



Forecasting exchange rates better than the random walk thanks to machine learning techniques

Christophe Amat, Tomasz Michalski, Gilles Stoltz

► To cite this version:

Christophe Amat, Tomasz Michalski, Gilles Stoltz. Forecasting exchange rates better than the random walk thanks to machine learning techniques. 2014. halshs-01003914v1

HAL Id: halshs-01003914

<https://shs.hal.science/halshs-01003914v1>

Preprint submitted on 10 Jun 2014 (v1), last revised 28 May 2018 (v6)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

FORECASTING EXCHANGE RATES BETTER THAN THE RANDOM WALK THANKS TO MACHINE LEARNING TECHNIQUES*

Christophe Amat

Tomasz Michalski

Gilles Stoltz

Ecole des Ponts ParisTech
Marne-la-Vallée, France

GREGHEC: HEC Paris – CNRS
Jouy-en-Josas, France

christophe.amat@polytechnique.edu

michalski@hec.fr

stoltz@hec.fr

June 10, 2014

Abstract

Simple exchange rate models based on economic fundamentals were shown to have a difficulty in beating the random walk when predicting the exchange rates out of sample in the modern floating era. Using methods from machine learning – adaptive sequential ridge regression with discount factors – that prevent overfitting in-sample for better and more stable forecasting performance out-of-sample we show that fundamentals from the PPP, UIRP and monetary models consistently improve the accuracy of exchange rate forecasts for major currencies over the floating period era 1973–2013 and are able to beat the random walk prediction giving up to 5 % improvements in terms of the RMSE at a 1 month forecast. “Classic” fundamentals hence contain useful information about exchange rates even for short forecasting horizons - and the Meese and Rogoff [1983] puzzle is overturned. Such conclusions cannot be obtained when rolling or recursive OLS regressions are used as is common in the literature.

JEL codes: C53, F31, F37

Keywords: exchange rates, forecasting, machine learning, purchasing power parity, uncovered interest rate parity, monetary exchange rate models

*We would like to thank Refet Gürkaynak, Robert Kollmann, Jose Lopez, Evren Örs, Cavit Pakel, Romain Rancière, Helene Rey and the seminar participants at Bilkent University and HEC Paris for helpful comments and discussions. All remaining errors are ours. We would like to thank Investissements d’Avenir (ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047) for financial support. Corresponding author: Tomasz Michalski, HEC Paris, 1, rue de la Libération, F-78351 Jouy-en-Josas Cedex. Phone: +33(0)139677240. E-mail: michalski@hec.fr.

1. Introduction

In this paper we reconsider the ability of simple models based on fundamentals to forecast exchange rates using statistical techniques borrowed from machine learning – adaptive sequential ridge regression with discount factors. We consider “traditional” fundamentals from the purchasing power parity models (PPP), uncovered interest rate parity (UIRP) and two standard monetary models and show that *all* beat the random walk without drift in forecasting forward the 1-month exchange rate for *all* currencies that we consider in the period 1973–2013 using the RMSE criterion. Statistical significance of this improvement depends on the test considered. For the Clark-West test (see Clark and West [2006]), we obtain that all of our forecasts are significantly better at a 10 % significance level than that of a random walk; for the Diebold-Mariano-West test (see Diebold and Mariano [1995] and West [1996]) we can find at least one fundamental for each considered currency that beats the random walk in a statistically significant way at a 10 % level. For 3 out of 7 currency pairs considered (USD/GBP, JPY/USD, SEK/USD) we obtain forecasts that beat the random walk at the 5 % significance level no matter what test is used and what fundamental is considered. We obtain similar success with predicting the direction of change of the exchange rates. We conduct several robustness checks – such as trying out different (sub)samples, variations on the exact form of the fundamentals included etc. and our results are upheld. Furthermore, investigation of statistical properties of the errors shows that our improvement is uniform rather than driven by a few well predicted instances. Our conclusion is therefore that “classic” fundamentals hence contain useful information about exchange rates even at short forecasting horizons.

These findings are contrary to the consensus findings in the literature started by the famous Meese and Rogoff [1983] result of the inability of the “classic” fundamentals based on PPP, UIRP or monetary models from the 1970s to

outperform the random walk for short time periods (see Rossi [2013] for a recent comprehensive review). One exception is that of Clark and West [2006] that demonstrate the predictability of a UIRP-based model (though with less sharp results than the ones obtained by us). We also consider exchange rate models based on Taylor-rule fundamentals that were found recently to assure predictability at the short time horizon (Engel and West [2006], Molodtsova and Papell [2009], Molodtsova et al. [2011], Giacomini and Rossi [2010], Rossi and Inoue [2012] though Rogoff and Stavrakeva [2008] disagree) but do not find that these perform much better in terms of forecasting and sometimes give strictly worse results¹. Our study is not comprehensive nor exhaustive in nature. Our goal is not to provide the best forecasts of exchange rates possible by finding fundamentals from a wide menu that was considered in the vast literature that have the strongest predictive power (or aggregating the information contained therein) but rather to show the usefulness of methods of machine learning in application to well-known economic problems.

A seminal recent study on predictability at a short horizon is Molodtsova and Papell [2009] that considered the ability of a model with Taylor rule reactions of central banks to deliver exchange rate predictability for 1 month horizons. Predictability, however, is a different concept than forecastability. In Molodtsova and Papell [2009] it is about testing whether the estimated coefficients of a considered model are jointly significantly different from zero when explaining changes in the exchange rate. It does not mean that a model that exhibits predictability necessarily provides better forecasts. The focus of these and many other attempts in general is rather to assess whether fundamentals play a role in exchange rate determination as it can be motivated theoretically why the forecasts they produce in terms of an evaluation criterion such as root mean square errors (RMSE) may fare worse than those of a random walk process (see

¹We do not use any models based on the evolution of net foreign assets such as Gourinchas and Rey [2007] because the fundamentals to conduct these tests are available at 3-month frequencies resulting in fewer observations that can be used.

Rossi [2013] for a discussion). In this paper, however, we are not interested in such understood predictability – we actually are interested whether we can produce better forecasts of the exchange rates than the random walk without drift (and the predictability follows). We are not interested in validating a particular model – that is trying to fit the coefficients and check whether the signs and magnitudes are that those posited by theories; we simply try to extract from the fundamentals in question information useful for the behavior of exchange rates. Our methods allow us to ascertain that the “classic” fundamentals contain valuable information about the behavior of exchange rates in the short run allowing to produce short-term forecasts beating the random walk. In particular, the good performance of the UIP relationship is encouraging also for practical reasons: interest rate data is easily available without any lags so can be effectively used for forecasting.

The intuition for the methods we use is the following. Our basic method – adaptive sequential ridge regression with discount factors – prevents overfitting in-sample for better and more stable performance out-of-sample. This is achieved by including a regularization term in an otherwise standard ordinary least squares (OLS) problem that prevents wild swings of coefficients when reestimating. We also weigh the observations that are included in the estimation by a discounting factor – which allows for our algorithm to accommodate quickly any structural breaks that might be present in the data (such as changes in the conduct of monetary policy, crisis times etc.). It is to be stressed that both these terms are estimated from the data up to the period for which the forecast is given rather than exogenously imposed². The method is hence entirely data-driven and does not rely on some educated guesses about important parameters as for example in Wright [2008] for Bayesian Model Averaging methods. What is important, this method also exhibits the following theoretical guarantee of

²Unsurprisingly, models that we call “oracles” – acting as if these optimal parameters were known beforehand – perform better.

performance, described in Section A: its RMSE converges, as the number of instances to be predicted increases, to the RMSE of the best fixed linear model (which is in particular weakly smaller than the RMSE of the random walk). One can prove that no such guarantee can be achieved for (recursive or rolling) OLS. We note that this adaptive sequential ridge regression with discount factors method was not specifically designed for the forecasting of exchange rates but is a general method that proved to work well in other problems (forecasting of air quality, electricity consumption or the production of oil reservoirs; see the monograph Cesa-Bianchi and Lugosi [2006] as well as the papers Mauricette et al. [2009] and Stoltz [2010]).

In our study, apart from the particular statistical methods, we used data and forecast evaluation methods that are standard in the literature as discussed in Rossi [2013]. We use lagged fundamentals and do not detrend, filter or seasonally adjust the data in any way. We use single-equation methods. Our benchmark for comparison is the random walk without drift. Our main sample is 1973–2013. We also use an auxiliary sample 1973–2006 taken from Molodtsova and Papell [2009] to allow for comparability with the Taylor-rule models they suggest and not to be accused of data snooping. Second, we have two samples that include the pre- and post-Great Recession period so that we can see whether our methods can cope effectively with forecasting in crisis periods with nonconventional monetary policies applied by central banks.

There are several themes discussed by Rossi [2013] that hold in our study. Traditional linear estimation methods (rolling or recursive OLS) with our fundamentals fail to forecast exchange rates better than the random walk. With the addition of several fundamentals (aggregating information) we do not get gains in the precision of predictions – hence parsimonious models work well. Fundamentals contemporaneous with the forecast period (actually realized values, as used in Meese and Rogoff [1983]) perform poorly no matter what is the model considered. We find also that a simple model based on the random walk with

“drift” (with the drift parameter estimated using our method) does a very good job in forecasting. A general interpretation of such a drift parameter is actually difficult: given the data-based choice of the regularization parameter and the discount factor it may not represent a long run trend but actually a short one with short-memory – which for a given currency pair may change depending on the exact forecasting time period in our sample. Combining our “classic” experts together with such a “drift” (so estimating the models in the affine and not the linear form) does not lead necessarily to better forecasts, however. Fundamentals seem to contain thus similar information as the estimated drift parameter. One possible explanation of this phenomenon is that fundamentals may contain the information about the drift themselves. The Swiss–U.S. currency pair studied in our sample may serve as an example (see Figure 1): a permanently lower inflation rate in Switzerland than in the USA and the constant appreciation of the CHF/USD exchange rate go hand in hand. We ask thus whether the fundamentals provide some other information on the behavior of the exchange rate beyond that included in the “drift”. To test this we use a two-step procedure to test the predictability of the fundamentals against the random walk with drift (estimated via our adaptive sequential ridge regression with discount factors). We obtain as a result weak evidence of predictability of monetary and Taylor-rule based experts versus the model with the “drift”.

The improvements in the forecasts using fundamentals we document (up to 5 % in terms of RMSE) are not large; the fundamentals we consider do not help to add a lot of predictive content relative to the random walk. The observations of Engel and West [2005] may well be valid: current and past fundamentals may have low correlations with future exchange rate realizations. Fundamentals may have little predictive power against an exchange rate that can be approximated by a random walk. What we show is that the basic fundamentals considered in the literature still do have *some* predictive power even at short horizons which was deemed a lost cause in international economics and may be useful for con-

ducting policy-oriented exchange rate forecasts. A valid question arises whether the gains from prediction that we obtain could constitute an important source of utility improvements for investors using portfolio investment strategies following our forecasts – akin to the exercise undertaken by Corte et al. [2012] for the Gourinchas and Rey [2007] model.

1.1. Related literature

Rossi [2013] provides a comprehensive recent review of the literature so we keep ours at a bare minimum. There were different ways in which researchers tried to cope with the negative result of Meese and Rogoff [1983] that showed that the simple exchange rate models from the 1970s did poorly in comparison to a random walk without drift in forecasting the exchange rates in the floating period after 1973 for short forecast horizons. One way, that we pursue here, was to use better econometric tools to extract information from the data. Some of the solutions proposed involved for example cointegration techniques as in Mark [1995] combined with the usage of panel data such as in Mark and Sul [2001], long samples in panels as in Rapach and Wohar [2002] and including a large set of countries as in Cerra and Saxena [2010]. These techniques typically are used for longer-horizon (above 1 year) forecasts when it is believed that the long-run relationship as modeled by the cointegrating equations kicks in. We focus on short-term forecasts and use single-equation methods. Two approaches related to ours are nonlinear models and models with time-varying parameters. Rossi [2013] discusses that different nonlinear methods were not particularly successful in forecasting exchange rates while Rossi [2006] questions the robustness of the time-varying parameter models. Bacchetta et al. [2010] argue that the gain from using such an approach would be practically minimal. These authors find on simulated data that the benefits from using such models in terms of greater explanatory power are in practice outweighed by additional estimation errors

of time-varying parameters. Schinasi and Swamy [1989] reassess the study of Meese and Rogoff [1983] using various nonlinear methods, including an early version of a ridge regression. In this paper we use ridge regression methods developed and improved afterwards (a sequential ridge regression with discount factors). Engel [1994] documents a failure of a Markov-switching model to beat the random walk in forecasting. Perhaps the closest effort to ours – insofar as the idea of extracting information from fundamentals is concerned – is Wright [2008] that uses Bayesian Model Averaging methods. However, his method combines a great number of predictors many of which do not come from the standard models considered in this paper. His approach does not improve the forecasts upon a random walk in a statistically significant way in most cases he considers. Moreover, it is not clear how to choose the “shrinkage” parameter that guards the informativeness of priors properly not knowing the properties of the data *ex ante*. Our method, in contrast, is entirely data driven.

Another way that researchers tried to improve the ability of “fundamentals” to forecast exchange rates is to consider different models. It turned out in the most recent years that exchange rate models based on Taylor-rule fundamentals perform well in ascertaining the predictability of the exchange rates at a short horizon (Molodtsova and Papell [2009]). For this reason we try out our methods on the same fundamentals though we do not find them better (and in fact they fare worse) than the “classic” fundamentals. Another successful fundamental was the behavior of net foreign assets as in Gourinchas and Rey [2007] or Corte et al. [2012]. The fundamentals to conduct these tests are available at 3-month frequencies resulting in fewer observations that can be used so we do not investigate them here. Other studies assessed the forecasting ability of exchange rate models of the 1990s such as Cheung et al. [2005], or differences in the term structures of forward premia such as Clarida et al. [2003]. Given the scope of our exercise, we did not evaluate these models with the machine-learning methods, but it may be a useful research agenda for the future. Some other attempts, that

are less relevant for our current study involved work on what data was available at what instant while forming the forecasts.

1.2. Organization of the paper

The paper is organized as follows. First, in Section 2 we lay out the fundamentals that we shall consider. Then, in Section 3 we discuss the methods that we use in the analysis and their difference with respect to the standard linear regressions. Next, in Section 4 we discuss the data while in Section 5 we discuss the results. Section 6 concludes. An Appendix contains further details regarding the methods and results tables.

2. Considered fundamentals

Given the heuristics of the statistical methods we use, we need “experts” that are going to predict the exchange rate in the future periods. Fundamentals from simple exchange rate models are going to serve for this purpose. In what follows, we will use “fundamentals” and “experts” interchangeably.

