Article

Timing Foreign Exchange Markets

Samuel W. Malone [1], Robert B. Gramacy [2,*] and Enrique ter Horst [3]

[1] Moody's Analytics, 121 N Walnut St., West Chester, PA 19380, USA; Samuel.Malone@moodys.com
[2] Booth School of Business, The University of Chicago, 5807 S, Woodlawn Ave, Chicago, IL 60637, USA
[3] Colegio de Estudios Superiores de Administración, Bogotá, Colombia and Instituto de Estudios Superiores de Administración, Caracas 1010, Venezuela; enriqueterhorst@gmail.com
* Correspondence: rbgramacy@chicagobooth.edu; Tel.: +1-773-702-0739

Abstract: To improve short-horizon exchange rate forecasts, we employ foreign exchange market risk factors as fundamentals, and Bayesian treed Gaussian process (BTGP) models to handle non-linear, time-varying relationships between these fundamentals and exchange rates. Forecasts from the BTGP model conditional on the carry and dollar factors dominate random walk forecasts on accuracy and economic criteria in the Meese-Rogoff setting. Superior market timing ability for large moves, more than directional accuracy, drives the BTGP's success. We explain how, through a model averaging Monte Carlo scheme, the BTGP is able to simultaneously exploit smoothness and rough breaks in between-variable dynamics. Either feature in isolation is unable to consistently outperform benchmarks throughout the full span of time in our forecasting exercises. Trading strategies based on *ex ante* BTGP forecasts deliver the highest out-of-sample risk-adjusted returns for the median currency, as well as for both predictable, traded risk factors.

Keywords: foreign exchange; speculation; Bayesian treed Gaussian process; Anatolyev-Gerko statistic; Giacomini-White statistic

JEL: F31, G15, G17

1. Introduction

We take a novel approach to short-horizon exchange rate forecasting by using priced, predictable, and traded foreign exchange rate factors as fundamentals. Specifically, we use the carry and dollar factors proposed by [1,2]. The dollar factor is the average excess return of a portfolio that goes long in all foreign currencies versus the US dollar, whereas the carry factor captures the return of a portfolio that funds long positions in high interest rate currencies with short positions in low interest rate currencies. Linear time series factor models with perfect foresight of these factors achieve RMSEs that are between 10%–70% lower than those of the random walk benchmark, depending on the currency [2].

This result is a notable success in relation to literature in the Meese and Rogoff tradition [3], but such success raises new and important questions. Do *ex post* factor model forecasts based on perfect foresight of the carry and dollar factors have economic value? If so, what are the main drivers of this value? Can *ex ante* forecasts of individual currency excess returns, based on combining rolling factor model estimates and factor forecasts, still beat the random walk benchmark relative to economic criteria? Can a model that captures potential non-stationarity and non-linearity in the relationship between factors and exchange rates improve the economic value of factor forecasts? These are the questions we address in this paper.

We focus on the economic value of short-horizon, out-of-sample forecasts of individual exchange rate excess returns, as measured by the directional accuracy, market timing ability, profitability, and Sharpe ratios of simple directional strategies in individual exchange rates versus the US dollar. To measure market timing ability, we employ a profitability-based statistic [4], which is positive when a given directional strategy achieves superior profitability compared to a baseline strategy that randomly predicts a positive excess currency return with the same probability as the strategy being tested. Intuitively, the statistic captures the ability of a directional strategy to make money by guessing the direction correctly when it matters, which corresponds to situations when the realized return magnitudes are high. By using the covered interest rate parity (CIP) relationship, which implies that realized carry trade excess returns are equal to the interest rate differential minus the realized depreciation of the (log) spot exchange rate, we map the problem of forecasting the spot exchange rate return into the problem of "timing" carry trades in individual currencies.

Our strategies, similar to those studied recently in [5,6], fund a long position in a given currency by a short position in the US dollar, if the expected next-month excess return of doing so is positive, and take the opposite positions otherwise. As demonstrated in [7], forecast breakdowns can arise due to the instability of the relationship between exchange rates and fundamentals over time. To deal with this problem, we contribute by employing a non-parametric Bayesian treed Gaussian process (BTGP) model [8] to form conditional forecasts, in addition to (Bayesian) linear models. The BTGP relaxes stationarity and heteroskedasticity assumptions underlying the relationship between exchange rate returns and fundamentals over time, and allows this relationship to exhibit a substantial degree of non-linearity. The BTGP is, in some sense, simply a representative from a wide class of modern flexible Bayesian nonparametric methods, many others of which may suit our motivating application well, and which we believe are under utilized in this context. However two features make the BTGP particularly attractive: (1) open source software; (2) an efficient model averaging Monte Carlo scheme, over sharp breaks (trees) and smooth transitions (GPs), allowing model fits and subsequent forecasts to strike a robust balance between dynamic shifts and long term trends.[1]

We make a distinction between *ex post* forecasting exercises, in which agents have perfect foresight of next-month fundamentals, following the Meese and Rogoff tradition, and *ex ante* forecasting exercises, in which they do not. Our exercises are designed sequentially, so that the difference between success in *ex post* and *ex ante* forecasting depends solely on the quality of the factor forecasts themselves. To form factor forecasts, we exploit predictability results [10] which demonstrate the average developed country forward premium *vs.* the US dollar is a successful predictor of the dollar factor, combined with the fact that both the carry and dollar factors exhibit some moderate degree of persistence.

