Inter-American Development Bank

Brazilian Exchange Rate Forecasting in High Frequency

Country Department Southern Cone

José Rossi (Inter-American Development Bank)
Carlos Piccioni (University of Brasilia)
Marina Rossi (University of Brasilia)
Daniel Cajueiro (University of Brasilia)

TECHNICAL NOTE Nº IDB-TN-2561

September 2022

Brazilian Exchange Rate Forecasting in High Frequency

José Rossi (Inter-American Development Bank)
Carlos Piccioni (University of Brasilia)
Marina Rossi (University of Brasilia)
Daniel Cajueiro (University of Brasilia)

September 2022

Inter-American Development Bank

Cataloging-in-Publication data provided by the
Inter-American Development Bank
Felipe Herrera Library

Brazilian Exchange Rate Forecasting in High-Frequency / José Rossi, Carlos Piccioni, Marina Rossi, Daniel Cajueiro.
p.cm. - (IDB Tecnical Note ; 2561)
includes bibliographical references.
1. Foreign exchange rate-Forecasting-Brazil. 2. Machine Learning. I. Rossi Júnior, José Luiz. II. Piccioni, Carlos. III. Rossi, Marina. IV. Cajueiro, Daniel. V. Country Department Southern Cone. VI. Serie.
IDB-TN-2561

http://www.iadb.org

Brazilian Exchange Rate Forecasting in High Frequency

José Rossi (Inter-American Development Bank)
Carlos Piccioni (University of Brasilia)
Marina Rossi (University of Brasilia)
Daniel Cajueiro (University of Brasilia)

Abstract

We investigated the predictability of the Brazilian exchange rate at High Frequency (1, 5 and 15 minutes), using local and global economic variables as predictors. In addition to the Linear Regression method, we use Machine Learning algorithms such as Ridge, Lasso, Elastic Net, Random Forest and Gradient Boosting. When considering contemporary predictors, it is possible to outperform the Random Walk at all frequencies, with local economic variables having greater predictive power than global ones. Machine Learning methods are also capable of reducing the mean squared error. When we consider only lagged predictors, it is possible to beat the Random Walk if we also consider the Brazilian Real futures as an additional predictor, for the frequency of one minute and up to two minutes ahead, confirming the importance of the Brazilian futures market in determining the spot exchange rate.

JEL CLASSIFICATION: N76; O13; C22; C53; Q47
KEYWORDS: Exchange Rate; Forecasting; High Frequency; Brazil

1. Introduction

It is practically a consensus in the literature that foreign exchange rates are hard to predict. Meese and Rogoff (1983) showed that traditional macroeconomic models do not outperform a naive random walk in pseudo out-of-sample forecasting exercise, in medium and short-term prediction horizons (one month to one year). Cheung et al. (2005) and Rossi (2013) evaluated the models developed in the following decades and conclude that no model performed very well: exchange rates predictability, for some models, depend on the choice of forecast horizon, exchange rates, sample period and the toughest benchmark to beat is the random walk without drift. In general, the macroeconomic models used in these studies make use of macroeconomic variables available only at low frequencies.

At higher frequencies, Evans and Lyons (2002) opened a new research area: how microstructure (order flows) affects the exchange rates. Evans and Lyons (2005) showed that daily customer order flow, from one day to one month, can beat the Random Walk in a real out-of-sample setup, same conclusion as Rime et al. (2010) with daily data. However, in opposition, Danielsson et al. (2012) concludes that predictability is only valid at higher frequencies (intraday data).

In this paper we go in a complementary direction to the micro-structure literature. Our contribution to the literature is to evaluate the predictability of the Brazilian exchange rate[1] at high frequency (1, 5 and 15 minutes) using high frequency local and global economic variables, namely: short and long-term Brazilian interest rates, Brazilian stock market index, gold price, oil price, stock market option-based implied volatility (VIX) and exchange rates of 17 other countries. In addition to the linear regression method, we also evaluate the use of Machine Learning algorithms such as Ridge, Lasso, Elastic Net, Random Forest and Gradient Boosting in the prediction exercises.

We choose the variables based on their availability in the desired frequency (maximum frequency of 1 minute) and their possible relationship with the Brazilian exchange rate. For example, short and long-term interest rates on Interbank Deposit (DI) contracts reflect expectations regarding monetary and fiscal policy, both of which affect the exchange rate. The relationship between the stock exchange and the exchange rate is also a recurrent theme of research, as in Tabak (2006). Ferraro et al. (2015) conjecture that, in small open commodity-exporting economies, the exchange rate is expected to reflect the movement of commodity prices. Beckmann et al. (2017) also shows that oil price is a potentially exchange rate predictor for the short run. We can also see VIX as a measure of uncertainty (Bekaert et al. (2013)),

[1] The Brazilian case is interesting since its exchange rate is floating according to the IMF classification, Brazil is one of the largest emerging economies, and the Brazilian Real has shown high volatility and moments of great depreciation in recent years, with significant impacts on the macroeconomic environment. However, exchange rate forecasting papers for the Brazilian currency focus on monthly and daily frequencies, as in Gaglianone and Marins (2017) and Moura et al. (2008).

and it can explain part of the daily variation of the nominal exchange rate (Kohlscheen et al. (2017)). Motivated by Felício and Rossi Júnior (2014), who extract common factors from a set of floating exchange rates and assess their predictive capacity, we consider 17 foreign exchange rates as possible predictors. We also investigate gold as a possible predictor, due to a possible bi-directional causal relationship with exchange rates in emerging countries (Gürış and Kiran (2014), Nair et al. (2015)).

We perform two types of forecasting exercises. The first, called out-of-sample fit, uses contemporary data as predictors (realized values of the predictors variables). This type of analysis captures correlations or co-movements. As put by Ferraro et al. (2015) in the forecasting literature, this type of prediction can be useful when we are interested in evaluating the predictive capacity of a model given a trajectory for some unmodeled set of variables. In other words, if it is possible to obtain a good model to predict this variable, then this model can be exploited to predict the exchange rate. Important examples of its use are the Meese and Rogoff (1983) and Cheung et al. (2005), which showed that, even using realized values of the predictors variables, traditional models were unable to beat the Random Walk in the exchange rate prediction. The second exercise is the real out-of-sample forecast, in which we seek to forecast the exchange rate at $t+1$ using information available only up to t.

We verified that, in the out-of-sample fit exercise, for each variable, it is possible to predict the Brazilian exchange rate at high frequency with less error than the Random Walk. We also verified that the local economic variables present lower mean squared error than the global variables. The augmented model, considering all variables, outperforms all individual models. Machine learning models further reduce the MSE when we apply them to the augmented model. When applying the Gradient Boosting, we also estimate the relative importance of each predictor, showing that the order of importance of the predictors changes according to the frequency considered, with Ibovespa, short-term interest rates and the Mexican exchange rate occupying the top positions.

