulos, A., Verdelhan, A., 2019. The term structure of currency Carry Trade Risk Premia. Am. Econ. Rev. 109 (12), 4142–4177.
Mark, N.C., 1995. Exchange rates and fundamentals: evidence on long-horizon predictability. Am. Econ. Rev. 85 (1), 201–218.
McCracken, M.W., Sapp, S.G., 2005. Evaluating the predictability of exchange rates using long-horizon regressions: mind your p's and q's! J. Money Credit Bank. 37 (3), 473–494.
Meese, R.A., Rogoff, K., 1983. Empirical exchange rate models of the seventies: do they fit out of sample? J. Int. Econ. 14 (1), 3–24.
Newey, W.K., West, K.D., 1987. A simple, positive, semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55 (3), 703–708.
Ng, S., Perron, P., 2001. Lag length selection and the construction of unit root tests with good size and power. Econometrica 69 (6), 1519–1554.
Park, C., Park, S., 2013. Exchange rate predictability and a monetary model with time-varying cointegration coefficients. J. Int. Money Financ. 37 (1), 394–410.
Perron, P., Ng, S., 1996. Useful modifications to some unit root tests with dependent errors and their local asymptotic properties. Rev. Econ. Stud. 63 (3), 435–463.
Pesaran, H.M., Timmermann, A., 1992. A simple nonparametric test of predictive performance. J. Bus. Econ. Stat. 10 (4), 461–465.
Phillips, Peter, C.B., Perron, P., 1988. Testing for a unit root in time series regression. Biometrika 75 (2), 335–346.
Pincheira, P.M., West, K.D., 2016. A comparison of some out-of-sample tests of predictability in iterated multi-step-ahead forecasts. Res. Econ. 70 (2), 304–319.
Rossi, B., 2007. Expectation hypothesis tests at long horizons. Econometr. J. 10 (3), 554–579.
Rossi, B., 2013. Exchange rate predictability. J. Econ. Lit. 51 (4), 1063–1119.
Sarno, L., Taylor, M.P., 2002. The Economics of Exchange Rates. Cambridge University Press, Cambridge, MA.
Timmermann, A., 2006. Forecast combinations. In: Elliott, G., Granger, C.W.J., Timmermann, A. (Eds.), Handbook of Economic Forecasting, Volume 1. North Holland, Amsterdam, pp. 136–196.
Valchev, R., 2020. Bond convenience yields and exchange rate dynamics. Am. Econ. J. Macroecon. 12 (2), 124–166.
Wright, J.H., 2008. Bayesian model averaging and exchange rate forecasts. J. Econ. 146 (2), 329–341.