The first fundamentals are going to be the inflation differentials, coming from the relative purchasing power parity (PPP) model, where a change in the exchange rate is given by

$$s_{t+1} - s_t = \alpha + \beta (\pi_{t+1,t} - \pi_{t+1,t}^*) \quad (1)$$

where s_t is the logarithm of the exchange rate (home currency units per unit of the foreign currency) at time t , $\pi_{t+1,t}$ is the inflation rate between time $t, t+1$, α and β are parameters and the “*” denotes the variables for the foreign country.

Next, we consider the uncovered interest rate parity (UIRP) model. The

exchange rate change is given by

$$s_{t+1} - s_t = \alpha + \beta (i_t - i_t^*) \quad (2)$$

where i is the short run interest rate.

Then, we consider the simple flexible-price monetary model of the 1970s vintage (Frenkel-Bilson)

$$s_t = \alpha + \beta_1 (m_t - m_t^*) + \beta_2 (y_t - y_t^*) \quad (3)$$

where m is the logarithm of the money stock and y is the logarithm of output. After differencing (3) we obtain

$$s_{t+1} - s_t = \beta_1 (\widehat{m}_{t+1,t} - \widehat{m}_{t+1,t}^*) + \beta_2 (\widehat{y}_{t+1,t} - \widehat{y}_{t+1,t}^*) \quad (4)$$

where $\widehat{x}_{t+1,t}$ over a variable x denotes the change in the logarithms between periods t and $t+1$. We also consider a more extensive version of this model including interest rate differentials

$$s_t = \alpha + \beta_1 (m_t - m_t^*) + \beta_2 (y_t - y_t^*) + \beta_3 (i_t - i_t^*) \quad (5)$$

which after differencing gives an additional expert – changes in interest rate differentials.

Hence, we use as experts: (i) inflation differentials $(\pi_{t+1,t} - \pi_{t+1,t}^*)$ or (ii) interest rate differentials $(i_t - i_t^*)$ or (iii) differences in money stock growth $(\widehat{m}_{t+1,t} - \widehat{m}_{t+1,t}^*)$, the differences in the output growth $(\widehat{y}_{t+1,t} - \widehat{y}_{t+1,t}^*)$ and changes in interest rate differentials $(i_{t+1} - i_{t+1}^*) - (i_t - i_t^*)$. We also investigate whether all the fundamentals mentioned above taken together are able to predict changes in the exchange rate³.

³The form in which the fundamentals are included does not seem to matter. We reran the PPP and monetary models also using price or output indices in levels or the monetary stock and

Finally, we also consider whether fundamentals from a Taylor-rule based exchange rate model are useful in forecasting as well. We use equation (7) from Molodtsova and Papell [2009]

$$s_{t+1} - s_t = \alpha + \beta_1 (\pi_t^* - \pi_t) + \beta_2 (\tilde{y}_t^* - \tilde{y}_t) + \beta_3 (i_{t-1}^* - i_{t-1}) \quad (6)$$

where \tilde{y} is a measure of the output gap. Therefore, our Taylor-rule experts are the inflation differentials, output gap differentials and lagged interest rate differentials.

We estimate the above models in the linear (without α , implying that there should be no change in the exchange rate if the fundamentals do not indicate a change) or in the affine form (with a possible “drift”). In our estimations we use lagged fundamentals available up to time t to forecast the changes in the exchange rates for the period t to $t + 1$. The exact data series used are further discussed in Section 4.

3. Combining macro-economic variables with regularized linear regressions

The past section indicated that several macro-economic variables were of interest while forecasting exchange rates. We will use a generic notation to refer to each of them, indexing them by subscripts $j \in \{1, 2, \dots\}$ instead of indexing them with different (Greek or Roman) letters. Thus, $f_{j,t+1}$ will denote⁴ (depending on the value of j) the quantities $\pi_{t+1,t} - \pi_{t+1,t}^*$, or $i_t - i_t^*$, etc.

We will restrict our attention to linear (not affine) forecasts, and output forecast our qualitative conclusions do not change.

⁴We assume here implicitly that the origin of the forecasting period was set such that fundamental forecasts for all macro-economic variables are available already at month $t = 0$, the month before the first month of the period.

casts \widehat{s}_{t+1} for s_t of the form

$$\widehat{s}_{t+1} - s_t = \sum_{j=1}^N u_{j,t+1} f_{j,t+1}, \quad (7)$$

where the summation is to be understood on the N fundamentals considered (their choice is up to us). The question is how to set the weights $u_{j,t+1}$ depending on the past. (Note, for later purposes, see Section 5.6.1, that adding extra constant fundamental forecasts $f_{0,t+1} \equiv 0.01$, one can create affine forecasts. We do not do so for the time being.)

We will review several methods to do so: the classical recursive and rolling regressions, and a family of regressions stemming from machine learning — sequential ridge regressions.

3.1. How to pick the weights (the aggregation strategies)

We denote by $\mathbf{u}_{t+1} = (u_{j,t+1})_{j \leq N}$ the vector of the linear weights to be picked for forecasting the rate of month $t + 1$.

Random walk. The first strategy consists of choosing $\mathbf{u}_{t+1} = (0, \dots, 0)$ at each round $t + 1$, that is, of forecasting \widehat{s}_{t+1} by s_t . We call it the random-walk strategy.

Recursive linear regressions. This is one of the most standard techniques in the literature. It consists of choosing, for months $t \geq 1$,

$$\mathbf{u}_{t+1} \in \arg \min_{\mathbf{u} \in \mathbb{R}^N} \sum_{\tau=1}^t \left(s_{\tau} - s_{\tau-1} - \sum_{j=1}^N u_j f_{j,\tau} \right)^2.$$

(The choice \mathbf{u}_1 is irrelevant as the root mean squared error discards anyway the forecast output for the first month, because of the existence of the training period.)

The issue with the chosen \mathbf{u}_{t+1} is that they tend to overfit past data, i.e., they lead to good in-sample predictions but poor out-of-sample ones. Little theoretical guarantees as to the out-of-sample performance of such plain linear regressions can be given.

Two ways to prevent this overfitting are the following:

- either consider only a fraction of the past data, typically the past H data points (in particular, if one believes that the exchange rate is a process with a memory bounded by H); or, in the same vein, put a higher weight on more recent observations;
- and/or add what is called a regularization term to the squared error, to help controlling (reducing) the range of the components of the vector \mathbf{u}_{t+1} .

We recall below two methods to do so: rolling regressions for the former way and the ridge regression for the latter way.

Rolling linear regressions. This is the other most standard technique in the literature. It consists of choosing, for months $t \geq 1$,

$$\mathbf{u}_{t+1} \in \arg \min_{\mathbf{u} \in \mathbb{R}^N} \sum_{\tau=\max\{1, t+1-H\}}^t \left(s_{\tau} - s_{\tau-1} - \sum_{j=1}^N u_j f_{j,\tau} \right)^2.$$

That is, at most H past instances are used to compute the linear regression (when $t \geq H$, then exactly the past H instances are used).

Adaptive sequential ridge regression. It was introduced by Hoerl and Kennard [1970] in a stochastic setting. What follows relies on recent new analyses of the ridge regression in the machine learning community; see Vovk [2001], Azoury and Warmuth [2001], Devaine et al. [2013]. The sequential ridge re-

gression with a constant regularization factor $\lambda \geq 0$ picks the weights

$$\mathbf{u}_{t+1}(\lambda) \in \arg \min_{\mathbf{u} \in \mathbb{R}^N} \left\{ \lambda \sum_{j=1}^N u_j^2 + \sum_{\tau=1}^t \left(s_\tau - s_{\tau-1} - \sum_{j=1}^N u_j f_{j,\tau} \right)^2 \right\}. \quad (8)$$

The recursive linear regression corresponds to the special case when $\lambda = 0$.

Vovk [2001], Azoury and Warmuth [2001] provide theoretical guarantees as to the out-of-sample performance of the sequential ridge regression, for the best regularization factor $\lambda \geq 0$; we state them in Appendix A. They entail in particular that its RMSE converges, as the number of instances to be predicted increases, to the RMSE of the best fixed linear model (which is in particular smaller than the RMSE of the random walk). However, determining ex ante this best constant regularization factor is impossible in general and therefore, Devaine et al. [2013] proposed an effective way of calibrating the regularization factors $\lambda_{t+1} \geq 0$ over time: at month $t + 1$, the vector used is

$$\mathbf{u}_{t+1}(\lambda_{t+1}) \in \arg \min_{\mathbf{u} \in \mathbb{R}^N} \left\{ \lambda_{t+1} \sum_{j=1}^N u_j^2 + \sum_{\tau=1}^t \left(s_\tau - s_{\tau-1} - \sum_{j=1}^N u_j f_{j,\tau} \right)^2 \right\}$$

where λ_{t+1} corresponds to the sequential ridge regression with constant regularization factor $\lambda > 0$ with current best overall performance:

$$\lambda_{t+1} \in \arg \min_{\lambda \geq 0} \sum_{\tau=1}^t \left(s_\tau - s_{\tau-1} - \sum_{j=1}^N u_{j,\tau}(\lambda) f_{j,\tau} \right)^2.$$

This last minimization problem may be computationally challenging and is often simplified by considering a finite grid Λ of possible values of λ and setting

$$\lambda_{t+1} \in \arg \min_{\lambda \in \Lambda} \sum_{\tau=1}^t \left(s_\tau - s_{\tau-1} - \sum_{j=1}^N u_{j,\tau}(\lambda) f_{j,\tau} \right)^2$$

instead. We do so with the grid

$$\Lambda = \{0\} \cup \left\{ m \times 10^k : m \in \{1, 2, 5\} \text{ and } k \in \{-4, -3, \dots, 1\} \right\} \cup \{10^2\}.$$

We will refer to this method as the adaptive sequential ridge regression (where “adaptive” is a reminder that we calibrate the regularization parameters λ_t over time).

Adaptive sequential ridge regression with discount factors. Actually, the method above was also twisted so as to focus more on recent observations while still coming with the desired theoretical guarantees (entailing, in particular, the convergence of the RMSE to the RMSE of the best fixed linear model), see Mauricette et al. [2009] as well as Stoltz [2010]. This is obtained via discounting: each past value, say of month τ , is associated with a discount factor of the form $1 + \mu/(t + 1 - \tau)^2$, where $t + 1$ is the index of the month where the next forecast is to be issued. The sequential ridge regression with regularization factor $\lambda \geq 0$ and discount factor $\mu \geq 0$ picks the linear weights

$$\begin{aligned} \mathbf{u}_{t+1}(\lambda, \mu) \in \arg \min_{\mathbf{u} \in \mathbb{R}^N} & \left\{ \lambda \sum_{j=1}^N u_j^2 \right. \\ & \left. + \sum_{\tau=1}^t \left(1 + \frac{\mu}{(t+1-\tau)^2} \right) \left(s_\tau - s_{\tau-1} - \sum_{j=1}^N u_j f_{j,\tau} \right)^2 \right\}. \end{aligned}$$

Again, the parameters λ_{t+1} and μ_{t+1} to be used for month $t + 1$ are estimated from the data in the same way as above. We consider the grid

$$\begin{aligned} \Lambda \times \Gamma = & \left(\{0\} \cup \left\{ m \times 10^k : m \in \{1, 2, 5\} \text{ and } k \in \{-4, -3, \dots, 1\} \right\} \cup \{10^2\} \right) \\ & \times \{0, 1, 2, 5, 10, 20, 50, 100\}. \end{aligned}$$

We choose

$$(\lambda_{t+1}, \mu_{t+1}) \in \arg \min_{(\lambda, \mu) \in \Lambda \times \Gamma} \sum_{\tau=1}^t \left(s_{\tau} - s_{\tau-1} - \sum_{j=1}^N u_{j,\tau}(\lambda, \mu) f_{j,\tau} \right)^2$$

and finally use the linear weights $\mathbf{u}_{t+1}(\lambda_{t+1}, \mu_{t+1})$ to issue our forecast:

$$\hat{s}_{t+1} - s_t = \sum_{j=1}^N u_{j,t+1}(\lambda_{t+1}, \mu_{t+1}) f_{j,t+1}.$$

3.2. Assessing the quality of the forecasts

We denote by T the total number of monthly values to be forecast, from months 1 to T . We allow for a short training period, of length $t_0 \geq 1$, and only evaluate the accuracy of the forecasts on months $t_0 + 1$ to T . We consider the root mean square error,

$$\begin{aligned} \text{RMSE} &= \sqrt{\frac{1}{T - t_0 + 1} \sum_{t=t_0}^T (\hat{s}_t - s_t)^2} \\ &= \sqrt{\frac{1}{T - t_0 + 1} \sum_{t=t_0}^T \left((\hat{s}_t - s_{t-1}) - (s_t - s_{t-1}) \right)^2}, \end{aligned} \tag{9}$$

and note that (by subtracting and adding the pivotal values s_{t-1}) this root mean square error is indifferently the one for the logarithms of exchange rates s_t or for the changes in logarithms of exchange rates $s_t - s_{t-1}$.

3.3. Study of statistical significance

We want to investigate whether the improvements in RMSE are statistically significant, that is, that the accuracy of the forecasts based on our adaptive sequential ridge regression with discount factors technique are significantly better than the one of the random walk. To test the hypothesis H_0 that the difference is not

significant against the alternative hypothesis H_1 that the new method is significantly better on average, the standard practice is to consider the instantaneous differences in accuracy between the random-walk strategy and the strategy under scrutiny,

$$d_t = (s_{t-1} - s_t)^2 - (\hat{s}_t - s_t)^2.$$

We denote by

$$\bar{d}_T = \frac{1}{T - t_0 + 1} \sum_{t=t_0+1}^T d_t$$

the empirical average of the differences and by

$$\bar{\sigma}_T^2 = \frac{1}{T - t_0 + 1} \sum_{t=t_0+1}^T (d_t - \bar{d}_T)^2$$

their empirical variance.

Some descriptive statistics. Table 8 reports the quantiles of the series of the d_t . We see that in the case of the adaptive sequential ridge regression with discount factors the distribution of the differences d_t is shifted toward positive values: the 75% and 90% quantiles are much larger in absolute values than the 25% and 10% quantiles, while in addition, the median is positive. This is not the case for rolling or recursive regressions. Additional comments on these matters are provided in Section 5.1.

Assuming i.i.d. differences. Under the assumption that the d_t be realizations of independent and identically distributed random variables D_t , the test statistic

$$S_{\text{iid}} = \sqrt{T - t_0 + 1} \frac{\bar{d}_T}{\sqrt{\bar{\sigma}_T^2}}$$

converges to a $\mathcal{N}(0, 1)$ distribution under H_0 while converging to $+\infty$ under H_1 . The p -values associated with the test thus constructed will be reported in

the tables in columns labeled “IID p -value”.