We compare the accuracy and economic value of factor forecasts for four models: the random walk with drift, the linear model, the BTGP, and a factor model, e.g., [11]. In particular we apply the procedure in that paper directly to the carry and dollar factors, rather than taking exchange rates as factors as in the original paper. As portfolios, the carry and dollar factors arguably do a better job of capturing distinct state variables that drive stochastic discount factors of investors in foreign exchange markets than do individual exchange rate returns, while diversifying away the sort of idiosyncratic risk known to degrade forecasting performance [1,10].

Our main results are as follows. We find that perfect foresight of the carry and dollar factors has significant economic value, both in absolute terms, and relative to the random walk benchmark and

---

[1] On a computational note, all code for this paper was written using the R language, with the BTGP model implemented using the `tgp` package [9]. To speed up back-testing exercises, we made use of the `snow` (now called `parallel`) and `snowfall` libraries in R. In our simple experimental setup, forecast generation for different currencies, as well as different time periods for a given currency, are independent. We experimented with parallelizing forecast generation over time periods, and found modest speed up gains on the order of 3.2 times using the socket method with four cores on a Lenovo laptop.

to models conditional on perfect foresight of macroeconomic fundamentals. The average annualized returns and Sharpe ratios of our strategies, across all 43 countries in our worldwide sample, are 22% and 2.1, respectively, compared to 2.8% and 0.37 for the zero-drift random walk benchmark, for *ex post* forecasting exercises. Linear and BTGP conditional models perform similarly in this setup, and both dominate the RW benchmark, as well as the random walk with drift (RWWD) benchmark, with respect to both directional accuracy and market timing ability. Upon considering subcases of the variables in a recently proposed factor model [2], we find that perfect foresight of both the carry and dollar factor confers added economic value relative to the benchmarks, but that the information in the dollar factor is substantially more valuable than the information contained in the carry factor.

To form *ex ante* exchange rate return forecasts, we first form *ex ante* forecasts of the carry and dollar factors. We find that traded factor forecasts formed using the [11] (GM) model, as well as conditional linear and BTGP competitors, are able to beat the random walk benchmarks for factor forecasting with respect to economic criteria, but not with respect to root mean squared error. The Sharpe ratios we obtain for factor timing strategies based on our forecasts, however, are consistent with recent results of [10,12] for the carry and dollar portfolio timing strategies. We find the novel result that the carry factor is much easier to time than the dollar factor: the BTGP obtains a Sharpe ratio of 1.23 for the carry factor timing strategy, *versus* only 0.38 for the dollar factor, for example, before transaction costs. All conditional models obtain similar RMSEs out-of-sample, and we use GM model factor forecasts in forming *ex ante* forecasts for individual currencies.

Our use of *ex ante* factor forecasts, combined with a rolling BTGP factor model for individual currencies, bears fruit in *ex ante* individual exchange rate forecasting exercises. *Ex ante* currency forecasts based on factor forecasts modestly outperform the random walk benchmark for the median currency with respect to directional accuracy, market timing ability, profitability, and Sharpe ratios. When we diagnose the source of BTGP outperformance by running subcases of the full BTGP model, we find that the modeling of potential non-stationarity via trees, and the modeling of potential non-linearity via Gaussian processes, are both useful in beating the benchmark. Like others, e.g., [13], we find no "magic bullet" solution to beating the benchmark, only a set of value-added innovations, which consist in our case of high quality carry and dollar fundamentals plus the BTGP model, capable of achieving superior performance when used together via Bayesian model averaging.

Finally, given our focus on out-of-sample performance with respect to multiple criteria, and our use of a diverse set of forecasting models estimated using rolling windows, we employ Giacomini–White [14] (GW) tests of unconditional and conditional predictive ability by country to evaluate our models in terms of error-based and profitability-based loss functions. Results from these tests largely confirm our conclusions from examining the performance of strategies on average and for the median currency.

The rest of this paper is organized as follows. Section 2 briefly reviews the relevant literature. Section 3 describes the dataset and fixes notation. Section 4 describes the forecasting approach and the BTGP model. Section 5 describes our trading strategy and the statistics we use to judge forecast performance. Section 6 presents empirical results. Section 7 concludes.

2. Related Literature

A substantial body of literature has focused on overturning the original [3] result that random walk models beat models conditional on perfect foresight of next-month macroeconomic fundamentals, with respect to RMSE and other traditional measures of forecasting accuracy. Such attempts have, by and large, met with little success. One class of approaches involves using more sophisticated model selection criteria (see e.g., [7,15,16]); others expand the range of exchange rates studied to include smaller country cross rates with the dollar or another major currency ([17–19]). A more detailed discussion of these papers, none of which manages to beat the random walk at the monthly frequency of greatest interest to the literature, can be found in [6]. Earlier work on exchange rate forecasting includes linear models with time-varying features ([20–22]), non-parametric

kernel regression models ([23,24]), and Markov-switching models ([25,26]). More recently, the use of neural network models to forecast exchange rates has met with some limited success [16]. While the application of neural networks to forecasting daily exchange rates out of sample [27–29] seems to have outperformed the random walk, an application of a neural network model with monetary fundamentals at the 1-month and longer horizons finds that the neural network performs no better than a random walk at these longer horizons [30]. Rogoff himself has noted that the inability of structural models to explain the movements of G3 exchange rates in particular remains a robust stylized fact [31].