However, no model maintains the predictive capacity in the real out-of-sample forecasting exercise, except when we also consider the Brazilian Real-U.S. dollar futures as an additional predictor variable. In this case, for some models, it is possible to beat the random walk at frequencies of 1 minute and for 1 and 2 minutes ahead. As stated by Ventura and Garcia (2012), in Brazil, the exchange rate is firstly determined at the exchange rate future market (the next maturity), and transmitted by arbitrage to the spot market. So, another contribution of this paper is to show how this relationship translates into out of sample prediction ability at high frequency.

In relation to the use of Machine Learning techniques, our paper differ from works like Colombo and Pelagatti (2020), Zhang and Hamori (2020) and Amat et al. (2018) in terms of frequency and set of economic variables. At high frequency, most works use technical trading strategy or univariate strategies for prediction, such as Manahov et al. (2014), Palikuca and Seidl

(2016), or combine microstructure with Machine Learning, as Choudhry et al. (2012). We differ from these evaluating a different set of predictors and benchmarking against random walk.

The paper is structured as follows. In Section 2 we describe the database. In Section 3 we present the models used for prediction and the methodology used to determine the parameters and hyper-parameters. In Section 4 we present and discuss the results of the out-of-sample fit and real out-of-sample prediction exercises. Section 5 concludes.

2. Data

We use intraday data from 2021-05-05 to 2021-11-12 (128 business days), 1 minute frequency (closing prices), from Bloomberg, resulting in a total of 46,080 observations. For exercises on frequencies of 5 and 15 minutes, we re-sample the data. On each day, the interval considered is from 10:00 am to 4:00 pm. Although the Brazilian foreign exchange market operates from 9:00 am to 6:00 pm, the stock exchange opens at 10:00 am and the negotiation of future interest rates contracts is halted at 4:00 pm.

Our interest is in forecasting the Brazilian Real/U.S. dollar nominal spot exchange rate. As candidates for predictor variables, we consider, for short and long-term interest rates, *DI Futures* contracts maturing in January 2023 (DI23) and January 2029 (DI29). The DI futures underlying asset is the average daily interest rate of interbank deposits (DI). The notional value of one DI future contract is R$100 thousand, and the value on the trade date is equivalent to this amount discounted at the negotiated rate. This rate reflects the expected evolution of the DI, that is, the expectations regarding the future interest rate (Vartanian et al. (2021)). As stated by Jeanneau et al. (2007) the Brazilian futures market is one of the main indicators of interest rates expectations, with the yield curve implicit in DI futures being the main benchmark for fixed income investment in Brazil. As a representative of the back end of the yield curve, we chose the contracts maturing in 2029 because they are the most liquid of the longer term futures contracts.

For oil price we use the West Texas Intermediate (WTI) crude oil price. The stock market index used was the Ibovespa from São Paulo Stock Exchange. We also include the gold price and the Chicago Board Options Exchange Volatility Index (VIX). For the true out-of-sample exercise, we use as predictor the Brazilian Real/U.S. dollar futures (next maturity).

We selected other 17 nominal exchange rates as predictors based on the following criteria: they are floating exchange rates; they are traded at the same times as the Brazilian exchange rate; they are from countries (and economic alliances) that have a GDP of at least 10% of the Brazilian GDP. The countries and economic alliances that met these criteria were: Australia (USDAUD), Canada (USDCAD), Czech Republic (USDCZK), Euro Area (USDEUR), Israel (USDILS), Japan (USDJPY), Mexico (USDMXN), New Zealand (USDNZD), Norway (USDNOK), Poland (USDPLN), Russian Federation (USDRUB), South Africa (USDZAR), Sweden (USDSEK), Switzerland (USDCHF), Thailand (USDTHB), Turkey (USDTRY) and England

(USDGBP).

So as to achieve stationarity, all variables are in (log) first differences. Only intraday differences are considered, that is, day-to-day differences are discarded.

3. Methods

In the next subsection, we present the models considered in this paper, followed by the strategy for determining the parameters and hyperparameters, as well as the performance evaluation methods used.

3.1. Exchange Rate Forecasting Model

Let s_t be the logarithm of the exchange rate, our interest is in predicting the change in the logarithm of the exchange rate one step ahead, that is, $\Delta s_{t+1} = s_{t+1} - s_t$. Let $\mathbf{x}_{t+1}$ be a set of predictors, with information up to $t+1$, then our general **out-of-sample fit** prediction model is defined by:

$$\Delta s_{t+1} = f(\mathbf{x}_{t+1}) + \varepsilon_{t+1} \tag{1}$$

where $f()$ is a mapping to be estimated, and ε_{t+1} is the prediction error. Note that the out-of-sample fit exercise use contemporary data (realized values of the predictors variables) [2].

The model of our second exercise, **out-of-sample** forecasting, is defined by:

[2]Note that this is not an in-sample prediction exercise, as $\mathbf{x}_{t+1}$ will not be used for parameter prediction.

$$\Delta s_{t+1} = f(\mathbf{x}_t) + \varepsilon_{t+1} \tag{2}$$

where $f()$ is also a mapping to be estimated and $\mathbf{x}_t$ a set of predictors with information only up to t.

For both exercises, we initially evaluate the predictive power of the variables using $f()$ as a Linear Regression Model, estimated on a rolling window, as presented in Subsection 3.2. So the forecasting model is defined, for the out-of-sample fit exercise, by:

$$\Delta s_{t+1}^{f} = \mathbf{x}_{t+1}'\hat{\beta} \tag{3}$$

where $\hat{\beta}$ is a $k \times 1$ vector of parameters to be estimated, and $\mathbf{x}_{t+1}$ a $k \times 1$ vector of the predictors, including a constant term. For the out-of-sample forecasting exercise, the Linear Regression Model is defined by:

$$\Delta s_{t+1}^{f} = \mathbf{x}_t'\hat{\beta} \tag{4}$$

For both the out-of-sample fit and out-of-sample forecasting exercises, we initially look at the individual predictive capabilities of the variables considered in this paper. Next, we group some of the variables, testing what we call augmented models. When increasing the models, we also start to consider for the regression $f()$ Machine Learning methods such as LASSO, Ridge, ElasticNet, Random Forest, Support Vector Regression and Extreme Gradient Boosting, described below. Equations (3) and (4) are also valid for regularized linear Machine Learning Models, differing, for each model, the form of estimation of $\hat{\beta}$, as will be seen in the next subsections.