More general stochastic modeling of the differences. Diebold and Mariano [1995] (see also West [1996]) relaxed the needed assumptions on the behavior of the differences d_t . We state their results with a rectangular lag, as they advocated to so. They showed that under an assumption of covariance stationary and short memory, for a truncation lag denoted by $H \geq 0$, the test statistics

$$S_{\text{DMW},H} = \sqrt{T - t_0 + 1} \frac{\bar{d}_T}{\sqrt{\bar{\sigma}_T^{H,2}}},$$

where

$$\bar{\sigma}_T^{H,2} = \frac{1}{T - t_0 + 1} \sum_{t=t_0+1}^T (d_t - \bar{d}_T)^2 + \frac{2}{T - t_0 + 1} \sum_{\tau=1}^H \sum_{t=\tau+t_0+1}^T (d_t - \bar{d}_T)(d_{t-\tau} - \bar{d}_T),$$

converge to a $\mathcal{N}(0, 1)$ distribution under H_0 while converging to $+\infty$ under H_1 . The choice of the truncation lag H was partially left open and Diebold and Mariano suggested to pick it as a function of the length of the short memory (of the autocovariance degree). Estimating by \hat{H} a proper H based on our data, substituting its value would lead to considering the test statistic $S_{\text{DMW},\hat{H}}$, which would not be guaranteed anymore converge to a $\mathcal{N}(0, 1)$ distribution under H_0 . We instead take a more difficult approach to reject H_0 , that is, we build a more conservative test than the original test based on some good value of H . Namely, we consider

$$S_{\text{DMW}} = \sqrt{T - t_0 + 1} \frac{\bar{d}_T}{\sqrt{\max_{H \in \{0,1,\dots,20\}} \bar{\sigma}_T^{H,2}}}, \quad (10)$$

which is smaller than any of the corresponding original statistic $S_{\text{DMW},H}$. The limiting distribution under H_0 is smaller than an a $\mathcal{N}(0, 1)$ distribution. Yet, we compute the p -values using quantiles of the normal distribution, which is very conservative. Despite all, we will be consistently able to reject the hypothesis

H_0 of equal accuracy abilities for our new method, the adaptive sequential ridge regression with discount factors. Note that the maximal value 20 for H was set on our data set because it corresponds roughly to \sqrt{T} . The p -values associated with the test thus constructed will be reported in the tables in columns labeled “DMW p -value”.

A specific behavior for the differences in the case of nested models. We are comparing here through the d_t two models that are nested: the larger model (7) (in which we use the adaptive sequential ridge regression with discount factors to compute the weights u_t) encompasses in particular the random-walk model. In this case, under an assumptions of residuals of the models forming a martingale difference sequence, Clark and West [2006] exhibited a test based on some adjusted differences, which in our context equal

$$a_t = d_t + (\hat{s}_t - s_{t-1})^2.$$

We denote by

$$\bar{a}_T = \frac{1}{T - t_0 + 1} \sum_{t=t_0+1}^T a_t \quad (11)$$

the empirical average of the adjusted differences a_t and by

$$\bar{S}_T^2 = \frac{1}{T - t_0 + 1} \sum_{t=t_0+1}^T (a_t - \bar{a}_T)^2 \quad (12)$$

their empirical variance. The statistic

$$S_{\text{CW}} = \sqrt{T - t_0 + 1} \frac{\bar{a}_T}{\sqrt{\bar{S}_T^2}} \quad (13)$$

converges to a $\mathcal{N}(0, 1)$ distribution under H_0 while converging to $+\infty$ under H_1 . The p -values associated with the test thus constructed will be reported in the tables in columns labeled “CW p -value”.

4. Data and experts used

Our main sample includes major floating exchange rates between March 1973–May 2013 (at most we have 483 data points). We also use a supplementary sample for March 1973–June 2006 taken directly from Molodtsova and Papell [2009] in order to compare our methods with “experts” taken from a Taylor-rule based exchange rate models that are deemed in the literature as generally being the most successful in obtaining predictability of the exchange rate at the short 1–month horizon. For the sample 1973–2013 we try to extend the same data series as used in Molodtsova and Papell [2009], but some of them were discontinued. As a result, we tried to find the closest substitutes possible for the entire period 1973–2013 from similar sources (IMF, OECD) through Datastream. The exchange rates are taken from the Federal Reserve Bank of Saint Louis database no matter what is the sample considered. A detailed description of the data is given in Appendix B.

We study the behavior of major floating currencies that are active throughout the 1973–2013 period which leaves out for example the DEM/USD exchange rate as the German Mark was included in the Euro in 1999⁵. The exchange rates considered are: USD/GBP, JPY/USD, CHF/USD, CAD/USD, SEK/USD, USD/AUD, DNK/USD. Even for some of these currencies, however, it was not possible to obtain fundamentals series for the entire 1973–2013.

Our experts were formed as follows. The inflation differentials are calculated as 12–month changes in consumer price indexes (CPI). We use a money market rate or 3–month interest rate differentials for the interest rate based experts. For differences in the money stock growth and output growth we use the preceding 12–month trends in these variables⁶. The output gaps for the Taylor-rule expert are percentage deviations from “potential” output that was estimated including

⁵We obtain qualitatively similar results for the DEM/USD exchange rate for the period 1973–1999 to the ones reported for other currencies. Available upon request.

⁶We also considered 6–month trends that did not change qualitatively our results. Available upon request.

a (i) a linear trend, (ii) a quadratic trend (iii) a linear and quadratic trends or (iv) a Hodrick-Prescott filter using the data available prior to the date for which the output gap was calculated. We do not detrend, filter or seasonally adjust the data in any way.

5. Results

5.1. Main results

In Table 1 we show the results for various subsets of our “classic” macro-economic variables (based on PPP, UIRP or monetary models). Only linear combinations were considered here. This table contains the RMSE of the random walk prediction (column 1) and the Theil statistics (the ratio of the RMSE of the forecasts from our experts and the RMSE of a random walk) respectively for the OLS rolling regressions (column 2), OLS recursive regressions (column 5) and the adaptive sequential ridge regression with discount factors (column 8). Next to the Theil statistics we show the corresponding p -values of the tests of the hypothesis H_0 that the difference in the forecasting performance is not significant against the alternative hypothesis H_1 that the method under scrutiny is significantly better on average, as discussed in Section 3.3: correcting for the fact that we are comparing two nested models (the Clark-West test, columns 3, 6, 10, labeled “CW p -value”) or that they adhere to the assumptions of Diebold and Mariano [1995] and West [1996] (columns 4, 7, 11, labeled “DMW p -value”)⁷. For the adaptive sequential ridge regression with discount factors we also show the p -values of the tests under the assumptions that the differences are i.i.d (column 9).

The conclusions from Table 1 are striking. Using the adaptive sequential

⁷We chose the truncation lag H for each Diebold and Mariano [1995] test in a conservative and data-driven way, see the maximum in the denominator of (10). Realized values of the argument of this maximum are available upon request.

ridge regression with discount factors we are able to beat the random walk without drift in terms of the RMSE improvement for *all* currency pairs for the period 1973–2013 for *all* fundamentals considered. The gains are not higher than ca. 5 % in terms of the RMSE: we are not able to extract much more information from the fundamentals than what a random walk contains. Yet, comparing to the existing literature as surveyed by Rossi [2013] these gains are substantial. These improvements are statistically significant according to the Clark-West tests for comparing predictive ability of nested models at a 5 % level for all fundamentals and at a 1 % level for most pairs. The Diebold-Mariano-West (further DMW) tests are more conservative, but still we find statistical significance at a 10 % level in most cases. In fact, we can beat the random walk without drift in forecasting the exchange rate in a statistically significant way for each currency pair in at least one of the sets of fundamentals considered. No set of fundamentals seems to be clearly better than the other. For the UIRP expert, the p -values for the DMW tests are statistically significant at a 10 % level in 6 out of the 7 pairs (it fails for the CAD/USD) while for the PPP and monetary experts they are so for 5 out of 7 pairs (they fail for CAD/USD, USD/AUD and for CHF/USD, USD/AUD respectively). The performance of the UIRP expert is encouraging also for practical reasons: interest rate data is available in real-time so can be effectively used for forecasting.

Interestingly, when we combine all fundamentals together – in the spirit of the information aggregation ability of machine learning techniques – we do not gain in terms of forecasting accuracy. Sometimes in aggregate the forecasts turn to be less actually less precise according to the RMSE criterion when comparing for example column 8 in panels 3–5 together. This may be an indication that the fundamentals contain pretty much the same information. There is a cost suffered by the method while adding experts that do not perform well (see Section A in the online appendix) – hence parsimony seems to be preferred in our setting as found in other studies of exchange rate forecasting.

The same fundamentals seem not to have much of forecasting power when evaluated using the classical methods considered in the literature – either the rolling or recursive OLS regressions. For neither of the currency pairs we obtain an improvement in terms of RMSE upon the random walk model (which is already rare) that is statistically significant at conventional levels using the DMW tests. We obtain predictability in 4 out of 7 currency pairs for the monetary experts exhibited in panel 4 of Table 1. If we mesh in all “classic” experts together in one regression (so estimating some atheoretical model) we obtain predictability at a 10 % level for all currency pairs for the OLS rolling regressions except for the DNK/USD pair. However, this does not mean that we get always improvements upon the random walk in terms of RMSE. This is in line with the common knowledge and the conclusion from the empirical literature that “classic” fundamentals do not allow for systematic improvements in prediction or forecasting (see Rossi [2013]).

Finally, we note that some descriptive statistics show that this average good performance of our new method does not come at the cost of local disasters, on the contrary: the forecasting errors seem to be uniformly better over time. Indeed, when computing the quantiles of the difference between the forecasting error of the random walk minus the one of the methods under scrutiny, as is reported in Table 8, we see that these differences are uniformly much large for the adaptive sequential ridge regression with discount factors than for the recursive or rolling regressions. By “uniformly” we mean here that quantiles of the same order for two series of differences are almost always ranked in the same manner, the ones corresponding to the adaptive sequential ridge regression with discount factors being larger than the ones for the rolling or recursive regressions.

5.2. Robustness checks

We tried to assess the robustness of our findings by conducting different exercises with our data. While doing so, we did not find any qualitatively different results than those reported above (all available upon request). We scrutinized the performance of the adaptive sequential ridge regression with discount factors method on different samples: the original 1973–2006 data from Molodtsova and Papell [2009]; post-Plaza accord period 1985–2013 and the post ERM-crisis period 1992–2013. Not much changed when we formulated the forecasts in levels (as originally in the Meese and Rogoff [1983]) and used price, money and output indices in levels. We also considered 6-month inflation, money stock or output changes which did not change qualitatively our results. Whenever there was no obvious choice for a substitute series from the IFS database spanning 1973–2013 we investigated several alternative series. We tried also to use “decoupled” experts, i.e., allowed for the coefficients on the fundamentals that belong to the same pair to vary. For example, for the PPP experts that involved treating inflation rates $\pi_{t+1,t}$ and $\pi_{t+1,t}^*$ as separate predictors. We also applied our methods to similar cross-currency pairs within our sample: low-inflation (JPY/CHF), resource-based (CAD/AUD) and nordics (DNK/SEK). In our setting, as found in the literature starting from Meese and Rogoff [1983], the forecasting performance deteriorates substantially while using the actually realized values of the fundamentals.

5.3. Directional tests

We also conduct tests whether the forecasts based on fundamentals can consistently predict the direction of change of the exchange rate better than the random walk (which does not predict any change). We report our findings for all sets of experts in Table 2. In columns 1–3, 4–6 and 7–9 we show the percentage of correctly predicted changes, the p -values of tests of superior predictive ability

under the assumptions that the differences are i.i.d. or of the Diebold-Mariano-West kind, respectively for the rolling and recursive OLS regressions and the adaptive sequential ridge regression with discount factors. As can be seen, the forecasts given from the adaptive sequential ridge regression with discount factors of the future direction of the exchange rate for monetary model experts improve upon a fair coin toss at a 2 % significance level for all currency pairs. With the simple Frenkel-Bilson flexible price monetary model experts we can predict between 55.9 % and 63.5 % of the directional changes correctly. The performance of the PPP or UIRP experts is a bit worse but still they are able to forecast the direction of change better than the coin toss for 6 out of 7 and 5 out of 7 currency pairs respectively according to the DMW tests. The rolling or recursive OLS methods fail to produce such improvements all across the board and for some currency pairs (USD/AUD) do very poorly. That having said, the monetary model experts for the rolling OLS achieve also some success in forecasting the directional change for 4 out of 7 currencies (and beating the adaptive sequential ridge regression with discount factors for the CAD/USD pair).

5.4. Taylor-rule fundamentals

In Table 3 we show the results for the experts from the Taylor-rule based exchange rate model tested by Molodtsova and Papell [2009]. Four different trends were considered to construct the output gaps (percentage deviations from the trend): linear, quadratic, linear and quadratic, and the Hodrick-Prescott filter⁸. We show the results for “decoupled” experts where each expert has its own coefficient (the “coupled” experts results are shown in Appendix C) because these “decoupled experts” led to a better overall performance in terms of predictability/forecastability both for the OLS regression and the adaptive sequential ridge

⁸The trends were estimated as in Molodtsova and Papell [2009] to allow for a comparison.

regression with discount factors than the “coupled” ones⁹. Again the models were estimated in the linear, not affine, form. The first observation is that the Taylor-rule experts do deliver “predictability” for especially OLS rolling regressions in at least 4 out of 7 currency pairs for each output gap, confirming thus the findings of Molodtsova and Papell [2009]. However, the rolling OLS regressions with all “classic” experts together from Table 1, lowest 5th panel, beat the Taylor model predictions both in terms of RMSE and the pattern of predictability (all CW statistic p -values below 10 %). Not even once, however, the Taylor-rule experts are able to beat the random walk in terms of the RMSE¹⁰ when the model is estimated by OLS methods. This changes when we reestimate the model with the adaptive sequential ridge regression with discount factors. However, we do not obtain a better performance in general than that observed already in Table 1 (for example in gains in the RMSE). Although the CW p -values are all below 5 % and are picture-perfect, the significance of the forecast improvement at the 10 % level according to the DMW tests is a bit worse: it fails in at least 3 out of 7 currencies depending on the output gap measure. No matter what the output gap considered, the Taylor-rule experts cannot provide a statistically significant better forecast than the random walk for the CAD/USD, USD/AUD or DNK/USD currency pairs. The best Taylor-rule based model seems to be the one with the output gap measure obtained after detrending output with a linear and quadratic trend: the forecasts improvements over the random walk are statistically significant at a 5 % for 4 out of 7 currencies and at a 15 % for 3 others which nears the performance of the PPP or monetary models experts from Table 1. Basing on our estimates we do not find that the Taylor-rule experts perform better in forecasting the exchange rates than “classic” fundamentals and in fact we could

⁹For example, the RMSE gains against the random walk for the adaptive sequential ridge regression with discount factors were statistically significant at a 10 % level according to the DMW test only for at most 2 out of 7 currency pairs.