Models that do beat the random walk with drift at longer time horizons, moreover, such as a segmented trends model [25] on quarterly data, and recent factor model approaches [11,32] at the semi-annual and longer time horizons, are still dominated by the zero-drift random walk at the monthly horizon. As noted in a recent survey of the literature on exchange rate predictability [33], the random walk without drift is the most difficult benchmark to beat, and the ability of alternative models to achieve superior forecasts depends significantly upon the choice of predictors, forecast horizon, sample period, and forecast evaluation method used, not to mention the model itself.

In this context, the ability to achieve lower RMSEs than the random walk for most currencies [2], using a linear time series factor model conditional on perfect foresight of the carry and dollar factors, represents a substantial new benchmark in the international finance literature. Importantly, in our work as in his, the carry and dollar factors used in factor models do not include the currency whose excess returns are being predicted. It has been cautioned [34] that success with respect to measures of forecasting accuracy does not guarantee success with respect to economic criteria. In this paper, we take this admonition seriously, and focus on a detailed examination of the economic value of forecasts formed using the carry and dollar factors. This focus, including our diagnosis of the drivers of economic value of forecasts gleaned from directional accuracy and market timing statistics, as well as the use of the BTGP model to take into account non-stationarity and non-linearity in the process linking these new fundamentals and exchange rates, all differentiate our paper from [2].

As the first study to examine the economic implications of individual exchange rate predictability based on carry and dollar factors, we measure economic value using carry trade timing strategies, e.g., [5,6]. In these papers and our own, the benchmark (random walk) strategy forecasts the log spot exchange rate return to be zero, then invokes CIP to set the expected excess return equal to the forward premium. We retain this benchmark, but differ from [5] by including the RWWD, linear, and BTGP models for forming forecasts of log currency returns, as well as paying careful attention to the sources of the economic value derived from forecasts. We differ from [6], which examines speculation in black markets for currency, and employ a set of macroeconomic fundamentals appropriate for that setting, by focusing on large, liquid foreign exchange markets. In this paper, we report results for the cross-section of out-of-sample return time series generated across currencies in a worldwide sample, similar to those studied by [1,2,10,35]. We do not study the implications of our work for improving optimal portfolio selection in foreign exchange markets [36], but leave that problem for future research.

One of the main innovations of our paper centers on the choice of fundamentals for our conditional models: the carry and dollar risk factors. Recent works [1,2] show that the carry and dollar factors are priced risk factors in foreign exchange markets, which provide a risk-based explanation for the profitability of the classic carry trade. In the carry trade, investors fund long positions in high interest rate currencies with short positions in low interest rate currencies. Such investors are, essentially, earning positive mean excess returns by loading up on exposure to the carry risk factor [1], whereas [2] demonstrates that investors who employ the same strategy, but with the dollar as their base currency, earn positive mean excess returns by loading up on exposure to the dollar risk factor.

It is *a priori* reasonable to expect that systematic risk factors in foreign exchange markets should be better candidates for forecasting individual currencies than largely idiosyncratic macroeconomic

fundamentals, for reasons familiar from the no-arbitrage logic of asset pricing theory [37]. However, suitable candidates for risk factors have only been discovered recently to serve this purpose.[2]

Our approach, and findings, echo recent factor-based approachs to bilateral exchange rate forecasting [11,32]. Unlike those papers, we use demonstrably superior carry and dollar factors to evaluate the economic value of individual exchange rate excess return forecasts at the monthly frequency, focus on the economic value of forecasts, and employ the flexible BTGP model for forecasting, which we show generally delivers superior results to the linear model.

3. Description of the Data

Our dataset is similar to those used in several recent works, e.g., [1,10]. We study 43 countries, including 14 developed and 28 emerging market countries, in addition to the Euro, during the period from November 1983 to December 2010. These currencies, which we refer to henceforth as "All" currencies, are: of Australia, Austria, Belgium, Brazil, Bulgaria, Canada, Croatia, Czech Republic, Denmark, Egypt, euro area, France, Germany, Hong Kong, Hungary, Iceland, India, Indonesia, Ireland, Israel, Italy, Japan, Kuwait, Mexico, Netherlands, New Zealand, Norway, Philippines, Poland, Portugal, Russia, Saudi Arabia, Singapore, Slovakia, South Africa, South Korea, Spain, Sweden, Switzerland, Taiwan, Thailand, Ukraine, and the United Kingdom.

Within the set of our 43 currency episodes, and following [1], we delete some observations from our sample due to large failures of covered interest rate parity for specific time periods and currencies. The deleted observations include the South African rand from the end of July 1985 to the end of August 1985, and the Indonesian rupiah from the end of December 2000 to the end of May 2007.

Spot and 1-month forward exchange rate data versus the US dollar was obtained from Reuters via Datastream. We conduct our analysis at the monthly frequency, starting from daily data to construct end-of-month exchange rates. Following the tradition in the finance literature since [38], and in the international economics literature since [3], our exposition uses spot and forward rates in logarithms rather than levels. Monthly excess returns for holding foreign currency k, denoted by rx^k, are given by

$$\text{rx}_{t+1}^k = i_t^k - i_t - \Delta s_{t+1}^k \approx \text{fp}_t^k - \Delta s_{t+1}^k = f_t^k - s_{t+1}^k$$

where s and f denote spot and forward exchange rates in logs, respectively. As shown by [39], covered interest parity holds closely in foreign exchange data at the monthly frequency, so the approximation above uses the fact that $i_t^k - i_t \approx \text{fp}_t^k$, where the forward premium in logarithms is defined as $\text{fp}_t^k \equiv f_t^k - s_t^k$.