3.1.1. Ridge

According to Masini et al. (2020), Ridge regression, proposed by Hoerl and Kennard (1970), seeks to combat problems generated by multicollinearity in Linear Regression, stabilizing the problem solution by introducing a small bias in exchange for reducing the variance of the estimator. Ridge penalizes the regression by the L_2 norm of the parameter vector. Therefore, $\hat{\beta}$ of the equations (3) and (4) are determined by [3]:

$$\hat{\beta} = \underset{\beta_1,\ldots,\beta_k}{\arg\min} \left[\frac{1}{T} \sum_{t=1}^{T} \left(\Delta s_{t+1} - \sum_{j=1}^{k} x_{j,t} \beta_j \right)^2 + \lambda \sum_{j=1}^{k} \beta_j^2 \right] \tag{5}$$

We determine λ, also called hyperparameter in Machine Learning literature, by cross-validation, as we present in the Subsection 3.2, or through the use of information criteria such as AIC or BIC. The greater the λ, the greater the shrinkage of the coefficients. For $\lambda = 0$ the model collapses to an OLS.

According to Zou and Hastie (2005), empirically it has been observed better performance of Ridge in relation to LASSO in the presence of significant multicollinearity of the predictors. However, Ridge does not generate parsimonious models: the least relevant predictors have their coefficients shrunk

[3] In the Ridge, Lasso and Elastic Net descriptions, we use the notation used by Costa et al. (2021).

towards zero, but never exactly zero.

3.1.2. LASSO

LASSO (Least Absolute Shrinkage and Selection Operator), proposed by Tibshirani (1996), penalizes the regression by the norm L_1 of the parameter vector $\hat{\beta}$. That is, $\hat{\beta}$ of the equations (3) and (4) are determined by:

$$\hat{\beta} = \underset{\beta_1,\ldots,\beta_k}{\arg\min} \left[\frac{1}{T} \sum_{t=1}^{T} \left(\Delta s_{t+1} - \sum_{j=1}^{k} x_{j,t}\beta_j \right)^2 + \lambda \sum_{j=1}^{k} |\beta_j| \right] \tag{6}$$

Given the penalty by the l_1 norm, the solution to the problem is sparse. So, in addition to performing shrinkage, LASSO selects variables, zeroing out the irrelevant coefficients. As such, it is one of the most popular regularization methods in data-rich environments (Masini et al. (2020)). As in Ridge, $\lambda = 0$ leads to OLS and the determination of λ can be done by cross-validation or through the use of information criteria such as AIC or BIC.

3.1.3. Elastic Net

Proposed by Zou and Hastie (2005), Elastic Net's idea is to combine the advantages of Ridge and LASSO, where $\hat{\beta}$ is determined by:

$$\hat{\beta} = \underset{\beta_1,\ldots,\beta_k}{\arg\min} \left[\frac{1}{T} \sum_{t=1}^{T} \left(\Delta s_{t+1} - \sum_{j=1}^{k} x_{j,t}\beta_j \right)^2 + \lambda \left(\alpha \sum_{j=1}^{k} \beta_j^2 + (1-\alpha) \sum_{j=1}^{k} |\beta_j| \right) \right] \tag{7}$$

with $\alpha \in [0, 1]$, that is, the Elastic Net penalty is a convex combination of the L_1 penalties, which performs variable selection, and the L_2 penalty, which stabilizes the solution of the problem (Masini et al. (2020)). λ, in turn, determines the overall strength of the penalties. Note that Ridge and LASSO regressions are special cases of Elastic Net, for $\alpha = 1$ and $\alpha = 0$ respectively.

Simulation studies have shown that Elastic Net has better predictive power than LASSO, while maintaining a similar sparse representation (Zou and Hastie (2005)).

3.1.4. Random Forest

Proposed by Breiman (2001), Random Forest is basically an ensemble of decision trees. Decision trees recursively partition the variables domain into rectangular regions, without overlapping and each region R_i is associated with a constant value in the case of a regression problem, so that visually they are step functions:

$$f(\mathbf{x}) = \sum_i c_i I(\mathbf{x} \in R_i) \tag{8}$$

In the Random Forest, each tree is built based on a bootstrap sample of training data. For each tree, the following steps are repeated for each terminal node, until a certain stopping criterion is met (for example, controlling the maximum depth or number of leaves) (Friedman (2017)):

(i) Randomly select l variables from the available k variables;

(ii) Select the best variable / split-point among the l variables (the combination variable / split-point that minimizes the mean squared error) and split the node into two child nodes.

Let B be the number of bootstrap samples, the ensemble of trees $\{T_b\}_1^B$ is stored, and the prediction about a new point $\mathbf{x}$ is given by the average of each tree prediction:

$$\Delta s_{t+1}^f = \frac{1}{B}\sum_{b=1}^{B} T_b(\mathbf{x}_t) \tag{9}$$

where $T_b(\mathbf{x})$ is the prediction of the b-th regression tree. That is, the general idea of Random Forest is to generate an average from several unbiased noisy models (trees), as in bagging (bootstrap aggregation), but reducing the correlation between the trees, so that the variance of the average is reduced. This is achieved by randomly selecting sets of variables in each division of the tree (Friedman (2017)).

3.1.5. Gradient Boosting

Proposed by Friedman (2001), Gradient Boosting is based on 'weak learners', trees with low predictive power, in an additive, step-by-step model. At each step, a weak learner is trained to compensate the constraints of previous weak learners. We can illustrate its operation with the algorithm presented in Friedman (2017), Section 10.10. For a loss function $L(y, f(x))$, N observations in training data, M weak learners:

1. Initialize the optimal constant model, which is just a simple terminal

node of a tree: $f_0(\mathbf{x}) = \arg\min_\gamma = \sum_{i=1}^{N} L(y_i, \gamma)$;

2. For $m = 1, 2, ..., M$:

   (a) For $i = 1, 2, ..., N$ compute:

$$r_{im} = -\left[\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)}\right]_{f=f_{m-1}}$$

   (b) Fit a regression tree to the targets r_{im}, giving terminal regions R_{jm}, $j = 1, 2, ..., J_m$.

   (c) For $j = 1, 2, ..., J_m$ compute:

$$\gamma_{jm} = \arg\min_{\gamma} \sum_{x_i \in R_{jm}} L(y_i, f_{m-1}(x_i) + \gamma)$$

   (d) Update $f_m(x) = f_{m-1}(x) + \sum_{j=1}^{J_m} \gamma_{jm} I(x \in R_{jm})$

3. The final forecast will be given by $f_M(\mathbf{x})$

In this work we use the Extreme Gradient Boosting (XGB) version developed by Chen and Guestrin (2016), which is an improved version of Gradient Boosting, both with regard to predictive performance and computational performance [4].