¹⁰For the coupled experts shown in Table 9 the exception are few results (4 out of 28 pairs) for the recursive OLS estimates. Still, they are not statistically significant at a 10 % level according to the DMW test.

argue they do slightly worse. Nevertheless, they still contain information that in general is valuable to allow for improving forecasts against the random walk.

5.5. Affine models and the clone problem

We estimated also affine versions of the models shown in Section 2. This means that we included a constant term – a “drift” – in each differenced equation alongside our fundamentals (Table 4). In doing so we discovered that a parsimonious model containing only a constant (“drift”) estimated by the adaptive sequential ridge regression with discount factors (see a more formal definition below, in Section 5.6.1) provides very good forecasts of the exchange rates (first upper panel, columns 8–11 of Table 4). This would be a source of worry if the random walk with “drift” estimated in the classical way would consistently provide better forecasts than the random walk in our sample, as this would indicate that there is something unusual with our data. However, the same model estimated by rolling or recursive OLS methods always performs worse than the random walk (first panel, columns 2–7 of Table 4) – which is consistent with the existing literature. Moreover, the “drift” estimated by our adaptive sequential ridge regression with discount factors (henceforth referred to as an adaptive drift) is not the typical “drift” term considered in the literature: weighing of the observations (based on the discounts estimated from the data) means that – depending on the past performance – effectively longer or shorter trends may be considered by the algorithm for forecasting purposes at different time points within the same sample. Unfortunately, adding our fundamentals does not improve predictive accuracy beyond that (lower panels 2–6 of Table 4)¹¹. Perhaps this is caused by the same problem faced by the traditional literature on exchange rate forecastability: adding predictors may cause the RMSE to increase in comparison to the nested model because of estimation inaccuracies even if these could

¹¹This is also true for the Taylor-rule based exchange rate model experts. An example is shown in Table XXX in Appendix C.

be informative (we explore this further in Section 5.6 below). It is to be noted that we obtain similar forecasting performances for the linear and affine models with the fundamentals: no constant (“drift”) term is used to obtain the RMSE improvements reported in Section 5.1 above. What is to be noted is that we obtain predictability for our classic experts for many currency pairs as witnessed by p -values of the CW-tests. Again as in Section 5.1, however, any gains in the forecasts for the OLS-estimated models are not statistically significant according to DMW tests.

The good performance of the random walk with adaptive drift seems to be similar to the “clone” problem known in the machine learning literature: the fundamentals we consider may carry similar information as the drift. The (adaptive) drift by itself has no economic interpretation. It may well be that the changes in the fundamentals cause the drift to appear. For example, the Purchasing Power Parity theory would predict that if one country has a consistently higher inflation rate than the other, its currency should depreciate. In an empirical investigation this could be perfectly approximated by a drift that would be completely explained by the changes in the price levels. A pattern like this appears in our data for some currency pairs – for example the Swiss Franc/U.S. dollar cross (Figure 1). An important question arises: are the “classic” fundamentals able to explain the changes in the exchange rates beyond the drift that they may carry? Given that the affine models with fundamentals in general do not produce better forecasts in terms of RMSEs than that with a pure “constant” (so the hope for better forecastability is lost), we want also to test for predictability of the fundamentals against the random walk with drift. In order to do this, we shall compare forecasting performance of two nested affine models using a 2-step procedure described in the next section.

5.6. Clark and West [2006] against the random walk with adaptive drift

We want to compare the predictive accuracy of the forecasts made by considering constant terms and linear combinations of the fundamentals to the ones made by considering but constant terms (the “random walk with adaptive drift”).

To be able to use the test by Clark and West [2006], it is necessary that the constants used in both formulations are the same, see Equation (2.6) in the mentioned reference. This is why we first explain how to adapt the aggregation method above to first pick the constant terms and then to possibly add some linear combination of the fundamentals.

5.6.1. Random walk with adaptive drift, with the possible addition of fundamentals

We call the random walk with adaptive drift the adaptive sequential ridge regression with discount factors solely based on constant experts $f_{0,t+1} \equiv 0.01$ to forecast the s_{t+1} . This leads to predictions that will be referred to as RWd_{t+1} in the sequel and that are of the form

$$\text{RWd}_{t+1} = s_t + \gamma_{t+1} .$$

Now, we can possibly include linear combinations of the fundamentals as follows. We use again the adaptive sequential ridge regression with discount factors, based this time on the fundamentals only, that is, on experts $f_{j,t+1}$ with $j \geq 1$ but not on the constant expert $f_{0,t+1}$, to forecast the $\Delta_{t+1} = s_{t+1} - \text{RWd}_{t+1}$ (instead of forecasting directly the s_{t+1}); that is, we issue first pre-forecasts of the form

$$\hat{\Delta}_{t+1} = \sum_{j=1}^N u_{j,t+1} f_{j,t+1} ,$$

where the vectors \mathbf{u}_{t+1} are obtained based on minimizations of the form

$$\mathbf{u}_{t+1}(\lambda, \mu) \in \arg \min_{\mathbf{u} \in \mathbb{R}^N} \left\{ \lambda \sum_{j=1}^N u_j^2 + \sum_{\tau=1}^t \left(1 + \frac{\mu}{(t+1-\tau)^2} \right) \left(s_\tau - s_{\tau-1} - \gamma_\tau - \sum_{j=1}^N u_j f_{j,\tau} \right)^2 \right\}.$$

We then substitute these linear combinations to output the final forecasts

$$\ln \tilde{s}_{t+1} = \hat{\Delta}_{t+1} + \text{RWd}_{t+1} = s_t + \left(\gamma_{t+1} + \sum_{j=1}^N u_{j,t+1} f_{j,t+1} \right).$$

5.6.2. The adapted Clark and West [2006] test

Thanks to the constraint that \tilde{s}_{t+1} and RWd_{t+1} possess some common part $\gamma_{t+1} + \ln s_t$ and only differ by (time-varying) linear combinations of the fundamentals, the assumptions of the setting of Clark and West [2006] may be satisfied and their methodology followed.

The latter can be implemented by considering the

$$a'_t = (\ln \text{RWd}_t - \ln s_t)^2 - (\ln \tilde{s}_t - \ln s_t)^2 + (\ln \text{RWd}_t - \ln \tilde{s}_t)^2$$

in lieu of the a_t and then, resorting to the same test statistic as in (11)–(13).

5.6.3. Results from our 2-step procedure

We report the results of our 2-step procedure described above for the 1973–2013 sample in Table 5. We show the best results obtained in terms of the Theil statistics against the random walk with adaptive drift – those from the flexible prices monetary model experts (the PPP and UIRP experts fare worse), all experts combined and for the best Taylor-rule based exchange rate model fundamentals. For the monetary experts we find predictability at a 10 % level against the random walk with adaptive drift for 3 currency pairs while for one other at a 12 % level.

We are unable to obtain any improvements for the CHF/USD, USD/AUD and DNK/USD currency pairs. Perhaps this is so because the industrial production figures for the first two countries in particular are available only at quarterly horizons, which worsens the informativeness of the production fundamental¹². The Taylor-rule experts (lowest panel of Table 5) exhibit predictability for 4 out of 7 currency pairs at a 10 % level against the random walk with adaptive drift and as such they do better than the monetary model. Interestingly, the monetary experts and Taylor-rule experts are able to allow predictability for different currency pairs. The only currency pair for which we cannot claim predictability even at a 20 % significance level is the CHF/USD currency cross, which incidentally is also the one for which there appears to be a drift in prices all across the sample period as shown in Figure 1. We conclude that we have weak evidence about predictability of the monetary and Taylor-rule experts against the random walk with adaptive drift (estimated with the adaptive sequential ridge regression with discount factors).

5.7. All experts combined

Machine-learning techniques are often used to efficiently aggregate information from many experts. We conduct such experiments, combining “classic”, Taylor-rule exchange rate models based experts, “classic” experts in the level (absolute) and the “drift” parameter. We show the results in Table 6. First, the adaptive sequential ridge regression with discount factors methods always deliver RMSE that are lower than those of the random walk. This improvement, however, is not much better than those observed for our “classic” experts only depicted in Table 1. Second, it is startling to observe that OLS methods exhibit often predictability of the exchange rate even though the Theil ratios of the RMSEs generated through these models are even 12 % higher than 1 and never

¹²Substitute series available at monthly intervals such as retail sales figures improve the p -values considerably but we cannot obtain predictability at a 10 % level for these currencies.

forecastability as witnessed by the DMW tests. We conclude that we cannot obtain great forecasting improvements using all the experts at our disposal.

6. Conclusions

In this paper we apply a method stemming from the field of machine learning – adaptive sequential ridge regression with discount factors – to the perennial problem of exchange rate forecasting. In doing so, we obtain gains in forecasting using the standardly applied RMSE criterion for PPP, UIRP and monetary-models based fundamentals that were not found using traditionally applied estimation methods based on OLS. We conclude thus that a major problem of international economics – whether there is a short-term relationship between “classic” fundamentals and exchange rates – is answered in the affirmative under the condition that proper statistical techniques, e.g., the adaptive sequential ridge regression with discount factors, are applied. Moreover, our success points to a possible potential of such techniques for improving the evaluation of economic problems.

Machine learning techniques serve also to effectively aggregate information from many sources. A tempting exercise, beyond the scope of this paper, would be to evaluate the forecasting performance including many more experts than the “classic” ones considered here that were suggested by the literature – for example those based on productivity, interest rate yield curves, net foreign assets etc. Venturing further one could consider many more experts – data series that are not typically associated with exchange rate forecasting in the true spirit of machine learning.

As with any new method applied to exchange rate forecasting, it remains to be seen whether our results could be replicated for different currencies, samples, forecasting periods and fundamentals. Given the robustness of the results shown in this paper, however, we hope that the application of these methods to

exchange rate forecasting will stand the test of time and will allow for better predictions and decision making in the future.

References

- K.S. Azoury and M. Warmuth. Relative loss bounds for on-line density estimation with the exponential family of distributions. *Machine Learning*, 43:211–246, 2001.
- Philippe Bacchetta, Eric van Wincoop, and Toni Beutler. Can parameter instability explain the Meese-Rogoff puzzle? In *NBER International Seminar on Macroeconomics 2009*, pages 125–173. University of Chicago Press, June 2010. URL <http://www.nber.org/chapters/c11912>.
- Valerie Cerra and Sweta Chaman Saxena. The monetary model strikes back: Evidence from the world. *Journal of International Economics*, 81:184–196, 2010.
- N. Cesa-Bianchi and G. Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, 2006.
- Yin-Wong Cheung, Menzie D.Chinn, and Antonio Garcia Pascual. Empirical exchange rate models of the nineties: Are any fit to survive? *Journal of International Money and Finance*, 24:1150–1175, 2005.
- Richard H. Clarida, Lucio Sarno, Mark P. Taylor, and Giorgio Valente. The out-of-sample success of term structure models: A step beyond. *Journal of International Economics*, 60:61–83, 2003.
- T. Clark and K. West. Using out-of-sample mean squared prediction errors to test the martingale difference hypothesis. *Journal of Econometrics*, 135(1–2):155–186, 2006.
- Pasquale Della Corte, Lucio Sarno, and Giulia Sestieri. The predictive information content of external imbalances for exchange rate returns: How much is it worth? *The Review of Economics and Statistics*, 94:100–115, 2012.
- M. Devaine, P. Gaillard, Y. Goude, and G. Stoltz. Forecasting the electricity consumption by aggregation of specialized experts; application to Slovakian and French country-wide (half-)hourly predictions. *Machine Learning*, 90(2):231–260, 2013.
- F. Diebold and R. Mariano. Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13(3):253–263, 1995.
- Charles Engel. Can the Markov switching model forecast exchange rates? *Journal of International Economics*, 36:151–165, 1994.

- Charles Engel and Kenneth D. West. Exchange rates and fundamentals. *Journal of Political Economy*, 113:485–517, 2005.
- Charles Engel and Kenneth D. West. Taylor rules and the Deutschmark-Dollar real exchange rate. *Journal of Money, Credit and Banking*, 38(5):1175–1194, 2006.
- Raffaella Giacomini and Barbara Rossi. Forecast comparisons in unstable environments. *Journal of Applied Econometrics*, 25(4):595–620, 2010.
- Pierre-Olivier Gourinchas and Helene Rey. International financial adjustment. *Journal of Political Economy*, 115(4):665–703, 2007.
- A.E. Hoerl and R.W. Kennard. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12:55–67, 1970.
- Nelson C. Mark. Exchange rates and fundamentals: Evidence on long-horizon predictability. *American Economic Review*, 85:201–218, 1995.
- Nelson C. Mark and Donggyu Sul. Nominal exchange rates and monetary fundamentals: Evidence from a small post-Bretton Woods panel. *Journal of International Economics*, 53:29–52, 2001.
- Boris Mauricette, Vivien Mallet, and Gilles Stoltz. Ozone ensemble forecast with machine learning algorithms. *Journal of Geophysical Research*, 114:D05307, 2009.
- Richard A. Meese and Kenneth Rogoff. Empirical exchange rate models of the seventies. Do they fit out of sample? *Journal of International Economics*, 14:3–24, 1983.
- Tanya Molodtsova and David H. Papell. Out-of-sample exchange rate predictability with Taylor rule fundamentals. *Journal of International Economics*, 77:167–180, 2009.
- Tanya Molodtsova, Alex Nikolsko-Rzhevskyy, and David H. Papell. Taylor rules and the Euro. *Journal of Money, Credit and Banking*, 43:535–552, 2011.
- David E. Rapach and Mark E. Wohar. Testing the monetary model of exchange rate determination: new evidence from a century of data. *Journal of International Economics*, 58:359–385, 2002.
- Kenneth S. Rogoff and Vania Stavrakeva. The continuing puzzle of short horizon exchange rate forecasting. Working Paper 14071, National Bureau of Economic Research, 2008.
- Barbara Rossi. Are exchange rates really random walks? Some evidence robust to parameter instability. *Macroeconomic Dynamics*, 10(1):20–38, 2006.

