Using Machine Learning to Set Exchange Rates for Medium Term Contracts

Eduardo Lorie
Georgia Institute of Technology
edlorie@gatech.edu

Matthew Robinson
Georgia Institute of Technology
mrobinson72@gatech.edu

1. INTRODUCTION

International business involves considerable risk. This is especially true for manufacturers, who often make production decisions well in advance of delivery dates. Consider the position of an airline that is preparing to expand its fleet. Such an airline might place an order for a passenger jet that will not be completed for a year or more, with payment due upon delivery of the jet. If the buyer is American and the manufacturer is also American, this is not a problem. All of the manufacturers production costs will be valued in US dollars, and the customer will pay in US dollars.

In contrast, consider the situation if the buyer is Canadian and the manufacturer is American. Since the manufacturer is American, the Canadian firm must pay for the jet in US dollars. Suppose the value of the US dollar appreciates 10% relative to the Canadian dollar while the American firm is filling the order. Then, when the Canadian firm pays for the jet, it is 10% more expensive in terms of Canadian dollars than when it ordered it. If the movement were in the opposite direction, the manufacturer would receive 10% fewer US dollars for the jet than it expected when it was ordered. Clearly, both firms would be interested in predicting adverse currency movements in advance, and may be interested in writing expected future currency valuations into contracts.

In order to resolve this dilemma, we estimate several models that predict the percent change in the exchange rate after one year, using only information that is currently available. The predictors for these models include the current exchange rate, the interest rate, inflation rate and government bond yield for both countries, and several variables related to international trade and financial flows. The remainder of this section will provide an overview of relevant economic theory, a summary of past statistical research on exchange rates and an overview of the data set. In the following section, we perform exploratory data analysis to develop a basic understanding of the trends in the data set. Next, we develop several models for predicting exchange rates and compare their effectiveness. Finally, we conclude by suggesting the most effective means to predict exchange rates and possible

directions for future research.

1.1 Related Work

Research on equilibrium in foreign exchange markets is a well developed component of classical economic theory. Classical models rest on two fundamental ideas. The first is that foreign exchange markets achieve equilibrium when the rate of return on deposits is the same across all currencies. The idea that investors will be indifferent between bank deposits denominated in different currencies is known as interest rate parity. The second is that the price of goods will be the same when valued in different currencies. The notion that a basket of goods should cost the same in all currencies is known as purchasing power parity.

Both of these concepts rely on the same basic premise, which is that price differentials create arbitrage opportunities. Interest rate and purchasing power parity hold that price differentials are self-correcting because, as investors move to take advantage of arbitrage opportunities, they push the market back toward equilibrium. To demonstrate this idea, suppose that a laptop costs $500 in the United States and the Euro equivalent of $550 in Germany. Someone in Germany could take advantage of this fact by buying cheap laptops in the US and selling them in Germany. This would create more demand for dollars, which are required to buy the laptops in the US, and push up US price levels. This trend would continue until price levels are high enough that laptops cost the same in the US as in Germany. When laptop prices are the same in both countries, the market is in equilibrium since there is no longer an opportunity for arbitrage.

The following equations summarize the relationship between exchange rates and interest rates under interest rate parity and the relationship between exchange rates and price levels under purchasing power parity, all else held equal. In these equations, e_t represents the exchange rate at time t in terms of the foreign currency, i is the real interest rate and π is the inflation rate.

$$e_{t+1} = e_t \left(\frac{1 + i_d}{1 + i_f} \right) \tag{1}$$

$$\frac{e_{t+1} - e_t}{e_t} = \frac{1 + \pi_d}{1 + \pi f} - 1 \tag{2}$$

Obstfeld and Taylor show that arbitrage opportunities between foreign and domestic assets are near zero under floating currency regimes, which suggests that exchange rates behave as predicted by interest rate parity. Likewise, Frankel and Rose provide evidence that exchange rates converge to

levels predicted by purchasing power parity in the long run. The relevant time-frame for the *long run* in this context is about four years. Under shorter time horizons, purchasing power parity performs considerably less well. Likewise, difficulty in predicting the relative performance of financial assets in different countries makes it difficult to predict interest rates using interest rate parity. In fact, Meese and Rogoff famously demonstrated that a random walk outperforms structural models for the exchange rate over a one to twelve month window. For this reason, the random walk is currently used as a benchmark for assessing the quality of models for predicting exchange rates.