[4]For example, regressions trees can include a regularization parameter, which influences the pruning tree mechanism, acting against overfiting, also reducing the sensitivity of predictions to individual observations. The algorithm is also known for optimizations for large datasets. For example, it can use an Approximate Greedy Algorithm, which does not test every possible threshold on each split of a tree, but instead use quantiles as candidate thresholds for each split. Using Parallel Learning, it also allows the dataset to be split across multiple computers at the same time. It also innovates with a native mechanism to build trees even when there are missing values (Sparsity-Aware Split Finding). It also seeks to optimize the use of computer hardware, for example, by storing gradients in

3.1.6. Support Vector Regression

Support Vector Regression (SVR) extends the concepts used by Support Vector Machines (SVM) in classification to regression problems. The idea is to form a 'tube' or 'band' around the actual regression function that contains most of the observations. We initially consider the linear case. Let the true Linear Regression function:

$$\mu(\mathbf{x}) = \beta_0 + \mathbf{x}'\beta \tag{10}$$

Define the following loss function (linear ε-Insensitive Loss Function):

$$L_1^{\varepsilon}(y, \mu(\mathbf{x})) = max\{0, |y - \mu(\mathbf{x})| - \varepsilon\} \tag{11}$$

That is, if the point $(\mathbf{x}, y)$ is such that $|y - \mu(\mathbf{x})| \leq \varepsilon$, then the loss is zero, and if $|y - \mu(\mathbf{x})| > \varepsilon$, the loss is $|y - \mu(\mathbf{x})| - \varepsilon$.

Points not inside the 'tube' are described by slack variables, ξ_i' and ξ_j. Define ξ_i' and ξ_j such that if the point lies above ε-tube, then $\xi_i' = y_i - \mu(\mathbf{x}_i) - \varepsilon \geq 0$, and if it is below, then $\xi_j = \mu(\mathbf{x}_j) - \varepsilon - y_j \geq 0$. For points inside ε-tube, the slack variables have a zero value. The primal optimization problem is to find β_0, β, $\xi_1', ..., \xi_n'$ and $\xi_1, ..., \xi_n$ for:

---

the processor's cache, so that it can more quickly calculate the scores used to define the split points of trees (Cache-Aware Access). It is also able to natively use techniques such as Sharding, which divides data between more than one storage unit in order to access them in parallel (Blocks for Out-of-Core Computation). Due to these characteristics - some of them implementation-specific and not related to statistics - which considerably improve the execution time, XGB is one of the favorite algorithms in Machine Learning competitions (Costa et al. (2021)).

$$
\begin{aligned}
& minimize \quad \frac{1}{2}||\beta||^2 + C\sum_{i=1}^{n}(\xi_i' + \xi_i) \\
& subject\ to \quad y_i - (\beta_0 + \mathbf{x_i'}\beta) \leq \varepsilon + \xi_i' \\
& \qquad (\beta_0 + \mathbf{x_i'}\beta) - y_i \leq \varepsilon + \xi_i \\
& \qquad \xi_i' \geq 0, \xi_i \geq 0, \quad i = 1, 2, ..., n
\end{aligned}
\tag{12}
$$

The solution to this problem is detailed in chapter 11 of Izenman (2008), and produces a linear function of $\mathbf{x}$ surrounded by a tube of radius ε. Points that do not fall into the tube are called *support vectors*. The formulation can also be extended to the non-linear case, through the use of kernels such as the radial basis function. In our exercises, we use the non-linear version of the SVR. C acts as a regularization parameter and, together with ε, are the SVR hyperparameters that can be determined by cross-validation.

3.2. Parameters and Hyper-parameters determination

For both Linear Regression and Machine Learning models, we estimate the model parameters by a rolling window, as follows: we split the data into training data, which correspond to the first half of the observations, and test data, the final half. We estimate the model parameters on the training data, and then the prediction of the next observation is made. The next observation of the testing data is then incorporated into the training data, the original first observation of the training data is discarded, we re-estimate the parameters over this rolling window, and we perform a new one-step-ahead prediction. We repeat this process until the rolling window has cycled through all of the test data.

In the case of Machine Learning models, we determine the hyperparameters in three different ways [5]. For Ridge, LASSO and Elastic Net, when each rolling window advances one step, a grid search algorithm tests possible hyperparameter combinations as follows: for each hyperparameter combination, we determine the model parameters over 62.5% of the initial rolling window observations and use them to predict the 12.5% of subsequent rolling window observations. We compute and store the mean squared error and repeat the process using 75% of the initial rolling window data for parameter estimation, with the prediction performed on the 12.5% of the following data, computing and storing the MSE. We repeat the process again, using 87,5% of the rolling window data for parameter estimation and the last 12,5% for prediction, so that all the rolling window observations are used. We select the combination of hyperparameters that presents the smallest mean squared error for the final estimation of the parameters over the entire rolling window / training data.

For algorithms with higher computational costs, such as Random Forest, Extreme Gradient Boosting and Support Vector Regression, we perform the process described in the previous paragraph only once, on the initial training data. In other words, the hyperparameters are not determined again at each advance of the rolling window: we determine them at the beginning of the

[5] We look for the best ways to determine hyperparameters given the restrictions in terms of computational costs.

process and we use same hyperparameters for every advance of the rolling window. We also use this strategy for Ridge, LASSO and Elastic Net for comparison purposes.

In the case of LASSO, as a third way of determining hyperparameters, we also use the information criteria AIC and BIC. For the Ridge, LASSO and Elastic Net methods, in which we use more than one method to determine the hyperparameters, we report the results of the method with the smallest mean squared error.

3.3. Model Valuation

The Random Walk without drift is used as the benchmark, the model to beat. It corresponds to the forecast of no change in the exchange rate one step ahead, that is:

$$\Delta s_{t+1}^{f} = 0 \tag{13}$$

For the Random Walk and for the estimated models, the mean squared errors (MSE) are computed and the ratios between the MSE of the models in relation to the Random Walk are reported. That is, values lower than 1 for this ratio suggest a better predictive capacity of the model in relation to Random Walk.

We apply the Diebold and Mariano test (Diebold and Mariano (1995)) to determine whether the model forecasting compared to the random walk forecasting are statistically different.

4. Results

Initially, we present the results for the out-of-sample fit exercise, and then the results for the out-of-sample forecasting exercise.

4.1. Out-of-sample-fit results

Table 1 presents the results of the out-of-sample fit prediction exercise of the Brazilian exchange rate, for frequencies of 1, 5 and 15 minutes, using Linear Regression and also Machine Learning algorithms for the augmented model. If it is possible to predict the values of any of these variables - changes in future interest rates, Ibovespa, oil price, VIX, gold price, or some of the 17 exchange rates considered - in $t+1$ (perfect foresight), then it is also possible to predict the Brazilian exchange rate at high frequency with less error than the Random Walk, for all frequencies tested, with exception for gold as a predictor for the 15-minute frequency [6].

Regarding the Ibovespa, our results on the intraday frequency are compatible with the conclusions of Tabak (2006) for the daily frequency, who found Granger causality from stock prices to the exchange rate. For oil price as a predictor, our results in high frequency are similar to Ferraro et al. (2015) in the Canadian dollar forecasting at daily frequency. For long-term interest rates and VIX, our intraday results are consistent with Rossi Júnior (2014) results for the weekly frequency.