- Barbara Rossi. Exchange rate predictability. *Journal of Economic Literature*, 51:1063–1119, 2013.
- Barbara Rossi and Atsushi Inoue. Out-of-sample forecast tests robust to the choice of window size. *Journal of Business and Economic Statistics*, 30(3):432–453, 2012.
- Garry J. Schinasi and P. A. V. B. Swamy. The out-of-sample forecasting performance of exchange rate models when coefficients are allowed to change. *Journal of International Money and Finance*, 8:375–390, 1989.
- Gilles Stoltz. Agrégation séquentielle de prédicteurs : méthodologie générale et applications à la prévision de la qualité de l’air et à celle de la consommation électrique. *Journal de la Société Française de Statistique*, 151(2):66–106, 2010.
- V. Vovk. Competitive on-line statistics. *International Statistical Review*, 69:213–248, 2001.
- Kenneth D. West. Asymptotic inference about predictive ability. *Econometrica*, 64:1067–1084, 1996.
- Jonathan H. Wright. Bayesian Model Averaging and exchange rate forecasts. *Journal of Econometrics*, 146:329–341, 2008.

Table 1: Main results: RMSE and Theil ratios of forecasts for linear models.

***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

Currency pair	Random walk RMSE	Rolling regression			Recursive regression			Sequential ridge adaptive			
		Theil ratio	CW p-value	DMW p-value	Theil ratio	CW p-value	DMW p-value	Theil ratio	IID p-value	CW p-value	DMW p-value
USD/GBP JPY/USD CHF/USD CAD/USD SEK/USD USD/AUD DNK/USD	2.4728 2.7562 2.9775 1.4581 2.5710 2.6155 2.6213	PPP expert									
		1.0014	0.2211	0.5798	1.0010	0.2309	0.5593	0.9721	0.0009 ***	0.0000 ***	0.0043 ***
		1.0013	0.4199	0.6425	1.0011	0.8643	0.8491	0.9917	0.0362 **	0.0057 ***	0.0361 **
		1.0069	0.5774	0.7820	1.0034	0.3416	0.6499	0.9790	0.0491 **	0.0005 ***	0.0618 *
		1.0082	0.6730	0.7611	1.0001	0.2451	0.5226	0.9959	0.1453	0.0141 **	0.1484
		1.0047	0.5941	0.7886	0.9981	0.0905 *	0.3287	0.9775	0.0013 ***	0.0000 ***	0.0137 **
		1.0020	0.3679	0.6806	0.9994	0.1439	0.4302	0.9930	0.1597	0.0405 **	0.1877
		1.0069	0.7438	0.8714	1.0016	0.4179	0.6204	0.9846	0.0147 **	0.0005 ***	0.0318 **
		UIRP expert									
USD/GBP	2.4728	1.0039	0.3651	0.6800	1.0040	0.3668	0.6814	0.9709	0.0190 **	0.0000 ***	0.0268 **
JPY/USD	2.7562	1.0005	0.3082	0.5396	1.0016	0.6919	0.7504	0.9788	0.0108 **	0.0001 ***	0.0107 **
CHF/USD	2.9775	1.0084	0.4682	0.7411	1.0053	0.3094	0.6574	0.9741	0.0413 **	0.0001 ***	0.0522 *
CAD/USD	1.4581	1.0039	0.5454	0.8076	1.0010	0.2531	0.6316	0.9990	0.4042	0.0397 **	0.4041
SEK/USD	2.5710	1.0039	0.2839	0.6383	0.9978	0.0674 *	0.4127	0.9666	0.0020 ***	0.0000 ***	0.0286 **
USD/AUD	2.6155	1.0031	0.2868	0.6845	1.0001	0.1246	0.5050	0.9850	0.0579 *	0.0022 ***	0.0719 *
DNK/USD	2.6213	1.0073	0.6207	0.8141	1.0041	0.4933	0.7043	0.9863	0.0725 *	0.0013 ***	0.0778 *
		Monetary model with flexible prices (Frenkel-Bilson) experts									
USD/GBP	2.4728	1.0086	0.1906	0.7819	1.0036	0.1757	0.6182	0.9493	0.0031 ***	0.0000 ***	0.0043 ***
JPY/USD	2.7562	1.0048	0.1219	0.7072	1.0004	0.1639	0.5331	0.9694	0.0076 ***	0.0000 ***	0.0075 ***
CHF/USD	2.9775	1.0071	0.2133	0.7146	1.0041	0.2859	0.6472	0.9880	0.2264	0.0004 ***	0.2621
CAD/USD	1.4581	1.0196	0.4242	0.9023	1.0040	0.4996	0.8168	0.9803	0.0437 **	0.0013 ***	0.0552 *
SEK/USD	2.5710	1.0189	0.8157	0.9609	1.0047	0.4375	0.7093	0.9631	0.0030 ***	0.0000 ***	0.0045 ***
USD/AUD	2.6155	1.0028	0.1626	0.6078	1.0027	0.5983	0.7820	0.9821	0.1032	0.0009 ***	0.1414
DNK/USD	2.6213	1.0229	0.6302	0.9533	1.0114	0.4932	0.7925	0.9805	0.0624 *	0.0001 ***	0.0685 *
		Monetary model with flexible prices experts									
USD/GBP	2.4728	1.0134	0.1273	0.8323	1.0101	0.2072	0.7601	0.9472	0.0041 ***	0.0000 ***	0.0058 ***
JPY/USD	2.7562	0.9912	0.0068 ***	0.2938	0.9868	0.0086 ***	0.1670	0.9701	0.0113 **	0.0000 ***	0.0113 **
CHF/USD	2.9775	0.9874	0.0016 ***	0.2306	0.9829	0.0016 ***	0.1324	0.9793	0.0897 *	0.0005 ***	0.1411
CAD/USD	1.4581	0.9773	0.0005 ***	0.1987	0.9993	0.0918 *	0.4764	0.9826	0.0618 *	0.0016 ***	0.0616 *
SEK/USD	2.5710	1.0244	0.4653	0.9155	1.0117	0.6846	0.8894	0.9668	0.0084 ***	0.0000 ***	0.0083 ***
USD/AUD	2.6155	0.9896	0.0032 ***	0.2958	1.0072	0.8296	0.9230	0.9813	0.0888 *	0.0007 ***	0.1276
DNK/USD	2.6213	1.0265	0.5932	0.9739	1.0145	0.5538	0.8450	0.9781	0.0596 *	0.0000 ***	0.0668 *
		All experts combined									
USD/GBP	2.4728	1.0271	0.0823 *	0.9260	1.0186	0.2656	0.8979	0.9548	0.0049 ***	0.0000 ***	0.0078 ***
JPY/USD	2.7562	1.0178	0.0519 *	0.8632	1.0047	0.0107 **	0.6308	0.9697	0.0020 ***	0.0000 ***	0.0020 ***
CHF/USD	2.9775	1.0015	0.0037 ***	0.5368	1.0003	0.0087 ***	0.5070	0.9853	0.1838	0.0003 ***	0.1992
CAD/USD	1.4581	0.9846	0.0004 ***	0.2224	1.0010	0.0987 *	0.5386	0.9845	0.0614 *	0.0019 ***	0.1054
SEK/USD	2.5710	1.0234	0.0261 **	0.8341	1.0080	0.0203 **	0.6596	0.9608	0.0008 ***	0.0000 ***	0.0022 ***
USD/AUD	2.6155	0.9853	0.0007 ***	0.3064	1.0004	0.0214 **	0.5119	0.9733	0.0104 **	0.0001 ***	0.0481 **
DNK/USD	2.6213	1.0385	0.2329	0.9817	1.0151	0.2671	0.8652	0.9784	0.0222 **	0.0000 ***	0.0249 **

Table 2: Directional tests: Percentages of changes predicted for linear models.

***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

Currency pair	Rolling regression			Recursive regression			Sequential adaptive regression		
	Percentage changes predicted	IID p-value	DMW p-value	Percentage changes predicted	IID p-value	DMW p-value	Percentage changes predicted	IID p-value	DMW p-value
	PPP expert								
USD/GBP	0.5200	0.1981	0.2702	0.5156	0.2546	0.3228	0.5622	0.0041 ***	0.0297 **
JPY/USD	0.4866	0.7146	0.6542	0.4844	0.7458	0.6720	0.5826	0.0002 ***	0.0018 ***
CHF/USD	0.5201	0.2042	0.3017	0.5035	0.4420	0.4657	0.5981	0.0000 ***	0.0003 ***
CAD/USD	0.5234	0.1608	0.2287	0.5011	0.4812	0.4876	0.5457	0.0265 **	0.0800 *
SEK/USD	0.5244	0.1498	0.2780	0.5111	0.3187	0.3950	0.5867	0.0001 ***	0.0047 ***
USD/AUD	0.4531	0.9764	0.9347	0.4665	0.9218	0.8426	0.5201	0.1975	0.2491
DNK/USD	0.4563	0.9619	0.8782	0.5049	0.4219	0.4493	0.5680	0.0029 ***	0.0104 **
	UIRP expert								
USD/GBP	0.5156	0.2546	0.3090	0.5089	0.3530	0.3866	0.5622	0.0041 ***	0.0183 **
JPY/USD	0.5045	0.4251	0.4460	0.4844	0.7458	0.6892	0.5826	0.0002 ***	0.0012 ***
CHF/USD	0.5177	0.2329	0.3127	0.5201	0.2042	0.3131	0.6028	0.0000 ***	0.0001 ***
CAD/USD	0.5100	0.3355	0.3951	0.4944	0.5933	0.5601	0.5278	0.1190	0.1781
SEK/USD	0.4933	0.6114	0.5716	0.4956	0.5748	0.5442	0.5778	0.0005 ***	0.0075 ***
USD/AUD	0.4710	0.8903	0.8193	0.4598	0.9555	0.8790	0.5134	0.2854	0.3067
DNK/USD	0.4782	0.8124	0.7325	0.4976	0.5392	0.5258	0.5534	0.0151 **	0.0269 **
	Monetary model with flexible prices (Frenkel-Bilson) experts								
USD/GBP	0.5178	0.2254	0.2499	0.5111	0.3187	0.3623	0.6067	0.0000 ***	0.0001 ***
JPY/USD	0.5335	0.0782 *	0.1395	0.5201	0.1975	0.2437	0.5915	0.0001 ***	0.0001 ***
CHF/USD	0.5177	0.2329	0.3336	0.5083	0.3668	0.4244	0.6265	0.0000 ***	0.0000 ***
CAD/USD	0.5590	0.0062 ***	0.0200 **	0.4967	0.5563	0.5382	0.5590	0.0062 ***	0.0087 ***
SEK/USD	0.4689	0.9066	0.7967	0.5022	0.4624	0.4795	0.6356	0.0000 ***	0.0000 ***
USD/AUD	0.4754	0.8507	0.7645	0.4442	0.9909	0.9514	0.5781	0.0005 ***	0.0014 ***
DNK/USD	0.4636	0.9303	0.8627	0.4976	0.5392	0.5271	0.6092	0.0000 ***	0.0000 ***
	Monetary model with flexible prices experts								
USD/GBP	0.5378	0.0545 *	0.0748 *	0.5022	0.4624	0.4713	0.6089	0.0000 ***	0.0000 ***
JPY/USD	0.5357	0.0653 *	0.1086	0.5134	0.2854	0.3134	0.5871	0.0001 ***	0.0002 ***
CHF/USD	0.5437	0.0360 **	0.1004	0.5414	0.0444 **	0.0981 *	0.5934	0.0001 ***	0.0002 ***
CAD/USD	0.6013	0.0000 ***	0.0002 ***	0.5278	0.1190	0.2148	0.5635	0.0036 ***	0.0167 **
SEK/USD	0.4756	0.8502	0.7526	0.5044	0.4252	0.4572	0.6311	0.0000 ***	0.0000 ***
USD/AUD	0.4777	0.8276	0.7488	0.4353	0.9969	0.9867	0.5781	0.0005 ***	0.0014 ***
DNK/USD	0.4709	0.8815	0.8227	0.5024	0.4608	0.4709	0.6092	0.0000 ***	0.0000 ***
	All experts combined								
USD/GBP	0.5289	0.1102	0.1506	0.5111	0.3187	0.3366	0.5933	0.0000 ***	0.0003 ***
JPY/USD	0.5179	0.2248	0.2661	0.5223	0.1724	0.2231	0.6094	0.0000 ***	0.0000 ***
CHF/USD	0.5225	0.1778	0.2546	0.5177	0.2329	0.2728	0.6052	0.0000 ***	0.0000 ***
CAD/USD	0.5702	0.0015 ***	0.0025 ***	0.5412	0.0404 **	0.0869 *	0.5278	0.1190	0.1747
SEK/USD	0.5378	0.0545 *	0.1265	0.4956	0.5748	0.5507	0.6022	0.0000 ***	0.0002 ***
USD/AUD	0.5134	0.2854	0.3483	0.5223	0.1724	0.2409	0.5714	0.0012 ***	0.0030 ***
DNK/USD	0.5461	0.0306 **	0.0550 *	0.5437	0.0381 **	0.0691 *	0.6214	0.0000 ***	0.0000 ***

Table 3: RMSE and Theil ratios of forecasts for the Taylor-rule based exchange models. “Decoupled” experts.