As in [1], we focus on trading strategies that involve positions exclusively in forward and spot currency markets, as carry trades are easy to implement in such markets, and forward contracts involve minimal default and counter-party risk. Unless otherwise stated, we drop the country superscript k in the rest of the paper for ease of exposition. In presenting our main results in the body of the paper, we ignore transaction costs and compute all strategies and returns using midpoint spot and futures prices. As robustness checks of the main exercises in this paper, we compute our individual country results net of transactions costs due to bid-ask spreads, as well as for subsets of developed and emerging market countries, in both cases following [1]. These results do not change any of our core conclusions, and the interested reader may find them in the Appendix.

To compute the carry and dollar factors we use to forecast exchange rates, following [2]. First, at the end of month t, we create a list of all currencies with data available in months t and $t + 1$, and remove the currency we are predicting from the list. Second, we place the remaining currencies in ascending order by the value of their log forward premium at time t and form six portfolios of roughly equal size. Let rx_{t+1}^{Pj}, denote the month $t + 1$ log returns of portfolios j, $j = 1, \ldots, 6$. Then the

---

[2]　We also study the non-traded volatility risk factor proposed by [35], but find that the carry factor slightly dominates this factor. These results are available upon request.

carry factor, which we denote Carry_{t+1}, is computed as $\text{Carry}_{t+1} = \text{rx}_{t+1}^{P6} - \text{rx}_{t+1}^{P1}$. The dollar factor, which we denote Dollar_{t+1} is computed as $\text{Dollar}_{t+1} = (1/6) \sum_{j=1}^{6} \text{rx}_{t+1}^{Pj}$.

4. The Forecasting Models

The forecasting models for exchange rate returns we consider are the random walk, the random walk with drift, a linear regression estimated via OLS, the BTGP model, and restrictions of the BTGP model. We apply these models to the prediction of next-month carry trade log returns as follows. First, since the random walk is equivalent to assuming $E_t(\Delta s_{t+1}) = 0$, this implies that

$$\text{rx}_{t+1}^{f,\text{RW}} = E_t(\text{rx}_{t+1}) = \text{fp}_t \tag{1}$$

by covered interest rate parity. Second, for the random walk with drift, we take the rolling mean of the past $W = 48$ months of the log spot currency return Δs, and form the forecast

$$\text{rx}_{t+1}^{f,\text{RWWD}} = \text{fp}_t - E_{t,W}(\Delta s_{t+1}), \tag{2}$$

where $E_{t,W}(\Delta s_{t+1}) \equiv (1/W) \sum_{i=t-W+1}^{t} \Delta s_i$. It is important to note that the forecasts of the random walk and random walk with drift models, respectively, is the same for both *ex post* and *ex ante* forecasting exercises, as both methods rely only upon information available at the end of month t. For the formation of *ex post* and *ex ante* forecasts using conditional models, we proceed as follows.

4.1. Ex Post Forecasting versus Ex Ante Forecasting

The *ex-post* out-of-sample model forecast $\text{rx}_{t+1}^{f,M}$, in which the next-month values of fundamentals are given, is defined for model M as

$$\text{rx}_{t+1}^{f,M} = \text{fp}_t - E_{t,W}^M(\Delta s_{t+1}|X_{t+1}), \tag{3}$$

where fp_t is the forward premium at time t, and $E_{t,W}^M(\Delta s_{t+1}|X_{t+1})$ denotes the estimate of the conditional expectation function of model M of exchange rate log returns on fundamentals, evaluated at the month $t + 1$ value of the fundamentals, X_{t+1}, whose elements may be computed using information available on or prior to date $t + 1$. Importantly, the parameters of model M for the conditional distribution of $\Delta s_{\tau+1}|X_{\tau+1}$ are estimated for the rolling observation set $\tau + 1 = t - W + 1, \ldots, t$, and are not a function of information at time $t + 1$.

The *ex-ante* out-of-sample model forecast $rx_{t+1}^{f,M}$, in which the next-month values of fundamentals are unknown, is defined for model M as

$$\text{rx}_{t+1}^{f,M} = \text{fp}_t - E_{t,W}^M(\Delta s_{t+1}|X_{t+1}^f), \tag{4}$$

where the conditional expectation function of model M is now evaluated at X_{t+1}^f, the forecast for month $t + 1$ fundamentals conditional on information available in month t. No information from date $t + 1$ is used to form expectations of date $t + 1$ excess returns. The parameters of model M for the conditional distribution of $\Delta s_{\tau+1}|X_{\tau+1}$ are identical to those in the *ex post* forecasting procedure above. We discuss the two stage factor forecasting procedure, inspired by [11], used to generate X_{t+1}^f in Section 6.2.