[6] We do not consider Brazilian Real-U.S. dollar futures in our out-of-sample fit exercise, as we already know of its natural correlation with the spot exchange rate by the non-arbitrage relationship imposed by the Covered Interest Rate Parity.

Table 1: Brazilian real / dollar nominal spot exchange rate forecasting exercise, out-of-sample fit results.

		MSE ratio, h = 1 step ahead		
Predictors	**Method**	**freq. 1 min**	**freq. 5 min**	**freq. 15 min**
DI23, DI29	Linear Regression	**0.8146*** **	**0.7332*** **	**0.7316*** **
Ibovespa	Linear Regression	**0.7942*** **	**0.7618*** **	**0.8095*** **
Oil price	Linear Regression	**0.9816*** **	**0.9812*** **	**0.9765** **
VIX	Linear Regression	**0.9654*** **	**0.9321*** **	**0.9375*** **
Gold price	Linear Regression	**0.9717*** **	**0.9729*** **	0.9770
17 currencies	Linear Regression	**0.8250*** **	**0.7943*** **	**0.8017*** **
DI23, DI29, Ibovespa	Linear Regression	**0.7063*** **	**0.6401*** **	**0.6777*** **
DI23, DI29, Ibovespa, Oil price, VIX, Gold price	Linear Regression	**0.6815*** **	**0.6119*** **	**0.6428*** **
DI23, DI29, Ibovespa, Oil price, VIX, Gold price, 17 currencies	Linear Regression	**0.6264*** **	**0.5626*** **	**0.5843*** **
	LASSO	**0.6260*** **	**0.5312*** **	**0.5449*** **
	Ridge	**0.6203*** **	**0.5347*** **	**0.5478*** **
	ElasticNet	**0.6825*** **	**0.5310*** **	**0.5444*** **
	SVR	**0.6761*** **	**0.5608*** **	**0.5802*** **
	RF	**0.6192*** **	**0.5515*** **	**0.5515*** **
	XGB	**0.5991*** **	**0.5349*** **	**0.5425*** **

MSE ratios in relation to Randow Walk without drift. ***, ** and * indicate rejection of the null hypothesis (model and Random Walk predictions are not different), at the 1%, 5% and 10% levels respectively. DI23 and D29 correspond to Interbank Deposit interest rates futures contracts maturing in January 2023 and January 2029. The 17 currencies correspond to the nominal foreign exchange rates listed in the section 2.

It is interesting to note that the global variables - oil price, gold price and VIX - have predictive power in this out-of-sample fit exercise, but to a lesser extent, as would be expected, than the local variables (short and long term interest rate futures and Ibovespa), which absorb local news that also impact the exchange rate.

Table 1 also presents the same exercise, but with grouping of variables: first, we test as predictors the set of local variables (interest rate futures - DI23, DI29 - and Ibovespa). We note that the forecast performance improves compared to previous results. By adding the global variables - oil price, gold and VIX - we further reduced the MSE compared to random walk. And, when we add the 17 exchange rates considered in this work, we reach the best prediction performance using Linear Regression as a method, for all frequencies considered.

We also tested the performance of Machine Learning algorithms for the last specification, containing all available variables. The idea is to verify if, in the possibility of predicting the covariates and then predicting the exchange rate, Machine Learning algorithms could be useful in improving the forecasting performance. We can see that, in general, they present better results than Linear Regression, with emphasis on Extreme Gradient Boosting, which presents the best result in two of the three frequencies tested, with Elastic Net with the best performance in the 5-minute frequency [7].

[7] It is important to point out that this result does not determine a final ranking of Machine Learning algorithms for this application, as the result depends on the methodology

Figure 1: Out-of-sample fit XGB predictors importance, 1 minute frequency.

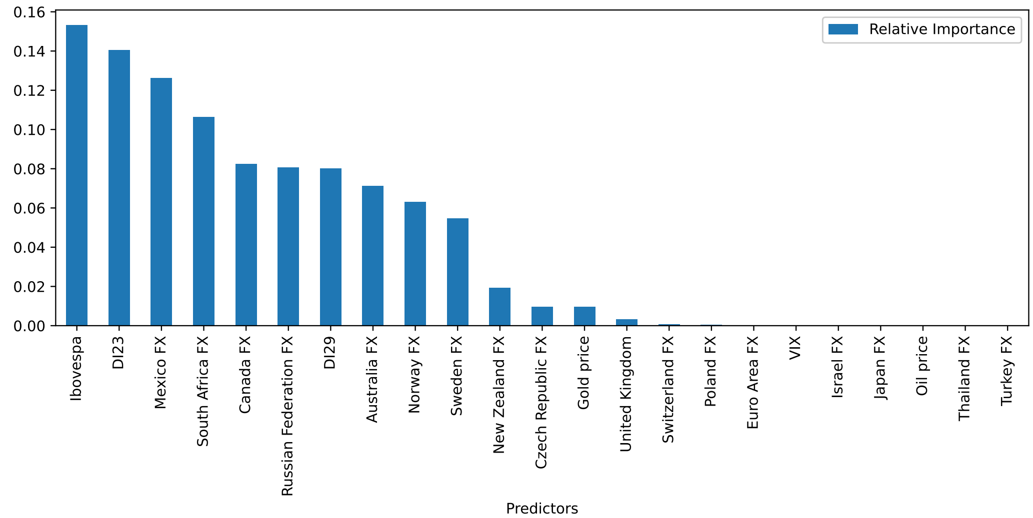
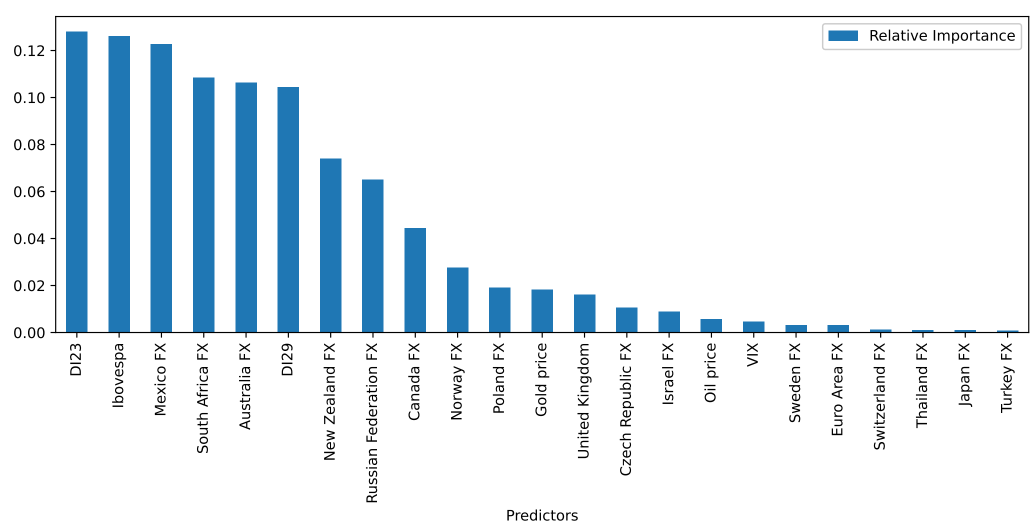


Figure 2: Out-of-sample fit XGB predictors importance, 5 minutes frequency.