Apart from random walks other approaches to have been studied to accurately determine future changes in exchange rates. Hsiech showed that there is little linear dependence in exchange-rate changes but there is a strong nonlinear dependence. On the other hand, Meese and Rose showed that nonlinear estimators to do not significantly improve popular models for predicting exchange rates. All these observations happened with the use of macroeconomic variables such as interest rates and inflation rates. Lyons and Evans showed that models become much more accurate with the introduction of microeconomics. They augmented previously used linear models with a new variable, order flow, which is the net of buyer and seller initiated orders. The introduction of this variable resulted in accurately predicting 60% of daily changes in Deutsche mark/US dollar rate.

Another popular model in determining future exchange rates is the artificial neural network. Panda and Narasimhan showed that Neural networks outperform both random walk and linear autoregressive models when predicting weekly changes in Indian rupee/US dollar exchange rates.

1.2 Data

This paper restricts its attention to changes in the exchange rate between the United States and its five biggest trading partners, excluding China. These trading partners are Europe, Canada, Mexico, Japan and South Korea. Europe in this context refers to the 19 countries that belong to the eurozone. Since they have a single monetary policy, they will be treated as a single unit for the purposes of this paper. China is excluded because it fixes its exchange rate. Because of this, its exchange rate is determined by government policy rather than economic variables.

The data used in this paper was collected entirely by central banks and compiled by Quandl, an economic data repository. The primary variables of interest in the data set are the inflation rate, interest rate and the yield on one year government bonds. In this data set, the inflation rate measures the annualized percent change in the consumer price index for each country. The interest rate is the interbank lending rate set by the central bank for the currency in question. For the United States, this is the federal funds rate. Other variables in the data set include an area's balance of trade, current account, foreign exchange reserves and other factors related to international financial flows. These variables were measured both for the United States and each currency pairing. Each variable is measured monthly, with the exception of GDP growth, current account balance and foreign direct investment. These variables are measured quarterly. In order to align this data with the rest of the data set, values for the intervening two months between each quarter were interpolated using cubic splines.

Figure 1: Percent Change over Time
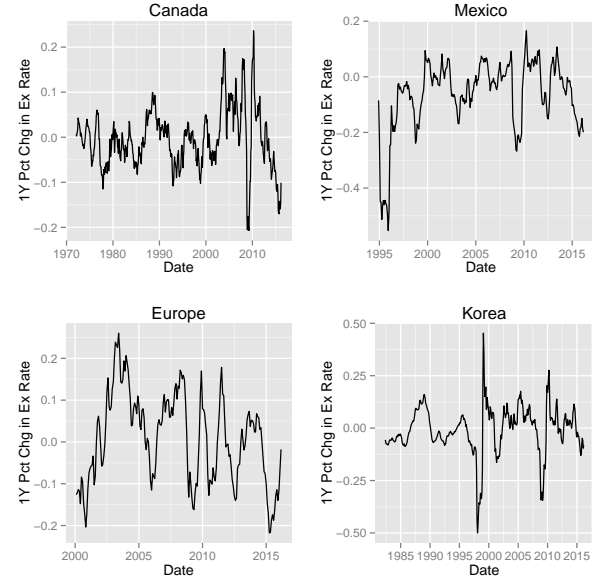
2. EXPLORATORY DATA ANALYSIS

2.1 Stationarity

2.2 Tests for Autocorrelation

Time series data often exhibits autocorrelation. Autocorrelation refers to correlation between values of a variable at different points in time. In order to assess the degree of autocorrelation in the exchange rate data set, we fit a simplified linear model that consisted of the current exchange rate, along with the interest rate, inflation rate and government bond yield for both countries. After fitting the model, we computed the Durbin-Watson test statistic with one to twelve month lags. Figure 2 depicts the value of the correlation coefficient for each country as a function of time.