4.2. The Bayesian Treed Gaussian Process Model

The Bayesian treed Gaussian process (BTGP) represents a flexible family of nonlinear, nonparametric, regression models built by combining the "choppy" divide-and-conquer thriftiness of trees with rich random-function smoothness of an infinite basis expansion by Gaussian processes (GPs). Although they sound complicated, treed partition models and GPs (taken separately) are

remarkably simple objects. Trees just divide up the input space recursively, usually fitting simple estimates (univariate Gaussian models—averages and residual sums of squares for means and variances respectively) at each element or bottom–node of the partition, called leaves. GPs are likewise conditionally (multivariate) Gaussian models where most of the modeling action involves learning a covariance matrix, specified *a priori* in terms of inverse distances between predictors in the input space. Although both trees and GPs lead to highly flexible non-linear predictors, they have their limitations: trees are inherently non-smooth, and GPs enforce an overly rigid sort of global smoothness in input–output relationships and can require expensive matrix decomposition.

The marriage of trees and GPs, with Bayesian model averaging as the bond, has been demonstrated to push nonlinear regression boundaries in terms of both computation (handling larger data sets) and flexibility, enabling wider application of nonlinear regression in challenging estimation and forecasting contexts.[3] More specifically, the BTGP model involves specifying a prior over regression models which, jointly, partition the input space hierarchically (via binary trees) and specify local mean structure (e.g., via GPs) independently on the elements of the partition depicted by the leaves of the tree. This defines a prior over functions which, when paired with a noise model, e.g., iid Normal (representing the likelihood), describes a posterior distribution. Markov chain Monte Carlo (MCMC) methods can then be deployed to numerically integrate over the *joint* posterior distribution of tree partitions and leaf-model parameters. Although any individual sample from the posterior predictive distribution would be discontinuous, due to the nature of the hard tree partitions and independent modeling at the leaves, averaging over many samples often leads to a surprisingly smooth surfaces provided the number of MCMC samples is large.

Simpler, earlier versions, of a similar idea include the Bayesian CART model and the Bayesian treed linear model (BTLM) [40,41]. These methods represent the Bayesian analogue of the classic CART (Classification and Regression Tree) methods [42] and a kind of Bayesian linear change-point model, respectively. Whereas change-point models are typically one-dimensional, e.g., in time, CART and the BTLM nicely generalize to multiple predictors, where the partitions are delineated by "straight" features in the next lower dimension (lines, planes, hyper-planes, *etc.*) By analogy, the BTGP can be thought of as a multi-dimensional change-point model too, facilitating regime changing nonlinear regressions. In all three examples, the flexibility afforded by trees, of being able to classify observations into different "leaves" of a partition, allows for modeling of differential behavior throughout an input or design space. With GPs and LMs, trees allow for the modeling of non-stationary and heteroskedastic relationships in the data which could not be easily accommodated in their typical, non-treed, formulation. In our particular forecasting setting, a treed LM/GP allows one to learn a functional relationship relating the carry trade return to fundamental variables (in a given month) that is inhomogenious in the spatial "geography" of that design space (of fundamentals and time) by indexing one of several GPs or LMs in the the leaves of a tree. The Bayesian inferential perspectives allows the posterior uncertainty in tree partition and leaf model specification—and therefore the degree of inhomogeniety—to be integrated over via model averaging.

In this paper we focus on the BTGP, but consider its special cases in some detail in order to explore how the different modeling layers, with increasing complexity, lend fidelity and thus accuracy in out of sample predictive tests. The full set of comparators are enumerated as follows:

$$M \in \{\mathrm{LM}, \mathrm{BLM}, \mathrm{BTLM}, \mathrm{BGP}, \mathrm{BTGP}\}, \tag{5}$$

where LM denotes the linear model estimated via ordinary least squares, BLM denotes the Bayesian linear model, BTLM denotes the Bayesian treed linear model, BGP denotes the Bayesian Gaussian process model. BLM, BTLM, and BGP models are restrictions of the full BTGP model: the non-treed

---

[3]   The BTGP has been applied successfully in areas such as computational fluid dynamics, genetics, climatology and political science.

version (omitting a "T" in the acronym) arise by specifying zero prior probability on trees with non-null splits, whereas the non GP options similarly limit the multivariate leaf model to an identity covariance structure. We do not explicitly consider a BCART alternative, although the `tgp` package does indeed provide one. The reason is that the default "ridge" prior on the regression coefficients deployed by BTLM applies substantial shrinkage when linear predictors are irrelevant, ultimately reducing the fit to a mean-only leaf model, *i.e.*, BCART. Such a reduction was never observed in any of our empirical work, and initial experiments showed that BCART was uniformly dominated by BTLM. It was therefore omitted from further consideration.

All of these variants are implemented in the `tgp` package [9] for R. For LM we use the `lm` function in R. In all applications of the BTGP and its GP-based subcases, we use the exponential correlation function and the Bayesian MLE prior discussed in depth in [8,9]. The package takes advantage of the tree partition structure to parallelize many aspects of inference, in order to increase both the Monte Carlo fidelity (more MCMC iterations) and to accomodate larger data sets than typically possible with GP-based models. It implements a "produced–consumer" model via `Pthreads` to queue evaluation, particularly of the posterior predictive equations, on separate threads in a hyperthreaded multi-core system. This is in addition to the threaded linear algebra libraries that are used to parallelize matrix computations. These enhancements are key to handling the sorts of data involved in the modeling of financial time series and other related tasks. Finally, we note that the BTGP prior has many tuning parameters, the choices of which may impact the performance of the method in any application, and our setting in this paper is no different. However, the `tgp` package defaults have been carefully engineered to give good (*i.e.*, non-pathalogical) behavior out of the box, and we exclusively use those settings throughout.