Algorithms based on regression trees such as Random Forest or Extreme Gradient Boosting can automatically provide estimates of the importance of each predictor. By importance we can understand how useful each variable was in reducing the MSE during the construction of trees in the algorithm training stage [8]. We present in figures 1 to 3 the average relative importance attributed by XGB to each predictor, for the frequencies of 1, 5 and 15 minutes.

It is interesting to note that for the 1 and 5 minute frequencies, the two most important variables are Ibovespa and the short-term interest rate (DI23), with the Mexican exchange rate in third place, indicating how the similarity with the Brazilian economy is reflected in the high-frequency exchanges rates co-movements. It is also interesting to highlight, in all frequencies, the importance attributed to the South African and Russian exchange rates, the two BRICS countries available in our database.

The importance of the long-term interest rate (DI29) grows with the decrease in frequency, becoming the third most important variable in the 15-minute frequency, behind the Mexican interest rate and the Ibovespa. It is also interesting to note that the importance of global variables such as VIX, gold and oil prices increases with decreasing frequency, as well as other exchange rates. That is, with decreasing frequency, the relative importance

for determining the hyperparameters presented in the Section 3.2.

[8]For more details on how the importance of each variable is determined, see Section 10.13.1 of Friedman (2017)

of variables becomes more homogeneous.

4.2. Real out-of-sample results

In the out-of-sample forecasting exercise, in which we seek to predict the Brazilian nominal exchange rate in $t+1$ with information available only up to t, in a first attempt we cannot repeat the same results of the out-of-sample fit exercise (beat the random walk). Table 2 presents the results when using only 1 lag of the predictors, including the lag of the Brazilian exchange rate. For the exchange rates of 17 other currencies as predictors, we also apply LASSO. In this case, at a frequency of 1 minute, we even found an MSE ratio of less than 1, but without statistical significance.

However, when considering as a possible predictor the Brazilian Real-U.S. dollar futures contracts (USDBRL futures) together with the lagged value of the spot exchange rate itself, it is possible to beat Random Walk at the significance level of 5% at the frequency of 1 minute, 1 step ahead. By including local variables such as future interest rates and Ibovespa, it is possible to marginally improve the MSE ratio at this frequency.

However, as we increase the model, including oil price, VIX, gold price and exchange rates of 17 other currencies, the results worsen and statistical significance is lost, indicating that these variables are introducing more noise than signal in the real out-of-sample forecast. Machine Learning algorithms are also not able to improve the result in relation to the Linear Regression for the model with all variables for the frequency of 1 minute. However, for the frequencies of 5 and 15 minutes, some ML methods are even able to

Table 2: Brazilian real / dollar nominal spot exchange rate forecasting exercise, out-of-sample results (1 lag).

		MSE ratio, h = 1 step ahead		
Predictors	**Method**	**freq. 1 min**	**freq. 5 min**	**freq. 15 min**
USDBRL	Linear Regression	0.9998	1.0007	1.0014
DI23, DI29	Linear Regression	0.9996	1.0013	1.0027
Ibovespa	Linear Regression	0.9998	1.0006	1.0007
Oil price	Linear Regression	0.9999	0.9995	1.0012
VIX	Linear Regression	1.0000	1.0006	1.0019
Gold price	Linear Regression	1.0005	1.0011	1.0014
17 currencies	Linear Regression	1.0007	1.0049	1.0051
	LASSO	0.9995	1.0001	1.0004
USDBRL future	Linear Regression	0.9992	1.0007	1.0014
USDBRL, DI23, DI29	Linear Regression	0.9996	1.0028	1.0051
USDBRL, Ibovespa	Linear Regression	0.9997	1.0014	1.0014
USDBRL, USDBRL future	Linear Regression	**0.9969****	0.9997	1.0019
USDBRL, USDBRL future, DI23, DI29	Linear Regression	**0.9967****	1.0019	1.0055
USDBRL, USDBRL future, Ibovespa	Linear Regression	**0.9968****	1.0005	1.0019
USDBRL, USDBRL future, DI23, DI29, Ibovespa	Linear Regression	**0.9967****	1.0026	1.0053
USDBRL, USDBRL future, DI23, DI29, Ibovespa, Oil price, VIX, Gold price, 17 currencies	Linear Regression	0.9978	1.0076	1.0124
	LASSO	0.9978	1.0044	1.0036
	Ridge	0.9982	1.0006	1.0004
	ElasticNet	**0.9979***	1.0006	1.0002
	SVR	0.9978	1.0023	0.9992
	RF	1.0016	1.0055	1.0061
	XGB	1.0000	1.0132	1.0146

MSE ratios in relation to Randow Walk without drift. ***, ** and * indicate rejection of the null hypothesis (model and Random Walk predictions are not different), at the 1%, 5% and 10% levels respectively. USDBRL is the nominal Brazilian real / dollar spot exchange rate. DI23 and D29 correspond to Interbank Deposit interest rates futures contracts maturing in January 2023 and January 2029. The 17 currencies correspond to the nominal foreign exchange rates listed in the section 2.

improve the MSE ratio in relation to the Linear Regression specification, but without surpassing Random Walk in most cases or, for MSE ratio less than 1, without reaching statistical significance, as is the case with the SVR for the specification with all variables at the frequency of 15 minutes.

As an additional exercise, we consider as predictors 5 lags [9] of the variables used so far, for the frequency of 1 minute, with the results shown in the table 3. We note that the specifications that combine exchange rate lags, future exchange rate, interest futures contracts and Ibovespa present an improvement in the MSE ratio in relation to the results of Table 2, with two of them now presenting statistical significance at 1%. The other specifications show worse results compared to those with only 1 lag.

We performed another exercise, to check in how many steps ahead it is possible to beat Random Walk at a frequency of 1 minute for the successful specifications so far, with 5 lags and Linear Regression as a method. Additionally, we included a new specification that uses 5 lags for the spot and future exchange rate and only 1 lag for the other local variables (interest rates and Ibovespa). Table 4 presents the results. We conclude that within 2 minutes it is possible to beat Random Walk with statistical significance for four specifications. For three minutes ahead, some specifications even maintain an MSE ratio less than 1, but lose statistical significance.

[9] Ventura and Garcia (2012) estimate that the total effect of an order flow shock occurs in the exchange rate within 5 minutes. We did not work with order flows in this paper, but we used this result as a motivation to test 5 lags of our predictors in the exercise with a frequency of 1 minute

Table 3: Brazilian real / dollar nominal spot exchange rate forecasting exercise, out-of-sample results (5 lags).