The test statistics show that the data set exhibits strong autocorrelation for several months, and that the severity of the autocorrelation decays over time. For each country, the correlation coefficient for a one month lag is in excess of 0.80, and is as high as 0.95 for Japan. This value drops below 0.50 for every country by month four, with the exception of Japan. The persistence of autocorrelation in the case of Japan may be due to its active currency manipulation policy. For each country, the correlation coefficient remains statistically significant for at least five months. The presence of autocorrelation in the data set argues for including lagged variables in the model, through at least several months.

3. METHODS

Constructing a model to predict exchanges requires the consideration of three important factors: which variables to include in the model, candidate methods for estimating the model and the criteria for evaluating each method. The variable selection process involves deciding which economic factors to include in the model, as well as how many months of lagged observations to include for each factor. Candidate methods were chosen by considering several statistical tech-

Figure 2: Severity of Autocorrelation
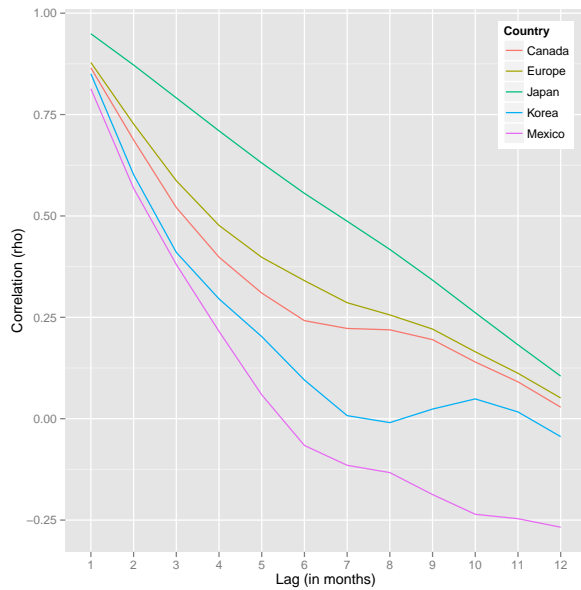
niques with properties that made them appealing for the exchange rate data set. Finally, a method for evaluating the candidate solutions that accounts for autocorrelation in the data set is developed.

3.1 Variable Selection

Selecting which variables to include in the model requires deciding which economic factors are relevant and for how many months prior to the current observation. To recap, the variables in the data set include the current exchange, interest, inflation and GDP growth rates, as well as several variables related to international financial flows. There is a strong theoretical justification for including each of these variables in the model. Nevertheless, overfitting is potentially problematic due to the small size of the data set, especially after the inclusion of lagged variables. The proper lag was selected by comparing the performance of a linear regression model across a variety of lag values. For each lag value, the variables were chosen by forward stepwise variable selection, using AIC as the selection criteria. Specifying the model in this manner allows us to balance the trade off between training error and model complexity.

3.2 Candidate Methods

After selecting a suitable subset of variables to include in the model, we fit the model using several statistical techniques. These techniques include standard linear regression, ridge regression, LASSO regression and principal components regression. Ridge regression and LASSO regression both include a regularization constant that gives preferences to models with smaller parameters in terms of absolute value. These models are appealing for the exchange rate data set because autocorrelation and the inclusion of lagged variables means the predictors in the model are highly correlated, meaning multicollinearity is a potential problem. Including a regularization constant helps prevent overfitting to the training data. For both of these models, the regularization constant is chosen using generalized cross validation.

Principal component regression is potentially useful due to the large number of predictors in the model and the high correlation between those predictors. The forward stepwise variable technique described above results in models with 10-20 predictors, which is relatively complex for a data set that has on the order of hundreds of observations. Principal component regression simplifies the model by considering a low rank approximation of the data set that consists of the k most important principal components. The number of components to include in the principal component regression was also chosen using cross validation.