5. The Trading Strategy and Measures of Forecasting Performance

For forecasts $\text{rx}_{t+1}^{f,M}$ of realized excess currency returns rx_{t+1}, we compute several statistics to evaluate the forecasting performance of model M for each individual currency versus the US dollar. First, as in [6,19,26,34,43], we report a directional accuracy statistic, which is equal to the percentage of months in which the prediction of the sign of the next-month carry trade return by a given model is correct.

Second, to measure the economic value of our forecasts of next-month carry trade returns, we employ a simple directional trading strategy of the sort analyzed in [4] and adapted to the foreign exchange market context in [6] for a study of black markets. In particular, given a forecast $\text{rx}_{t+1}^{f,M}$ for the next-period log return of a carry trade that goes short in the US dollar and long in the foreign currency, we consider an "artificial technical analyst" who takes a position according to the sign of the forecast, going long the foreign currency versus the US dollar if the forecast is positive, and going short the foreign currency versus the US dollar if the forecast is negative. The realized (log) returns to such a strategy are given by r_{t+1}, where

$$r_{t+1} = \text{sign}(\text{rx}_{t+1}^{f,M})\text{rx}_{t+1}. \tag{6}$$

From the log returns, we can calculate the discrete returns according to the formula

$$R_{t+1} = \exp(r_{t+1}) - 1. \tag{7}$$

$\text{RX}_{t+1}^{f,M}$ and RX_{t+1} denote the discrete forecast returns and carry trade returns, respectively, and are calculated in a similar manner. For a currency with T observations, realized carry trade returns RX_{t+1} and return forecast $\text{RX}_{t+1}^{f,M}$, the Anatolyev-Gerko statistic is computed as

$$\text{EP} = \frac{A_T - B_T}{\sqrt{\hat{V}_{\text{EP}}}} \xrightarrow{d} N(0,1), \tag{8}$$

where

$$A_T = \frac{1}{T-W} \sum_{t=W}^{T-1} \text{sign}(\text{RX}_{t+1}^{f,M}) \text{RX}_{t+1} \tag{9}$$

and

$$B_T = \left(\frac{1}{T-W} \sum_{t=W}^{T-1} \text{sign}(\text{RX}_{t+1}^{f,M}) \right) \left(\frac{1}{T-W} \sum_{t=W}^{T-1} \text{RX}_{t+1} \right), \tag{10}$$

$T - W$ is the number of out-of-sample periods, and $\hat{V}_{\text{EP}}$ is the feasible estimate of the variance of the numerator of the statistic, for example as is given in [4].

The EP statistic is useful because it measures the extent to which a given forecasting signal is able to generate trades whose profitability (before transactions costs) exceeds that of trades generated by a forecasting signal whose unconditional directional forecasting probabilities are equal to those of the model being tested. Models with significantly positive EP statistics, therefore, are often able to achieve a consistent market timing of nontrivial moves of the exchange rate.

Third, as a transparent measure of realized profitability, we calculate the cumulative monthly return, expressed as an annual percentage rate (APR). The out-of-sample APR, assuming an initial net worth of USD 1, is computed as APR $= 12((V_T)^{1/(T-W)} - 1)$, where V_T is the net worth at the end of the sample period for a given country.

Fourth, we compute the annualized Sharpe ratio for each model and country in the sample, where we apply an adjustment [44] when generating annualized Sharpe ratios from monthly Sharpe ratios to correct for autocorrelation in monthly strategy returns.

Fifth, we report the results of GW unconditional and conditional predictive ability tests, for both *ex post* and *ex ante* out-of-sample forecasting exercises, for each of the six pairs of distinct forecasting models in the set $\{\text{RW}, \text{RWWD}, \text{LM}, \text{BTGP}\}$. We focus on four loss functions of interest: squared error, absolute error, directional (in)accuracy, and (minus) profits:

$$L_{t+1}^{\text{SE}} = (\text{rx}_{t+1} - \text{rx}_{t+1}^f)^2 \qquad L_{t+1}^{\text{AE}} = |\text{rx}_{t+1} - \text{rx}_{t+1}^f|$$
$$L_{t+1}^{\text{DA}} = I(\text{sign}(\text{rx}_{t+1})\text{sign}(\text{rx}_{t+1}^f) < 0) \quad L_{t+1}^{\text{profit}} = -\text{RX}_{t+1}\text{sign}(\text{RX}_{t+1}^f)$$

The GW framework is ideal for our setting, which uses rolling window (limited memory) estimators of a mix of nested and non-nested conditional models, and the unconditional version of the GW test is, as they note, equivalent to the test of Diebold and Mariano [45] in such a setting.

6. Empirical Results

We first present *ex post* forecasting results for models conditioned on a set of traditional macroeconomic variables, and compare the performance of such forecasts against that of models conditioned on our set of global foreign exchange fundamentals. After establishing the dominance of the latter set of factors in *ex post* exercises, we present results on *ex ante* forecasting exercises, in which $t + 1$ values of fundamentals are unknown, following by GW test results for both *ex post* and *ex ante* exercises.