		MSE ratio, h = 1
Predictors (5 lag of)	**Method**	**freq. 1 min**
USDBRL	Linear Regression	1.0001
DI23, DI29	Linear Regression	1.0003
Ibovespa	Linear Regression	1.0000
Oil price	Linear Regression	1.0004
VIX	Linear Regression	1.0001
Gold	Linear Regression	1.0012
17 currencies	Linear Regression	1.0052
USDBRL future	Linear Regression	0.9995
USDBRL, DI23, DI29	Linear Regression	1.0007
USDBRL, Ibovespa	Linear Regression	1.0005
USDBRL, USDBRL future	Linear Regression	**0.9948*** **
USDBRL, USDBRL future, DI23, DI29	Linear Regression	**0.9955** **
USDBRL, USDBRL future, Ibovespa	Linear Regression	**0.9952*** **
USDBRL, USDBRL future, DI23, DI29, Ibovespa	Linear Regression	**0.9959** **
USDBRL, USDBRL future, DI23, DI29, Oil price, Ibovespa, VIX, Gold, 17 currencies	Linear Regression	1.0024
	LASSO	0.9996
	Ridge	1.0017
	ElasticNet	0.9996
	SVR	1.0003
	RF	1.0017
	XGB	1.0000

MSE ratios in relation to Randow Walk without drift. ***, ** and * indicate rejection of the null hypothesis (model and Random Walk predictions are not different), at the 1%, 5% and 10% levels respectively. USDBRL is the nominal Brazilian real / dollar spot exchange rate. DI23 and D29 correspond to Interbank Deposit interest rates futures contracts maturing in January 2023 and January 2029. The 17 currencies correspond to the nominal foreign exchange rates listed in the section 2.

Figure 3: Out-of-sample fit XGB predictors importance, 15 minutes frequency.

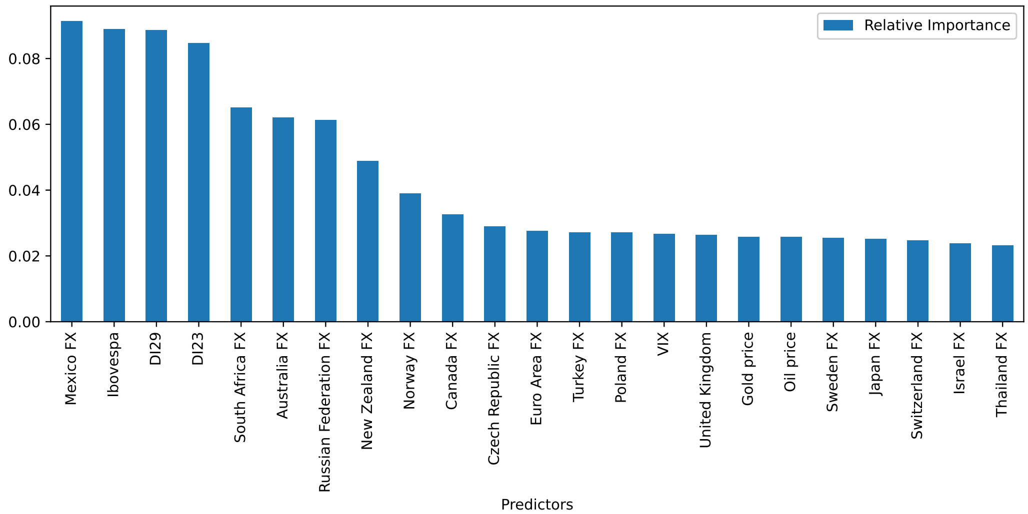


Table 4: Brazilian real / dollar nominal spot exchange rate forecasting exercise, out-of-sample results, 1 to 3 steps ahead.

		MSE ratio, frequency = 1 min		
Predictors (lags of)	**Method**	**h = 1**	**h = 2**	**h = 3**
USDBRL (5), USDBRL future (5)	Linear Regression	**0.9948*** **	**0.9973*** **	0.9987
USDBRL (5), USDBRL future (5), DI23 (5), DI29 (5)	Linear Regression	**0.9955** **	0.9983	0.9999
USDBRL (5), USDBRL future (5), Ibovespa (5)	Linear Regression	**0.9952*** **	**0.9979* **	0.9993
USDBRL (5), USDBRL future (5), DI23 (5), DI29 (5), Ibovespa (5)	Linear Regression	**0.9959** **	0.9989	1.0005
USDBRL (5), USDBRL future (5), DI23 (1), DI29 (1), Ibovespa (1)	Linear Regression	**0.9945*** **	**0.9973** **	0.9987

MSE ratios in relation to Randow Walk without drift. ***, ** and * indicate rejection of the null hypothesis (model and Random Walk predictions are not different), at the 1%, 5% and 10% levels respectively. USDBRL is the nominal Brazilian real / dollar spot exchange rate. DI23 and D29 correspond to Interbank Deposit interest rates futures contracts maturing in January 2023 and January 2029.

4.3. Robustness Check

For the real out-of-sample case, as a robustness exercise, we performed the prediction one step ahead for another five rolling window sizes for the two specifications with the best result at the frequency of 1 minute, with the results presented in table 5. We note that the MSE ratio remains at similar levels and with the same statistical significance.

Table 5: Brazilian real / dollar nominal spot exchange rate forecasting exercise, out-of-sample results, different rolling windows.

		MSE ratio, freq. 1 min, h = 1				
		Rolling Window size:				
Predictors (lags of)	**Method**	**18,75%**	**25%**	**31,25%**	**37,50%**	**43,75%**
USDBRL (5), USDBRL future (5)	Lin. Reg.	**0.9956*** **	**0.9952*** **	**0.9950*** **	**0.9952*** **	**0.9946*** **
USDBRL (5), USDBRL future (5), DI23 (1), DI29 (1), Ibovespa (1)	Lin. Reg.	**0.9953*** **	**0.9949*** **	**0.9947*** **	**0.9951*** **	**0.9944*** **

MSE ratios in relation to Randow Walk without drift. ***, ** and * indicate rejection of the null hypothesis (model and Random Walk predictions are not different), at the 1%, 5% and 10% levels respectively. USDBRL is the nominal Brazilian real / dollar spot exchange rate. DI23 and D29 correspond to Interbank Deposit interest rates futures contracts maturing in January 2023 and January 2029.

5. Conclusion

This article investigated the predictability of the Brazilian nominal exchange rate in out-of-sample fit and out-of-sample forecasting exercises using as set of predictor variables: short and long-term Brazilian interest rates, Brazilian stock market index, gold price, oil price, VIX and exchange rates of 17 other countries. In the out-of-sample fit exercise it is possible to show that, in case those variables were predictable, it would be possible to predict the exchange rate in frequencies 1, 5 and 15 minutes. However the relation between the exchange rate and those economic variables is ephemeral: the capability of prevision gets lost, for all frequencies, therefore we move from out-of-sample fit exercise to the out-of-sample forecasting. This shows how quickly the exchange rate really adjusts. Neither with 17 other exchange rates as predictors it is possible to forecast the movement of the Brazilian exchange rate 1 minute ahead better than Random Walk with statistical significance.