3.3 Model Evaluation

Evaluating the effectiveness of the candidate solution required splitting the data into training and test sets. In all cases described in this paper, the training set consists of 80% of the available data, with the remaining allocated to the test set. In order to effectively form the training and test sets, two aspects of the data required consideration: autocorrelation and the size of the data set.

As discussed above, the data set used in this paper exhibits autocorrelation that remains significant for at least several months. Since the economic variables in the data set are heavily correlated over time, this means that the data set cannot be split between training and test sets simply by separating the data randomly. To illustrate the nature of the problem, suppose that an observation from May 2012 is randomly assigned to the training set and an observation from June 2012 is randomly assigned to the test set. Since the May and June observations are correlated, a model that overfits the training data might perform artificially well on the test set because the training set contains observations that are similar to observations in the test set. In order to circumvent this problem, the test set was chosen by selecting a continuous block of time with a random start date. Additionally, six months of data before and after the test block was discarded. This data was used in neither the training nor the test set in order to ensure that the observations in the test set were as independent as possible from the observations in the training set.

Since this paper considers monthly data, the number of observations in each time series is on the order of hundreds, ranging from 121 for Mexico to 459 for Japan. Due to the relatively low number of observations in the data set, a true appraisal of the effectiveness of each method requires an evaluation of its performance over a large number of random partitions. In order to accomplish this, we randomly selected 100 start points for each data set and computed the training error for each method at each start point. The same 100 start points were used for each method. This allows us to test the hypothesis that one method outperforms another by conducting pairwise two sample t-tests. Using many randomly selected start dates has the added advantage of providing more variability for the test sets. Since continuous blocks of time are correlated, the peculiarities of a particular block of time may provide a distorted view of the effectiveness of a given method if only one block of time is used. As a result, using a continuous block of time with many randomly selected start points is a reasonable method for splitting the data between training and test sets.

The mean squared testing error over the 100 random partitions of the data set is used as a basis for comparing models. For brevity, we will refer to this value as the mean testing

Figure 3: Model Performance for Various Lags

Figure 4: Model Performance for Various Lags
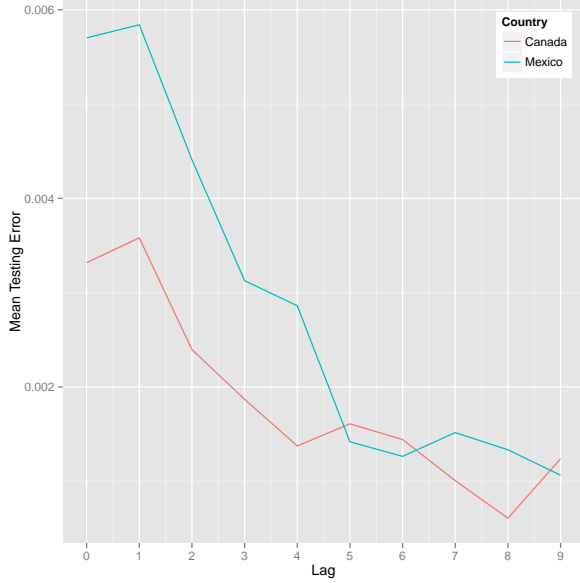
error in this paper. In all cases, two sample t-tests are used to determine whether the difference between two models is statistically significant. Comparison against a random walk, which predicts no change in the exchange rate after one year, is used as a benchmark for determining whether or not the model's predictions are useful.

4. RESULTS

4.1 Determining the Optimal Lag

The lag was determined by considering the performance of a standard linear regression model with a variety of lags. Since we want the simplest model specification that provides accurate predictions, we choose the point at which including an additional month of lag stops improving the performance of the model. This is accomplished by plotting the mean training error of the model as a function of the lag. These plots appear in Figures 3 and 4.