6.1. Ex Post Forecasting Results

6.1.1. Macroeconomic Determinants: The Dornbusch–Frankel Model

As a point of reference, we first test the *ex post* out-of-sample trading strategy performance of strategies conditional on perfect foresight of macroeconomic determinants. In particular, we focus on the Dornbusch–Frankel [46,47] set of exchange rate determinants tested in [3], and run the regression

$$\Delta s_{t+1}^{\text{LM}} = \alpha + \beta(i_t^* - i_t) + \gamma(\Delta m_{t+1}^* - \Delta m_{t+1})$$
$$+ \delta(\Delta y_{t+1}^* - \Delta y_{t+1}) + \tau(\pi_{t+1}^* - \pi_{t+1}) + \epsilon_{t+1}, \tag{11}$$

where $i_t^* - i_t$ is the interest rate differential between country k and the US at the end of month t, $\Delta m_{t+1}^* - \Delta m_{t+1}$ is the difference in the log change of the money supplies between months t and $t+1$, $\Delta y_{t+1}^* - \Delta y_{t+1}$ is the difference in the log changes in the industrial production levels between months t and $t+1$, and $\pi_{t+1}^* - \pi_{t+1}$ is the difference of the 12-month rolling average inflation rates (measured as log changes in the respective consumer price indices) in month $t+1$. The linear model forecast for Δs_{t+1} given by the above model is

$$
\begin{aligned}
E_{t,W}^{\mathrm{LM}}(\Delta s_{t+1}|X_{t+1}^{\mathrm{DF}}) = \hat{\alpha} + \hat{\beta}(i_t^* - i_t) + \hat{\gamma}(\Delta m_{t+1}^* - \Delta m_{t+1}) \\
+ \hat{\delta}(\Delta y_{t+1}^* - \Delta y_{t+1}) + \hat{\tau}(\pi_{t+1}^* - \pi_{t+1}),
\end{aligned}
\tag{12}
$$

where $X_{t+1}^{\mathrm{DF}} = (i_t^* - i_t, \Delta m_{t+1}^* - \Delta m_{t+1}, \Delta y_{t+1}^* - \Delta y_{t+1}, \pi_{t+1}^* - \pi_{t+1})^\top$ is the vector of month $t+1$ Dornbusch–Frankel fundamentals at which the conditional expectation is evaluated to produce the *ex post* linear model forecasts. In an analogous manner, we compute and test *ex post* forecasts $E_{t,W}^{\mathrm{BTGP}}(\Delta s_{t+1}|X_{t+1}^{\mathrm{DF}})$ from the Bayesian treed Gaussian process model.

The results of this exercise are displayed in Table 1. None of the models achieves statistically significant Anatolyev-Gerko statistics for many of the 28 countries in the sample, which is reduced somewhat from the full set of 43 countries due to data availability. The RW achieves an EP statistic that is positive and significant at the 5% level for only four countries, and other models fare worse. The economic value of perfect foresight of macro fundamentals, even with the BTGP model, is low.

Table 1. Performance statistics for *ex post* forecasting exercises: Linear and Bayesian treed Gaussian process (BTGP) models conditional on Dornbusch–Frankel fundamentals. Includes summary statistics for the the directional accuracy percentage, the Anatolyev-Gerko excess profitability (EP) test statistic, the cumulative strategy return, expressed as an annual percentage return (APR), and the annualized Sharpe ratio, corrected for auto-correlation of monthly returns as in [44], for strategies based on model forecasts. The summary statistics reported are the mean, median, minimum and maximum values attained across the sample of 28 countries for which data is available, with the exception of the EP test, for which we report the number of countries in each of three categories: those for which the test is positive and significant at the 5% level, those for which it is insignificant, and those for which it is negative and significant. We also report incidence of a positive EP statistic. The models considered are the benchmarks, the random walk (RW) and random walk with drift (RWWD), and the conditional models, the linear model (LM) and the Bayesian treed Gaussian process (BTGP). The LM and BTGP models are conditional on perfect foresight of the Dornbusch–Frankel model fundamentals discussed in the text, so that one-month ahead individual currency forecasts formed at time t use models estimated up to time t and perfect foresight of time $t+1$ values of fundamentals.

	Model	**RW**	**RWWD**	**LM**	**BTGP**
	Number of countries	28	28	28	28
Dir. accuracy (%)	Mean	0.5458	0.5496	0.5431	0.5406
	Median	0.5574	0.5575	0.5449	0.543
	Min	0.3824	0.4054	0.2353	0.2353
	Max	0.6818	0.6765	0.6475	0.6475
EP test	EP> 0, $p < 0.05$	4	2	2	1
	EP insignificant	24	25	26	26
	EP< 0, $p < 0.05$	0	1	0	1
	Incidence of EP > 0	14	17	17	19
Cum. APR	Mean	0.0187	0.0145	0.0052	0.0045
(pre-TC)	Median	0.0269	0.0217	0.0215	0.0261
	Min	-0.1563	-0.0971	-0.2433	-0.2568
	Max	0.1065	0.1306	0.1104	0.088
Sharpe ratio	Mean	0.2539	0.2257	0.1167	0.1591
(pre-TC)	Median	0.3206	0.285	0.3409	0.2802
	Min	-0.8995	-0.822	-4.0664	-2.5734
	Max	1.3098	1.3879	1.3317	0.9035