***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

Currency pair	Random walk RMSE	Rolling regression			Recursive regression			Sequential adaptive ridge			
		Theil ratio	CW p-value	DMW p-value	Theil ratio	CW p-value	DMW p-value	Theil ratio	IID p-value	CW p-value	DMW p-value
		Output gap experts: deviations from a linear trend									
USD/GBP	2.4728	1.0291	0.0001 ***	0.8785	1.0189	0.0003 ***	0.7845	0.9689	0.0663 *	0.0000 ***	0.0969 *
JPY/USD	2.7562	1.0210	0.0055 ***	0.8313	1.0130	0.0091 ***	0.7325	0.9726	0.0883 *	0.0000 ***	0.0880 *
CHF/USD	2.9775	1.0420	0.0735 *	0.9610	1.0258	0.1089	0.8986	0.9806	0.1434	0.0000 ***	0.1431
CAD/USD	1.4581	1.0387	0.3232	0.9009	1.0131	0.5613	0.9445	0.9956	0.3324	0.0109 **	0.3322
SEK/USD	2.5710	1.0397	0.1829	0.9440	1.0115	0.0166 **	0.6394	0.9639	0.0106 **	0.0000 ***	0.0302 **
USD/AUD	2.6155	1.0245	0.0210 **	0.8869	1.0205	0.2182	0.9287	0.9948	0.3308	0.0076 ***	0.3334
DNK/USD	2.6213	1.0211	0.0523 *	0.8657	1.0110	0.0336 **	0.7511	0.9921	0.3208	0.0003 ***	0.3411
		Output gap experts: deviations from a quadratic trend									
USD/GBP	2.4728	1.0262	0.0001 ***	0.8531	1.0246	0.0007 ***	0.8484	0.9619	0.0586 *	0.0000 ***	0.0809 *
JPY/USD	2.7562	1.0256	0.0071 ***	0.8689	1.0121	0.0113 **	0.7208	0.9756	0.0965 *	0.0000 ***	0.0962 *
CHF/USD	2.9775	1.0340	0.0895 *	0.9521	1.0241	0.2181	0.9148	0.9834	0.1591	0.0000 ***	0.1588
CAD/USD	1.4581	1.0374	0.2354	0.8850	1.0129	0.5471	0.9586	0.9839	0.1461	0.0020 ***	0.1459
SEK/USD	2.5710	1.0460	0.3032	0.9326	1.0164	0.0526 *	0.7481	0.9563	0.0162 **	0.0000 ***	0.0351 **
USD/AUD	2.6155	1.0078	0.0006 ***	0.6260	1.0187	0.0937 *	0.9028	0.9731	0.0652 *	0.0023 ***	0.1227
DNK/USD	2.6213	1.0212	0.0603 *	0.8576	1.0119	0.1031	0.7882	0.9807	0.1379	0.0000 ***	0.1521
		Output gap experts: deviations from a linear and a quadratic trend									
USD/GBP	2.4728	1.0358	0.0039 ***	0.9272	1.0242	0.0062 ***	0.8564	0.9590	0.0260 **	0.0000 ***	0.0446 **
JPY/USD	2.7562	1.0219	0.0093 ***	0.8618	1.0098	0.0291 **	0.7128	0.9660	0.0246 **	0.0000 ***	0.0245 **
CHF/USD	2.9775	1.0295	0.0723 *	0.9332	1.0248	0.1721	0.8934	0.9722	0.0331 **	0.0000 ***	0.0380 **
CAD/USD	1.4581	1.0390	0.1867	0.9051	1.0131	0.2099	0.8793	0.9858	0.1450	0.0027 ***	0.1447
SEK/USD	2.5710	1.0466	0.3659	0.9639	1.0218	0.1318	0.8252	0.9578	0.0210 **	0.0000 ***	0.0411 **
USD/AUD	2.6155	1.0082	0.0018 ***	0.6357	1.0164	0.0794 *	0.8705	0.9745	0.0825 *	0.0014 ***	0.1339
DNK/USD	2.6213	1.0261	0.1226	0.9134	1.0190	0.3901	0.9120	0.9816	0.1328	0.0001 ***	0.1372
		Output gap experts: deviations from a Hodrick-Prescott filtered trend									
USD/GBP	2.4728	1.0382	0.0110 **	0.9414	1.0252	0.0448 **	0.9016	0.9600	0.0173 **	0.0000 ***	0.0334 **
JPY/USD	2.7562	1.0231	0.0311 **	0.9158	1.0032	0.0209 **	0.6114	0.9728	0.0155 **	0.0000 ***	0.0154 **
CHF/USD	2.9775	1.0390	0.3809	0.9845	1.0233	0.5502	0.9358	0.9800	0.0293 **	0.0000 ***	0.0291 **
CAD/USD	1.4581	1.0321	0.0985 *	0.9034	1.0073	0.1768	0.7952	1.0036	0.6008	0.0316 **	0.6009
SEK/USD	2.5710	1.0488	0.2031	0.9533	1.0169	0.0818 *	0.7686	0.9559	0.0119 **	0.0001 ***	0.0327 **
USD/AUD	2.6155	1.0088	0.0074 ***	0.6335	1.0186	0.2327	0.9264	0.9814	0.0457 **	0.0061 ***	0.1201
DNK/USD	2.6213	1.0557	0.3920	0.9867	1.0187	0.5768	0.9255	0.9930	0.3323	0.0052 ***	0.3403

Table 4: RMSE and Theil ratios of forecasts for affine models.

***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

Currency pair	Random walk RMSE	Rolling regression			Recursive regression			Sequential ridge adaptive			
		Theil ratio	CW p-value	DMW p-value	Theil ratio	CW p-value	DMW p-value	Theil ratio	IID p-value	CW p-value	DMW p-value
		Constant									
USD/GBP	2.4728	1.0037	0.3380	0.6747	1.0044	0.3576	0.7059	0.9545	0.0023 ***	0.0000 ***	0.0044 ***
JPY/USD	2.7562	1.0009	0.1497	0.5418	1.0005	0.1750	0.5281	0.9574	0.0004 ***	0.0000 ***	0.0004 ***
CHF/USD	2.9775	1.0021	0.1554	0.5800	1.0014	0.1109	0.5520	0.9628	0.0049 ***	0.0000 ***	0.0106 **
CAD/USD	1.4581	1.0052	0.3251	0.7129	1.0024	0.3343	0.6928	0.9793	0.0092 ***	0.0001 ***	0.0307 **
SEK/USD	2.5710	1.0073	0.6713	0.8217	1.0021	0.3991	0.6331	0.9492	0.0001 ***	0.0000 ***	0.0052 ***
USD/AUD	2.6155	1.0026	0.2119	0.6367	1.0031	0.2827	0.6861	0.9767	0.0501 *	0.0008 ***	0.0759 *
DNK/USD	2.6213	1.0093	0.7311	0.8866	1.0051	0.6896	0.7623	0.9669	0.0018 ***	0.0000 ***	0.0020 ***
		PPP expert + constant									
USD/GBP	2.4728	1.0036	0.0619 *	0.6022	1.0055	0.4131	0.7192	0.9588	0.0033 ***	0.0000 ***	0.0055 ***
JPY/USD	2.7562	1.0118	0.2421	0.8005	1.0030	0.0784 *	0.5943	0.9659	0.0028 ***	0.0000 ***	0.0028 ***
CHF/USD	2.9775	1.0004	0.0249 **	0.5121	1.0041	0.1734	0.6473	0.9678	0.0141 **	0.0000 ***	0.0210 **
CAD/USD	1.4581	0.9965	0.0059 ***	0.4202	1.0029	0.1767	0.6990	0.9793	0.0103 **	0.0001 ***	0.0423 **
SEK/USD	2.5710	0.9990	0.0268 **	0.4695	1.0023	0.3960	0.6337	0.9474	0.0002 ***	0.0000 ***	0.0050 ***
USD/AUD	2.6155	0.9834	0.0010 ***	0.2954	1.0030	0.2350	0.6451	0.9780	0.0367 **	0.0011 ***	0.0920 *
DNK/USD	2.6213	1.0052	0.0243 **	0.6399	1.0030	0.1158	0.6154	0.9712	0.0038 ***	0.0000 ***	0.0042 ***
		UIRP expert + constant									
USD/GBP	2.4728	1.0165	0.3928	0.9067	1.0090	0.4767	0.8371	0.9591	0.0062 ***	0.0000 ***	0.0113 **
JPY/USD	2.7562	1.0049	0.0516 *	0.6246	1.0056	0.0583 *	0.6279	0.9645	0.0017 ***	0.0000 ***	0.0017 ***
CHF/USD	2.9775	1.0127	0.3400	0.8304	1.0066	0.2374	0.7022	0.9682	0.0198 **	0.0000 ***	0.0246 **
CAD/USD	1.4581	1.0124	0.2122	0.8024	1.0043	0.3308	0.7628	0.9846	0.0763 *	0.0019 ***	0.0976 *
SEK/USD	2.5710	1.0016	0.0315 **	0.5431	0.9957	0.0187 **	0.3888	0.9513	0.0002 ***	0.0000 ***	0.0040 ***
USD/AUD	2.6155	1.0030	0.0241 **	0.5790	1.0015	0.1523	0.5884	0.9757	0.0320 **	0.0002 ***	0.0505 *
DNK/USD	2.6213	1.0073	0.1115	0.7070	1.0074	0.1164	0.7174	0.9769	0.0206 **	0.0000 ***	0.0216 **
		Monetary model with flexible prices (Frenkel-Bilson) experts + constant									
USD/GBP	2.4728	1.0146	0.1757	0.8537	1.0083	0.1866	0.7143	0.9486	0.0025 ***	0.0000 ***	0.0059 ***
JPY/USD	2.7562	1.0127	0.0733 *	0.8445	1.0072	0.1575	0.7537	0.9643	0.0043 ***	0.0000 ***	0.0043 ***
CHF/USD	2.9775	1.0188	0.4903	0.9199	1.0136	0.3554	0.8337	0.9793	0.0971 *	0.0001 ***	0.1304
CAD/USD	1.4581	1.0230	0.4560	0.9358	1.0069	0.7590	0.9262	0.9796	0.0325 **	0.0007 ***	0.0323 **
SEK/USD	2.5710	1.0281	0.5600	0.9671	1.0124	0.2440	0.8033	0.9515	0.0015 ***	0.0000 ***	0.0055 ***
USD/AUD	2.6155	1.0042	0.0240 **	0.6234	0.9979	0.0111 **	0.4284	0.9806	0.1014	0.0016 ***	0.1455
DNK/USD	2.6213	1.0357	0.7260	0.9917	1.0158	0.5379	0.9080	0.9749	0.0332 **	0.0000 ***	0.0340 **
		Monetary model with flexible prices experts + constant									
USD/GBP	2.4728	1.0194	0.1282	0.8711	1.0149	0.2152	0.8026	0.9503	0.0040 ***	0.0000 ***	0.0062 ***
JPY/USD	2.7562	0.9975	0.0076 ***	0.4447	0.9920	0.0191 **	0.3047	0.9633	0.0036 ***	0.0000 ***	0.0036 ***
CHF/USD	2.9775	0.9994	0.0113 **	0.4845	0.9918	0.0057 ***	0.2944	0.9818	0.1298	0.0001 ***	0.1608
CAD/USD	1.4581	0.9758	0.0007 ***	0.1824	1.0017	0.1088	0.5506	0.9801	0.0402 **	0.0006 ***	0.0554 *
SEK/USD	2.5710	1.0298	0.2800	0.9565	1.0187	0.3913	0.8928	0.9571	0.0038 ***	0.0000 ***	0.0086 ***
USD/AUD	2.6155	0.9898	0.0005 ***	0.3242	1.0014	0.0207 **	0.5410	0.9821	0.1192	0.0015 ***	0.1575
DNK/USD	2.6213	1.0388	0.6534	0.9927	1.0190	0.5897	0.9380	0.9742	0.0384 **	0.0000 ***	0.0382 **
		All experts combined + constant									
USD/GBP	2.4728	1.0343	0.0254 **	0.9406	1.0244	0.1309	0.9077	0.9547	0.0040 ***	0.0000 ***	0.0054 ***
JPY/USD	2.7562	1.0256	0.0309 **	0.8975	1.0148	0.0135 **	0.7829	0.9658	0.0015 ***	0.0000 ***	0.0015 ***
CHF/USD	2.9775	1.0068	0.0034 ***	0.6326	1.0051	0.0103 **	0.6159	0.9784	0.0779 *	0.0001 ***	0.0991 *
CAD/USD	1.4581	0.9863	0.0001 ***	0.2728	1.0060	0.0704 *	0.6765	0.9829	0.0690 *	0.0009 ***	0.0961 *
SEK/USD	2.5710	1.0180	0.0056 ***	0.8190	1.0165	0.0353 **	0.8161	0.9571	0.0011 ***	0.0000 ***	0.0033 ***
USD/AUD	2.6155	0.9959	0.0004 ***	0.4431	1.0084	0.0234 **	0.6941	0.9719	0.0122 **	0.0001 ***	0.0580 *
DNK/USD	2.6213	1.0458	0.1479	0.9858	1.0225	0.2658	0.9214	0.9737	0.0106 **	0.0000 ***	0.0119 **

Table 5: RMSE and Theil ratios for the two-step procedure.

***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

Currency pair	Sequential adaptive ridge			
	RMSE Theil ratio	IID p-value	CW p-value	DMW p-value
Monetary model with flexible prices (Frenkel-Bilson) experts				
USD/GBP	0.9986	0.4317	0.0594 *	0.4317
JPY/USD	1.0000	0.5005	0.1452	0.5005
CHF/USD	1.0048	0.9484	0.9285	0.9024
CAD/USD	0.9970	0.2278	0.0692 *	0.2276
SEK/USD	0.9959	0.2311	0.0208 **	0.2309
USD/AUD	1.0070	0.7103	0.5926	0.7105
DNK/USD	1.0047	0.8574	0.6569	0.8577
All experts combined				
USD/GBP	0.9988	0.4196	0.1182	0.4195
JPY/USD	0.9999	0.4665	0.2474	0.4665
CHF/USD	1.0022	0.9064	0.8367	0.8985
CAD/USD	1.0016	0.6207	0.3926	0.6100
SEK/USD	0.9979	0.3272	0.0989 *	0.3270
USD/AUD	0.9966	0.0911 *	0.0398 **	0.1959
DNK/USD	1.0028	0.8154	0.6393	0.8036
Output gap experts: deviations from a linear and a quadratic trend				
USD/GBP	1.0002	0.5072	0.0614 *	0.5121
JPY/USD	0.9984	0.4122	0.0694 *	0.5000
CHF/USD	1.0032	0.7767	0.4184	0.9700
CAD/USD	0.9998	0.4841	0.1807	0.5000
SEK/USD	1.0005	0.5219	0.0587 *	0.5709
USD/AUD	0.9928	0.1066	0.0298 **	0.1063
DNK/USD	1.0045	0.7000	0.2002	0.7969

Table 6: RMSE and Theil ratios of forecasts for all experts.