The optimal lag varies by country. Noticeable improvements in the performance of the model cease after month four for Canada and month five for Mexico. The results justify including lagged variables out to nine months for Europe, Japan and Korea. In the case of Korea, the decision to include a lengthy lag on the variables is less clear cut, since the model also performs well with shorter lags. However, the mean testing error decreases by a factor of two after the ninth month is included, making the increased complexity of the model worthwhile.

Including lagged variables results in a large number of possible predictors for each model. Using the lags chosen above, there are 42 possible predictors for the Canada model, 49 for the Mexico model and 77 each for the Europe, Japan and Korea models. The variables used to predict the percent change in the exchange rate were chosen by forward stepwise variable selection, as described in the methods section. These variables and their lags appear in the appendix. Of note is that at least 9 of the 12 variables in the data set appear in each of the final models chosen by forward stepwise

Figure 5: Mean Testing Error

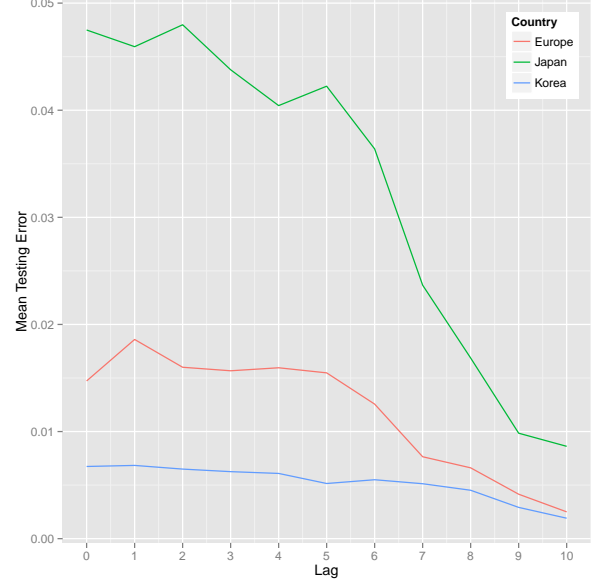| Country	Standard	Ridge	Lasso	PCR	Naive
Canada	0.001374	0.001395	0.001409	0.001532	0.00636
Europe	0.004150	0.004389	0.004343	0.003869	0.01159
Mexico	0.001418	0.001499	0.001735	0.001483	0.01146
Japan	0.009847	0.009840	0.009622	0.009602	0.02058
Korea	0.002919	0.002931	0.003109	0.002519	0.01288

variable selection. In addition, each variable appeared in at least two of the models. This indicates that each variable is useful as a predictor of exchange rates, as economic theory suggests they should be.

4.2 Method Comparison

The mean testing errors for the 100 random partitions of the data set appear in Figure 5 for each country and estimation technique. The naive estimate is a random walk and always predicts no change in the exchange rate after one year. This estimate is used as a baseline for assessing the usefulness of the statistical models. Each of the estimation techniques outperformed the naive estimation. The improvement in predictive power ranged from a factor of two for Japan to a factor of ten for Mexico. For each method, we tested the hypothesis that mean testing error for the model is less than the mean testing error for the random walk using a pairwise two sample t-test. The results were statistically significant at well below the 1% significance level.

A noteworthy observation from the results table is that the mean testing error for each model is robust to the technique used to estimate the model. For each country, we tested the hypothesis that the mean testing errors for the ridge, LASSO and principal component regressions were different than the mean testing error for the standard linear model. Although the differences were statistically significant in some cases, the magnitude of the difference was very small in every case. In practice, this means that each of the models will produce similar results. Since using more sophisticated regression

techniques does not improve the performance of the model, these results support estimating the model using a standard linear regression. In fact, the primary difficulty in producing a good model is not selecting the best estimation technique, but rather choosing the correct predictors. While the results were robust to the estimation technique, they were highly sensitive to the variables included in the model, particularly the months of lag. This makes preliminary work to determine a close to optimal lag value vital to producing a useful model for predicting exchange rates.

5. FINDINGS

REFERENCES

Frankel, J., and A. Rose. 1996. "A panel project on purchasing power parity: mean reversion within and between countries." *Journal of International Economics* 40: 209-224.