It is only possible to recover the prediction power, for the frequency of 1 minute, if we consider as predictor the Brazilian Real-U.S. dollar futures. This is probably due to Brazilian foreign exchange market characteristics. The futures market is much more liquid than the spot market, and according to Ventura and Garcia (2012), the exchange rate is firstly determined at the exchange rate future market and then transmitted by arbitrage to the spot market.

Machine Learning algorithms, in turn, improve the results in the out-of-sample fit prediction exercise. For the augmented models in the out-of-sample forecasting exercise, frequencies of 5 and 15 minutes, in general, also improve the MSE ratio in relation to the linear specification, however not beating the Random Walk.

References

Amat, C., Michalski, T., Stoltz, G., 2018. Fundamentals and exchange rate forecastability with simple machine learning methods. Journal of International Money and Finance 88, 1–24.

Beckmann, J., Czudaj, R., Arora, V., 2017. The relationship between oil prices and exchange rates: theory and evidence. US Energy Information Administration working paper series , 1–62.

Bekaert, G., Hoerova, M., Duca, M.L., 2013. Risk, uncertainty and monetary policy. Journal of Monetary Economics 60, 771–788.

Breiman, L., 2001. Random forests. Machine learning 45, 5–32.

Chen, T., Guestrin, C., 2016. Xgboost: A scalable tree boosting system, in: Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, pp. 785–794.

Cheung, Y.W., Chinn, M.D., Pascual, A.G., 2005. Empirical exchange rate models of the nineties: Are any fit to survive? Journal of international money and finance 24, 1150–1175.

Choudhry, T., McGroarty, F., Peng, K., Wang, S., 2012. High-frequency exchange-rate prediction with an artificial neural network. Intelligent Systems in Accounting, Finance and Management 19, 170–178.

Colombo, E., Pelagatti, M., 2020. Statistical learning and exchange rate forecasting. International Journal of Forecasting 36, 1260–1289.

Costa, A., Ferreira, P., Gaglianone, W., Guillén, O., Issler, J., Lin, Y., 2021.

Machine Learning and Oil Price Point and Density Forecasting. Technical Report. Central Bank of Brazil, Research Department.

Danielsson, J., Luo, J., Payne, R., 2012. Exchange rate determination and inter-market order flow effects. The European Journal of Finance 18, 823–840.

Diebold, F.X., Mariano, R.S., 1995. Comparing predictive accuracy. Journal of Business and Economic Statistics 13, 253–263.

Evans, M.D., Lyons, R.K., 2002. Order flow and exchange rate dynamics. Journal of political economy 110, 170–180.

Evans, M.D., Lyons, R.K., 2005. Meese-rogoff redux: Micro-based exchange-rate forecasting. American Economic Review 95, 405–414.

Felício, W.R.d.O., Rossi Júnior, J.L., 2014. Common factors and the exchange rate: results from the brazilian case. Revista Brasileira de Economia 68, 49–71.

Ferraro, D., Rogoff, K., Rossi, B., 2015. Can oil prices forecast exchange rates? an empirical analysis of the relationship between commodity prices and exchange rates. Journal of International Money and Finance 54, 116–141.

Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Annals of statistics , 1189–1232.

Friedman, J.H., 2017. The elements of statistical learning: Data mining, inference, and prediction. springer open.

Gaglianone, W.P., Marins, J.T.M., 2017. Evaluation of exchange rate point and density forecasts: an application to brazil. International Journal of Forecasting 33, 707–728.

Güriş, B., Kiran, B., 2014. The price of gold and the exchange rate: Evidence from threshold cointegration and threshold granger causality analyses for turkey. Acta Oeconomica 64, 91–101.

Hoerl, A.E., Kennard, R.W., 1970. Ridge regression: Biased estimation for nonorthogonal problems. Technometrics 12, 55–67.

Izenman, A.J., 2008. Modern Multivariate Statistical Techniques: Regression, Classification and Manifold Learning. Springer.

Jeanneau, S., Araujo, M., Amante, A., 2007. The search for liquidity in the brazilian domestic government bond market. BIS Quarterly Review, June .

Kohlscheen, E., Avalos, F.H., Schrimpf, A., 2017. When the walk is not random: commodity prices and exchange rates. International Journal of Central Banking 13, 121–158.

Manahov, V., Hudson, R., Gebka, B., 2014. Does high frequency trading affect technical analysis and market efficiency? and if so, how? Journal of International Financial Markets, Institutions and Money 28, 131–157.

Masini, R.P., Medeiros, M.C., Mendes, E.F., 2020. Machine learning advances for time series forecasting. arXiv preprint arXiv:2012.12802 .

Meese, R., Rogoff, K., 1983. Empirical exchange rate models of the seventies: Do they fit out of sample? Journal of International Economics 14, 3–24.

Moura, M.L., Lima, A.R., Mendonça, R.M., 2008. Exchange rate and fundamentals: the case of brazil. Economia Aplicada 12, 395–416.

Nair, G.K., Choudhary, N., Purohit, H., 2015. The relationship between gold prices and exchange value of us dollar in india. EMAJ: Emerging Markets Journal 5, 17–25.

Palikuca, A., Seidl, T., 2016. Predicting high frequency exchange rates using machine learning.

Rime, D., Sarno, L., Sojli, E., 2010. Exchange rate forecasting, order flow and macroeconomic information. Journal of International Economics 80, 72–88.

Rossi, B., 2013. Exchange rate predictability. Journal of economic literature 51, 1063–1119.

Rossi Júnior, J.L., 2014. The usefulness of financial variables in predicting exchange rate movements. Insper Working Paper WPE: 332/2014 .

Tabak, B.M., 2006. The dynamic relationship between stock prices and exchange rates: Evidence for brazil. International Journal of Theoretical and Applied Finance 9, 1377–1396.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological) 58, 267–288.

Vartanian, P.R., Citro, S.G., Scarano, P.R., et al., 2021. Determinants of spot and forward interest rates in brazil in an international liquidity scenario: An econometric analysis for the period 2007-2019. International Journal of Applied Economics, Finance and Accounting 10, 19–31.

Ventura, A., Garcia, M., 2012. Mercados futuro e à vista de câmbio no brasil: o rabo abana o cachorro. Revista Brasileira de Economia 66, 21–48.

Zhang, Y., Hamori, S., 2020. The predictability of the exchange rate when combining machine learning and fundamental models. Journal of Risk and Financial Management 13, 48.

Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. Journal of the royal statistical society: series B (statistical methodology) 67, 301–320.