***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

Currency pair	Random walk RMSE	Rolling regression			Recursive regression			Sequential adaptive ridge			
		Theil ratio	CW p-value	DMW p-value	Theil ratio	CW p-value	DMW p-value	Theil ratio	IID p-value	CW p-value	DMW p-value
		Coupled classic + coupled Taylor-rule + decoupled output gaps									
USD/GBP	2.4728	1.0502	0.1098	0.9814	1.0269	0.2101	0.9414	0.9547	0.0053 ***	0.0000 ***	0.0057 ***
JPY/USD	2.7562	0.9941	0.0000 ***	0.3848	0.9947	0.0000 ***	0.4043	0.9697	0.0026 ***	0.0000 ***	0.0025 ***
CHF/USD	2.9775	1.0041	0.0006 ***	0.5690	0.9988	0.0010 ***	0.4798	0.9902	0.2985	0.0008 ***	0.3069
CAD/USD	1.4581	1.0053	0.0009 ***	0.5863	1.0027	0.0230 **	0.5892	0.9858	0.1025	0.0020 ***	0.1022
SEK/USD	2.5710	1.0745	0.1878	0.9481	1.0291	0.1587	0.8935	0.9553	0.0013 ***	0.0000 ***	0.0082 ***
USD/AUD	2.6155	1.0013	0.0023 ***	0.5174	1.0097	0.0499 **	0.7479	0.9736	0.0137 **	0.0001 ***	0.0458 **
DNK/USD	2.6213	1.0481	0.3937	0.9921	1.0210	0.3194	0.9151	0.9773	0.0225 **	0.0000 ***	0.0250 **
		Coupled classic + coupled Taylor-rule + decoupled output gaps + constant									
USD/GBP	2.4728	1.0553	0.0344 **	0.9811	1.0312	0.1131	0.9487	0.9541	0.0071 ***	0.0000 ***	0.0083 ***
JPY/USD	2.7562	1.0109	0.0000 ***	0.7037	1.0047	0.0000 ***	0.5835	0.9682	0.0033 ***	0.0000 ***	0.0033 ***
CHF/USD	2.9775	1.0104	0.0006 ***	0.6682	1.0038	0.0009 ***	0.5622	0.9801	0.0997 *	0.0001 ***	0.1266
CAD/USD	1.4581	0.9990	0.0004 ***	0.4845	1.0088	0.0362 **	0.7387	0.9840	0.0874 *	0.0012 ***	0.1065
SEK/USD	2.5710	1.0544	0.0632 *	0.9811	1.0294	0.0819 *	0.9132	0.9508	0.0010 ***	0.0000 ***	0.0059 ***
USD/AUD	2.6155	1.0020	0.0015 ***	0.5200	1.0169	0.0484 **	0.8172	0.9721	0.0142 **	0.0001 ***	0.0521 *
DNK/USD	2.6213	1.0509	0.1367	0.9884	1.0287	0.3526	0.9572	0.9755	0.0223 **	0.0000 ***	0.0240 **
		Decoupled classic, Taylor-rule and output gaps experts + constant									
USD/GBP	2.4728	1.0675	0.0002 ***	0.9813	1.0461	0.0117 **	0.9372	0.9573	0.0107 **	0.0000 ***	0.0117 **
JPY/USD	2.7562	1.0642	0.0002 ***	0.9842	1.0266	0.0000 ***	0.8109	0.9661	0.0029 ***	0.0000 ***	0.0029 ***
CHF/USD	2.9775	1.0567	0.0007 ***	0.9587	1.0304	0.0005 ***	0.8382	0.9801	0.1152	0.0001 ***	0.1379
CAD/USD	1.4581	0.9953	0.0000 ***	0.4398	1.0359	0.0045 ***	0.8868	0.9835	0.0643 *	0.0007 ***	0.0777 *
SEK/USD	2.5710	1.1275	0.0819 *	0.9558	1.1296	0.1445	0.9094	0.9514	0.0005 ***	0.0000 ***	0.0047 ***
USD/AUD	2.6155	1.0336	0.0001 ***	0.7561	1.1286	0.2253	0.9571	0.9761	0.0742 *	0.0004 ***	0.1221
DNK/USD	2.6213	1.0729	0.0060 ***	0.9826	1.2190	0.0084 ***	0.8822	0.9773	0.0372 **	0.0000 ***	0.0598 *
		Coupled classic + coupled Taylor-rule + coupled absolute (level) + decoupled output gaps + constant									
USD/GBP	2.4728	1.1045	0.2305	0.9995	1.0645	0.3078	0.9948	0.9578	0.0089 ***	0.0000 ***	0.0161 **
JPY/USD	2.7562	1.0746	0.0020 ***	0.9968	1.0321	0.0000 ***	0.8907	0.9695	0.0279 **	0.0000 ***	0.0277 **
CHF/USD	2.9775	1.0666	0.0004 ***	0.9755	1.0243	0.0006 ***	0.7925	0.9698	0.0196 **	0.0000 ***	0.0318 **
CAD/USD	1.4581	1.0318	0.0193 **	0.8833	1.0284	0.2340	0.8946	0.9884	0.0204 **	0.0002 ***	0.0250 **
SEK/USD	2.5710	1.0775	0.2691	0.9957	1.0385	0.2096	0.9371	0.9538	0.0009 ***	0.0000 ***	0.0075 ***
USD/AUD	2.6155	1.0334	0.0055 ***	0.8445	1.0407	0.4361	0.9811	0.9860	0.2297	0.0059 ***	0.2449
DNK/USD	2.6213	1.0837	0.2522	0.9990	1.0511	0.2509	0.9883	0.9754	0.0470 **	0.0000 ***	0.0469 **

Appendix for online publication

- A. Overview of the theoretical guarantees for ridge regression methods
- B. Detailed data description
- C. Additional graphs and tables

A. Theoretical guarantees for ridge regression

We provide here an idea of the theoretical guarantees associated with the ridge regression methods; we do so mostly to show that the machine-learning spirit on which these methods are built differs from the classical econometric assumptions and aims. In particular, note that the bounds presented below are extremely robust, because they hold no matter what are the logarithms s_t of the exchange rates or the values $f_{j,t}$ of the fundamentals: no stochastic modeling of any of the latter is required.

In words. In words, the guarantees described below indicate that the RMSE of the methods converge, as the number of instances to be predicted increases, to the RMSE achieved by the best fixed linear model, denoted below by u^* . The latter is therefore smaller than the RMSE of the random walk, as the random walk is a particular linear model (with all coefficients equal to 0 but one equal to 1).

In equations. The bound for the sequential ridge regression (8) with constant regularization factor $\lambda > 0$ follows from the papers by Vovk [2001] and Azoury and Warmuth [2001], see also the summary presented in the monograph by Cesa-Bianchi and Lugosi [2006]. We fix a bound B on the range of the logarithms of the exchange rates and of the fundamentals. For all $\lambda > 0$, for all possible sequences of logarithms of exchange rates $s_t \in [-B, B]$ and of values of the fundamental $f_{j,t} \in [-B, B]$, for all constant linear combinations $\mathbf{u} \in \mathbb{R}^N$, the cumulative square loss of the sequential ridge regression (8) with constant regularization factor $\lambda > 0$ is such that

$$\begin{aligned} \sum_{t=1}^T \left(s_t - s_{t-1} - \sum_{j=1}^N u_{j,t}(\lambda) f_{j,t} \right)^2 - \sum_{t=1}^T \left(s_t - s_{t-1} - \sum_{j=1}^N u_j f_{j,t} \right)^2 \\ \leq \lambda \sum_{j=1}^N u_j^2 + 2NB^2 \left(1 + \frac{NTB^2}{\lambda} \right) \ln \left(1 + \frac{TB^2}{N\lambda} \right). \quad (14) \end{aligned}$$

In particular, when λ is well chosen (e.g., of the order of $1/\sqrt{T}$), one has

$$\sum_{t=1}^T \left(s_t - s_{t-1} - \sum_{j=1}^N u_{j,t}(\lambda) f_{j,t} \right)^2 \leq \sum_{t=1}^T \left(s_t - s_{t-1} - \sum_{j=1}^N u_j^* f_{j,t} \right)^2 + O(\sqrt{T}),$$

where u^* denotes some optimal (in hindsight) constant linear combination. The first term in the right-hand side of the above display corresponds to an approximation error (how well in hindsight the fundamentals can predict the exchange rates) while the second $O(\sqrt{T})$ term is a sequential estimation error (the price to pay for facing a sequential rather than a batch problem).

In particular, as mentioned above, by dividing both members by T and taking roots, this cumulative guarantee leads to a guarantee in terms of RMSE: the RMSE of the sequential ridge regression (8) with a well-chosen constant regularization factor is less than the RMSE achieved using the best linear model u^* plus some term of the order of $T^{-1/4}$, which vanishes as $T \rightarrow +\infty$.

Dependency on the number of fundamentals. The regret bound (14) increases with the number N of experts, while the approximation error decreases with N . As the error suffered by our method is the sum of these two quantities, its behavior as N increases is not clear—and actually, it can either increase or decrease. The theory is mute about its behavior. The practice shows in our case that parsimony is preferred.

Discount factors. The bound (14) can be adapted to handle discount factors (see Mauricette et al. [2009]). Because of the proof technique known so far for this adaptation, the bound always worsens when discount factors are considered (it is maybe an artifact of the analysis). As documented in Mauricette et al. [2009], the practical performance, however, improves with the addition of discounting factors.

Choosing λ . The theoretical guarantees exhibited above heavily rely on picking the right λ . This choice can only be made in hindsight (knowing T , B and the other parameters) and thus, we have no fully sequential method yet. To circumvent that, the articles Devaine et al. [2013], Stoltz [2010] introduced a data-driven meta-method for picking the parameters λ (and μ) of adaptive sequential ridge regression with discount factors, which we detailed in the main body of the paper. Unfortunately, this method comes with no clean theoretical guarantee yet but showed good performance in other contexts (such as the forecasting of electricity consumption, air quality or the production of oil reservoirs).

B. Detailed data description

The dataset of Molodtsova and Papell [2009] was extended to the period 1973–2013 whenever possible – when the series were not discontinued. Series that had to be substituted have Datastream as the principal source. For those, as there were typically more series available without a clear advantage over one another, only the ones that were eventually used to generate predictions presented in the paper are listed (as previously noted, we did not get qualitatively different results using the alternatives). Data was obtained through Datastream on 13/07/2013. When the code from the original data series was known it was given instead. All series are monthly unless noted. Quarterly series were transformed into monthly ones by simple interpolation. The output gap was estimated for each country with at least 27 data points.

Given the availability of the data and the desire to compare different models (“classic” fundamentals, Taylor-rule based exchange rate models) we were able to run the predictions for different currency pairs for the following periods:

Currency pair	Forecasting period
USD/GBP	04/1973 – 04/2013
JPY/USD	04/1973 – 02/2013
CHF/USD	02/1974 – 11/2011
CAN/USD	04/1973 – 03/2013
SEK/USD	04/1973 – 04/2013
USD/AUD	04/1973 – 02/2013
DNK/USD	04/1976 – 02/2013

Table 7: Write a title; refer to the table in the text.

Original Data Series	Source and series name (original series given if possible)
Nominal exchange rate, Japan -- U.S.	FRED, EXJPUS
Nominal exchange rate, Canada -- U.S.	FRED, EXCAUS
Nominal exchange rate, Switzerland – U.S	FRED, EXSZUS
Nominal exchange rate, U.S. – U.K.	FRED, EXUSUK
Nominal exchange rate, Sweden – U.S.	FRED, EXSDUS
Nominal exchange rate, U.S. – Australia	FRED, EXUSAL
Nominal exchange rate, Denmark – U.S.	FRED, EXDNU5
M1 money supply, n.s.a., U.S.	IFS, 11159MA.ZF...
M1 money supply, billions of yens, n.s.a, Japan	IFS, 15834...ZF...
M1 money supply, billions of Canadian dollars, n.s.a., Canada	IFS, 15634...ZF...
Narrow money, billions of Swiss francs, n.s.a., Switzerland	Datastream: SWI34...A, original source: IFS
M0 money supply, millions of British pounds, n.s.a., U.K.	IFS, 11219MC.ZF...
M0 money supply, millions of Swedish kronors, n.s.a., Sweden	Datastream: SDM0....A, original source: Statistics Sweden
M1 money supply, millions of Australian dollars, n.s.a., Australia	IFS, 19334...ZF...
M1 money supply, millions of Danish kroners, n.s.a., Denmark	Datastream: DKOMA027A, original source: OECD, MEI
Federal Funds Rate, United States	IFS, 11160B..ZF...
Money Market rate, Canada	Datastream: CNBKRATER, original source: Statistics Canada
Money Market Rate, Switzerland	IFS, 14660B..ZF...
Money Market Rate, U.K.	IFS, 11260B..ZF...
Money Market Rate, Sweden	IFS, 14460B..ZF...
Money Market Rate, Australia	IFS, 19360B..ZF...
Money Market Rate, Denmark	IFS, 12860B..ZF..
Money Market Rate, Japan	IFS, 15860B..ZF..
Industrial production, s.a., U.S.	IFS, 11166..CZF...
Industrial production, s.a., Japan	IFS, 15866..CZF...
Industrial production, s.a., Canada	Datastream: CNI66..CE, original source: IFS
Industrial production, s.a., Switzerland, quarterly	IFS, 14666..BZF...
Industrial production, s.a., U.K.	IFS, 11266..CZF...
Industrial production excluding construction, s.a., Sweden	Datastream: SDOPRI35G, original source: OECD, MEI
Industrial production, s.a., Australia, quarterly	IFS, 19366..CZF...
Industrial production, s.a., Denmark	IFS, 12866..BZF...
Consumer price index, U.S.	IFS, 11164...ZF...
Consumer price index, Japan	IFS, 15864...ZF...
Consumer price index, Canada	IFS, 15664...ZF...
Consumer price index, Switzerland	IFS, 14664...ZF...
Consumer price index, U.K.	IFS, 11264...ZF...
Consumer price index, Sweden	IFS, 14464...ZF...
Consumer price index, quarterly, Australia	IFS, 19364...ZF...
Consumer price index, Denmark	IFS, 12864...ZF...

C. Additional graphs and tables

This section provides additional tables and figures.

Table 8 reports the quantiles of the differences in forecasting error between the random walk and the methods under scrutiny.

Table 9 shows the results for the Taylor-rule based exchange rate model experts in their “coupled” (homogenous) form.

Figure 1 traces the evolution of the Swiss franc – U.S. Dollar exchange rate and the price level ratio for the period 1973–2013.

Table 8: Quantiles of the differences in forecasting error between the random walk and the method under scrutiny. Distributions of errors that are shifted towards positive values indicate methods that outperform the random walk.