Hsieh, D. 1989 "Testing for Nonlinear Dependence in Daily Foreign Exchange Rate Changes." *Journal of Business* 62: 329-368.

Krugman, P., M. Obstfeld and M.Melitz. 2014. *International Economics: Theory and Policy.* 10th ed. Upper Saddle River, New Jersey: Pearson Education.

Juselius, K. 1995. "Do purchasing power parity and uncovered interest rate parity hold in the long run? An example of likelihood inference in a multivariate time-series model.". *Journal of Econometrics* 69: 211-240.

Lyons, K.R. and M.D.D. Evans. 2002 "Order Flow and Exchange Rate Dynamics." *Journal of Political Economy* 110: 170-180

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

Meese, R.A. and A.K. Rose. 1990 "Nonlinear, Nonparametric, Nonessential Exchange Rate Estimation." *The American Economic Review* 80: 192-196

Obstfeld, M., and A. Taylor. 2003. "Globalization and capital markets." *Globalization in historical perspective* University of Chicago Press: 121-188.

Panda, C. and V. Narasimhan 2007 "Forecasting exchange rate better with artificial neural network." *Journal of Policy Modeling* 29: 227-236

Wooldridge, J. 2009. *Introductory Econometrics.* 4th Ed. Mason, Ohio: South-Western Cengage Learning.

Figure 6: Variables and Lags – Canada

Variable	Lags
Exchange Rate	0,2,4
Domestic Bond Yield	1,2,3,4
Foreign Bond Yield	0,1,3,4
Domestic Interest Rate	0,3,4
Foreign Interest Rate	0,2
Domestic Inflation Rate	2,4
Foreign Inflation Rate	0,3,4
Domestic GDP Growth	0
Foreign GDP Growth	0
Domestic Balance of Trade	0
Foreign Balance of Trade	None
Foreign Exchange Reserves	0

Figure 7: Variables and Lags – Europe

Variable	Lags
Exchange Rate	0,9
Domestic Bond Yield	0,5,7
Foreign Bond Yield	1,4,9
Domestic Interest Rate	3, 6,8,9
Foreign Interest Rate	5,6
Domestic Inflation Rate	0, 6
Foreign Inflation Rate	4,5
Domestic GDP Growth	0
Foreign GDP Growth	0
Domestic Balance of Trade	0
Foreign Balance of Trade	0
Foreign Exchange Reserves	None

Figure 8: Variables and Lags – Mexico

Variable	Lags
Exchange Rate	0,2
Domestic Bond Yield	5
Foreign Bond Yield	3,4
Domestic Interest Rate	0,1,5
Foreign Interest Rate	3,5
Domestic Inflation Rate	0,3,5
Foreign Inflation Rate	0,2
Domestic GDP Growth	0
Foreign GDP Growth	0
Domestic Balance of Trade	None
Foreign Balance of Trade	None
Foreign Exchange Reserves	None

Figure 9: Variables and Lags – Japan

Variable	Lags
Exchange Rate	0,4,8,9
Domestic Bond Yield	0,4
Foreign Bond Yield	None
Domestic Interest Rate	3,6,9
Foreign Interest Rate	2,7,9
Domestic Inflation Rate	0
Foreign Inflation Rate	0,8
Domestic GDP Growth	None
Foreign GDP Growth	0
Domestic Balance of Trade	0
Foreign Balance of Trade	0
Foreign Exchange Reserves	0

Figure 10: Variables and Lags – Korea

Variable	Lags
Exchange Rate	0,6,9
Domestic Bond Yield	0,5
Foreign Bond Yield	8
Domestic Interest Rate	2,6,9
Foreign Interest Rate	7,9
Domestic Inflation Rate	1,3,5,6,9
Foreign Inflation Rate	0,9
Domestic GDP Growth	None
Foreign GDP Growth	None
Domestic Balance of Trade	0
Foreign Balance of Trade	None
Foreign Exchange Reserves	0