Currency pair	Rolling regression					Recursive regression					Ridge regression				
	10%	25%	50%	75%	90%	10%	25%	50%	75%	90%	10%	25%	50%	75%	90%
	PPP expert														
USD/GBP	-0.98	-0.31	0.00	0.27	1.01	-0.78	-0.32	0.00	0.26	0.81	-0.80	-0.17	0.01	0.32	1.90
JPY/USD	-0.84	-0.32	-0.02	0.25	0.85	-0.27	-0.10	0.00	0.07	0.23	-0.27	-0.04	0.01	0.16	0.74
CHF/USD	-1.32	-0.45	0.00	0.27	0.99	-1.19	-0.47	-0.02	0.40	1.32	-1.61	-0.23	0.05	0.60	2.68
CAN/USD	-0.40	-0.11	0.00	0.10	0.36	-0.17	-0.07	0.00	0.07	0.17	-0.15	-0.04	0.01	0.07	0.21
SEK/USD	-1.19	-0.33	0.00	0.26	0.87	-0.64	-0.27	-0.01	0.24	0.72	-0.49	-0.11	0.02	0.38	1.23
USD/AUD	-0.95	-0.36	-0.03	0.24	0.86	-0.75	-0.37	-0.05	0.28	0.76	-0.38	-0.10	0.00	0.13	0.52
DNK/USD	-1.49	-0.53	-0.03	0.50	1.30	-0.93	-0.41	0.00	0.33	0.98	-0.55	-0.15	0.03	0.50	1.43
	UIRP expert														
USD/GBP	-1.32	-0.23	0.00	0.19	1.03	-1.32	-0.32	0.00	0.27	1.15	-1.52	-0.24	0.01	0.53	3.02
JPY/USD	-0.90	-0.33	0.00	0.25	1.00	-0.56	-0.24	0.00	0.15	0.54	-0.82	-0.09	0.01	0.38	2.07
CHF/USD	-1.55	-0.54	0.00	0.32	1.27	-1.48	-0.60	0.00	0.44	1.56	-2.34	-0.28	0.03	0.68	4.02
CAN/USD	-0.29	-0.10	0.00	0.09	0.28	-0.22	-0.08	0.00	0.08	0.22	-0.22	-0.04	0.00	0.07	0.27
SEK/USD	-1.47	-0.32	-0.01	0.29	1.23	-1.21	-0.42	-0.02	0.42	1.24	-0.88	-0.22	0.02	0.56	2.30
USD/AUD	-1.37	-0.57	-0.06	0.34	1.22	-1.16	-0.57	-0.09	0.41	1.30	-0.82	-0.18	0.00	0.26	1.18
DNK/USD	-1.64	-0.57	-0.02	0.42	1.44	-1.04	-0.31	0.00	0.26	1.05	-0.82	-0.13	0.01	0.63	2.03
	Monetary model with flexible prices (Frenkel-Bilson) experts														
USD/GBP	-2.01	-0.65	-0.02	0.59	1.81	-1.53	-0.46	-0.01	0.40	1.37	-2.16	-0.52	0.06	0.88	3.42
JPY/USD	-2.10	-0.64	0.00	0.51	2.14	-1.21	-0.49	0.00	0.45	1.39	-2.10	-0.48	0.03	1.08	3.35
CHF/USD	-2.48	-0.80	-0.02	0.69	2.68	-2.06	-0.65	-0.01	0.62	1.77	-3.30	-0.76	0.11	1.35	3.72
CAN/USD	-0.82	-0.27	0.01	0.22	0.67	-0.33	-0.14	0.00	0.13	0.32	-0.40	-0.09	0.01	0.19	0.57
SEK/USD	-1.84	-0.67	-0.06	0.38	0.98	-1.19	-0.38	-0.01	0.32	0.87	-1.29	-0.25	0.08	0.88	2.42
USD/AUD	-1.67	-0.52	-0.04	0.43	1.28	-0.76	-0.25	-0.03	0.10	0.65	-1.35	-0.24	0.02	0.47	2.17
DNK/USD	-2.77	-0.92	-0.09	0.57	1.85	-1.54	-0.32	-0.01	0.23	1.06	-1.71	-0.33	0.06	1.00	3.50
	Monetary model with flexible prices experts														
USD/GBP	-2.29	-0.78	-0.01	0.80	1.98	-1.43	-0.47	-0.03	0.39	1.44	-2.30	-0.57	0.06	0.90	3.53
JPY/USD	-2.85	-0.81	-0.01	0.80	2.92	-2.02	-0.58	-0.02	0.67	2.17	-2.25	-0.48	0.03	1.11	3.52
CHF/USD	-3.94	-1.48	0.00	1.43	4.47	-3.08	-0.94	0.00	1.07	4.28	-2.68	-0.74	0.09	1.32	3.79
CAN/USD	-0.93	-0.26	0.03	0.33	1.07	-0.37	-0.11	0.01	0.14	0.40	-0.40	-0.12	0.01	0.18	0.63
SEK/USD	-2.11	-0.87	-0.08	0.42	1.27	-1.32	-0.41	-0.02	0.30	0.90	-1.31	-0.26	0.08	0.89	2.53
USD/AUD	-1.72	-0.65	-0.06	0.47	1.84	-0.93	-0.34	-0.04	0.09	0.62	-1.35	-0.24	0.02	0.47	2.17
DNK/USD	-3.07	-1.11	-0.10	0.58	2.36	-1.78	-0.34	-0.01	0.26	1.07	-2.09	-0.42	0.06	1.14	3.79
	All experts combined														
USD/GBP	-4.20	-1.35	-0.08	0.85	3.05	-2.06	-0.57	-0.02	0.47	1.52	-2.05	-0.42	0.04	0.91	3.85
JPY/USD	-3.89	-1.29	-0.10	0.99	3.57	-2.74	-1.01	-0.03	1.02	3.12	-1.81	-0.45	0.05	0.94	3.12
CHF/USD	-4.31	-1.33	-0.03	1.09	4.69	-3.33	-1.17	-0.09	1.08	3.80	-3.32	-0.91	0.12	1.59	4.23
CAN/USD	-1.14	-0.33	0.02	0.39	1.14	-0.47	-0.18	0.01	0.19	0.52	-0.52	-0.13	0.00	0.17	0.56
SEK/USD	-3.71	-1.35	-0.06	0.89	3.03	-2.51	-0.88	-0.07	0.64	2.59	-1.31	-0.33	0.06	0.88	2.60
USD/AUD	-3.02	-1.02	-0.04	0.67	3.31	-2.25	-0.65	-0.01	0.62	1.98	-1.20	-0.28	0.02	0.48	2.26
DNK/USD	-5.05	-1.70	0.00	1.41	3.94	-2.11	-0.63	0.02	0.71	2.45	-1.52	-0.36	0.12	1.13	3.29

Table 9: RMSE and Theil ratios of forecasts for the Taylor-rule based exchange models. “Coupled” experts.

***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels.

Currency pair	Random walk RMSE	Rolling regression			Recursive regression			Sequential adaptive ridge			
		Theil ratio	CW p-value	DMW p-value	Theil ratio	CW p-value	DMW p-value	Theil ratio	IID p-value	CW p-value	DMW p-value
		Output gap experts: deviations from a linear trend									
USD/GBP	2.4728	1.0220	0.8014	0.9802	1.0140	0.4952	0.9149	0.9754	0.0663 *	0.0000 ***	0.0797 *
JPY/USD	2.7562	1.0157	0.4115	0.9082	1.0083	0.3633	0.7971	0.9838	0.2082	0.0002 ***	0.2079
CHF/USD	2.9775	1.0215	0.1559	0.9293	1.0155	0.2245	0.8853	0.9868	0.2357	0.0001 ***	0.2495
CAD/USD	1.4581	1.0137	0.1844	0.8731	1.0043	0.1616	0.7489	0.9927	0.2024	0.0049 ***	0.2022
SEK/USD	2.5710	1.0300	0.3927	0.9173	0.9985	0.0933 *	0.4770	0.9669	0.0146 **	0.0002 ***	0.0576 *
USD/AUD	2.6155	1.0161	0.0682 *	0.8610	1.0086	0.1073	0.7702	1.0231	0.7586	0.4166	0.7588
DNK/USD	2.6213	1.0079	0.2940	0.7808	0.9984	0.0992 *	0.4208	0.9957	0.3637	0.0300 **	0.3956
		Output gap experts: deviations from a quadratic trend									
USD/GBP	2.4728	1.0216	0.7842	0.9789	1.0132	0.4953	0.9018	0.9840	0.1995	0.0000 ***	0.1992
JPY/USD	2.7562	1.0117	0.0453 **	0.7706	1.0064	0.1199	0.6989	0.9836	0.1560	0.0001 ***	0.1557
CHF/USD	2.9775	1.0191	0.3375	0.9641	1.0140	0.5631	0.9484	0.9894	0.2329	0.0004 ***	0.2326
CAD/USD	1.4581	1.0132	0.2526	0.8660	1.0036	0.2515	0.7283	0.9969	0.3547	0.0099 ***	0.3546
SEK/USD	2.5710	1.0307	0.4223	0.9239	1.0003	0.0899 *	0.5048	0.9698	0.0110 **	0.0001 ***	0.0315 **
USD/AUD	2.6155	1.0154	0.0673 *	0.8582	1.0085	0.1121	0.7728	1.0229	0.7400	0.3577	0.7402
DNK/USD	2.6213	1.0082	0.2979	0.7887	1.0009	0.1793	0.5389	0.9978	0.4097	0.0450 **	0.4385
		Output gap experts: deviations from a linear and a quadratic trend									
USD/GBP	2.4728	1.0216	0.7462	0.9854	1.0107	0.3737	0.8878	0.9743	0.0588 *	0.0000 ***	0.0633 *
JPY/USD	2.7562	1.0108	0.0157 **	0.7382	1.0074	0.1599	0.7395	0.9858	0.2347	0.0001 ***	0.2344
CHF/USD	2.9775	1.0188	0.4387	0.9742	1.0144	0.7482	0.9780	0.9967	0.4141	0.0002 ***	0.4140
CAD/USD	1.4581	1.0127	0.3073	0.8487	1.0048	0.5449	0.8764	0.9992	0.4701	0.0225 **	0.4700
SEK/USD	2.5710	1.0244	0.4747	0.9131	1.0028	0.1138	0.5531	0.9726	0.0219 **	0.0008 ***	0.0640 *
USD/AUD	2.6155	1.0122	0.0526 *	0.8291	1.0072	0.0674 *	0.7243	1.0173	0.6994	0.2303	0.6996
DNK/USD	2.6213	1.0108	0.5356	0.9014	1.0032	0.4397	0.7117	0.9950	0.2360	0.0175 **	0.2755
		Output gap experts: deviations from a Hodrick-Prescott filtered trend									
USD/GBP	2.4728	1.0175	0.5368	0.9398	1.0101	0.3582	0.8475	0.9705	0.0523 *	0.0000 ***	0.0521 *
JPY/USD	2.7562	1.0012	0.0306 **	0.5539	1.0034	0.2014	0.6908	0.9876	0.1380	0.0001 ***	0.1378
CHF/USD	2.9775	1.0173	0.5770	0.9771	1.0138	0.7907	0.9742	0.9928	0.2468	0.0020 ***	0.2474
CAD/USD	1.4581	1.0040	0.0517 *	0.6384	0.9990	0.1311	0.4324	1.0024	0.5738	0.0388 **	0.5739
SEK/USD	2.5710	1.0288	0.6488	0.9389	0.9997	0.1011	0.4947	0.9737	0.0220 **	0.0007 ***	0.0964 *
USD/AUD	2.6155	1.0198	0.1185	0.7793	1.0092	0.1464	0.7967	1.0086	0.6021	0.0702 *	0.6022
DNK/USD	2.6213	1.0191	0.8561	0.9834	1.0056	0.5828	0.7927	1.0011	0.5308	0.1150	0.5275

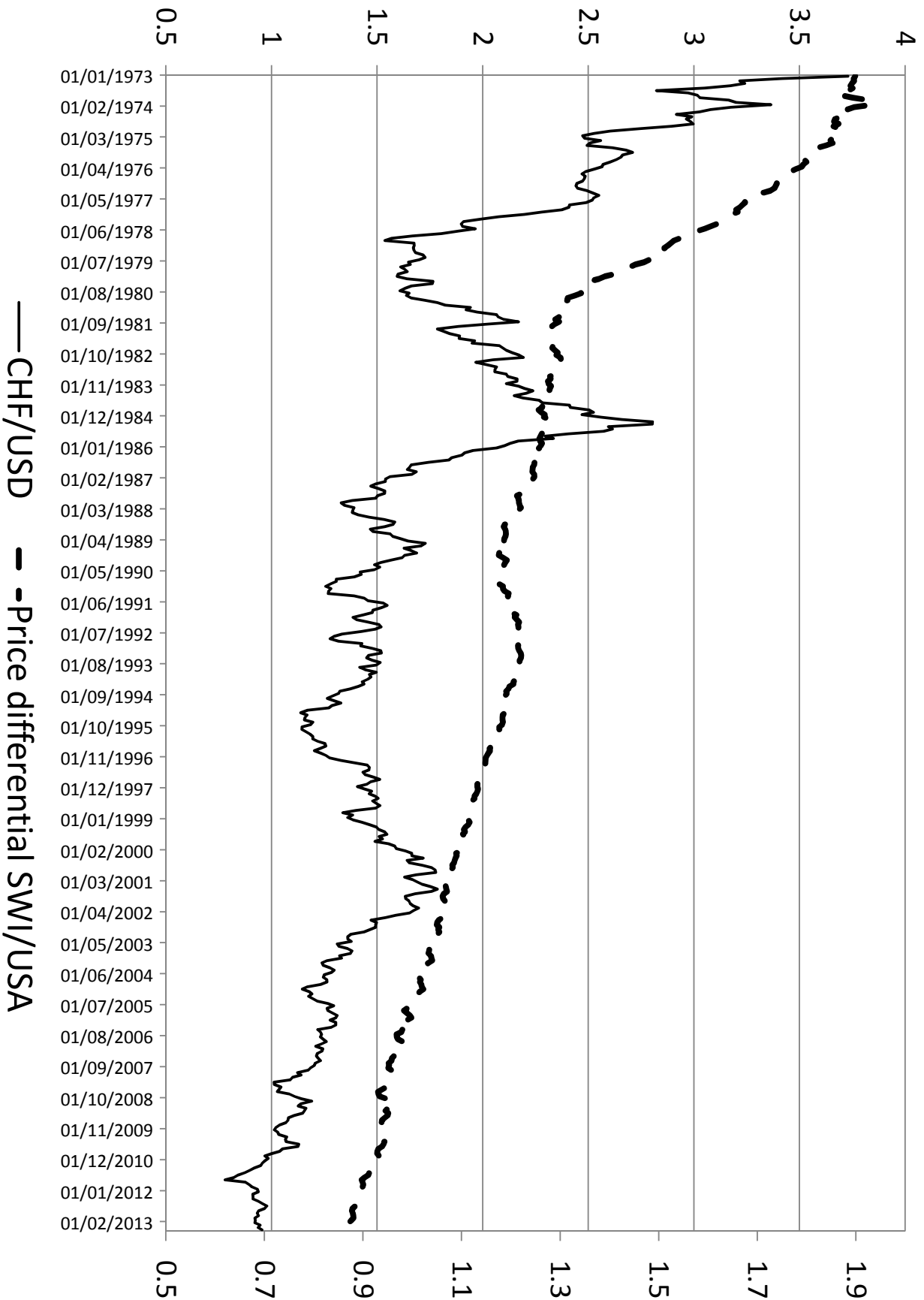


Figure 1: Ratio of Swiss vs. US price level (right-hand scale) and Swiss Franc/U.S. Dollar exchange rate (left-hand scale), 1973–2013