6.1.2. Global Foreign Exchange Market Fundamentals: Carry and Dollar Factors

We focus on the forecasting ability of the carry and dollar factors proposed and studied in [1] and [2].[4] Our linear forecasting model, identical to the linear model considered in [2], implies a forecast $\Delta s_{t+1}^{f,\text{LM}}$ equal to:

$$E_{t,W}^{\text{LM}}(\Delta s_{t+1}|X_{t+1}^{\text{GFX}}) = \hat{\alpha} + \hat{\beta}(i_t^* - i_t) + \hat{\gamma}(i_t^* - i_t)\text{Carry}_{t+1} \\ + \hat{\delta}\text{Carry}_{t+1} + \hat{\tau}\text{Dollar}_{t+1}, \tag{13}$$

where $X_{t+1}^{\text{GFX}} = (i_t^* - i_t, (i_t^* - i_t)\text{Carry}_{t+1}, \text{Carry}_{t+1}, \text{Dollar}_{t+1})^{\top}$ is the set of global foreign exchange (GFX) factors studied by [2]. Note the conditional carry factor, $(i_t^* - i_t)\text{Carry}_{t+1}$, allows for the possibility that the beta of the currency excess returns with the carry factor may depend upon the currency's forward premium at time t.

The Economic Value of Perfect Foresight of Dollar and Carry Fundamentals

To measure the value of perfect foresight of the carry and dollar factors, we pose the following question: "If a speculator employing simple directional strategies could estimate the parameters of our conditional models using only data up to the end of month t, but could foresee the end-of-month $t + 1$ values of the fundamentals upon which the models are conditioned, what economic value would such clairvoyance bring her?" The answer to this question may be found in Table 2, which displays directional accuracy statistics, [4] market timing statistics, realized profitabilities and Sharpe ratios for simple directional strategies that take the *ex post* forecasts of our models conditional on the information in X_{t+1}^{GFX} as inputs.

Table 2. Performance statistics for *ex post* forecasting exercises: Linear and BTGP models conditional on global foreign exchange (GFX) factors. The metrics are the same as for Table 1 across the sample of 43 countries for which data is available, with the exception of the excess profitability (EP) test, for which we report the number of countries in each of three categories: those for which the test is positive and significant at the 5% level, those for which it is insignificant, and those for which it is negative and significant. We also report incidence of a positive EP statistic. The models are the same as for Table 1.

	Model	**RW**	**RWWD**	**LM**	**BTGP**
	Number of countries	43	43	43	43
Dir. accuracy (%)	Mean	0.5792	0.5601	0.7648	0.7665
	Median	0.5523	0.5414	0.7978	0.7879
	Min	0.3824	0.3636	0.5455	0.6061
	Max	0.9083	0.9	0.963	0.9444
EP test	EP > 0, $p < 0.05$	9	3	38	38
	EP insignificant	34	39	5	5
	EP < 0, $p < 0.05$	0	1	0	0
	Incidence of EP > 0	25	21	43	43
Cum. APR	Mean	0.0279	0.0098	0.2229	0.2235
(pre-TC)	Median	0.0285	0.0125	0.2452	0.2422
	Min	-0.0714	-0.1701	0.0005	0.0013
	Max	0.1139	0.1306	0.4742	0.5017
Sharpe ratio	Mean	0.374	0.1958	2.0704	2.1052
(pre-TC)	Median	0.3369	0.1809	1.8147	1.6784
	Min	-0.33	-1.6726	0.0827	0.2587
	Max	1.3752	1.3879	4.5951	4.5871

---

4   Our results indicate that the carry factor slightly but unambiguously dominates the nontraded volatility factor in [35] in achieving lower RMSEs and performance statistics. These results are available upon request.

As the results from Table 2 make clear, perfect foresight of the Global Forex set of fundamentals delivers a high degree of prediction accuracy, market timing ability, and realized profitability for simple directional strategies. Somewhat remarkably, the EP test statistics for the conditional models are positive for all 43 countries, and most are significantly different from zero at the 5% level. The average realized annualized returns for the directional strategies based on conditional model forecasts are 22% before transaction costs, compared to just 2.8% for the RW, and the average (median) annualized Sharpe ratios, corrected for autocorrelation of monthly returns as in [44], are 2.1 (1.7) for the BTGP, 2.1 (1.8) for the linear model, 0.37 (0.34) for the RW, and 0.20 (0.18) for the RWWD. These results, importantly, do not reflect gains from diversification of timing strategies across currencies, as in [5].

The Source of *Ex Post* Forecasting Superiority: The Forward Premium, the Carry Factor, and the Dollar Factor

Having established that the models conditional on X_{t+1}^{GFX} significantly outperform the random walk benchmark, it is useful to test the relative contributions of the forward premium, the carry factor terms, and the dollar factor term to achieving this result.

To that end, Table 3 repeats the exercise of Table 2 for linear and BTGP models conditional on three subsets of the Global Forex (GFX) set of factors. The first subset, for which results are reported in panel A, consists of the forward premium (interest rate differential) alone, $X_{t+1}^{\text{fp}} = i_t^* - i_t$. The second set of factors, for which results are reported in panel B, is $X_{t+1}^{\text{GFX}_{\text{Carry}}} = (i_t^* - i_t, (i_t^* - i_t)\text{Carry}_{t+1}, \text{Carry}_{t+1})^\top$. The third set of factors, for which results are reported in panel C, is $X_{t+1}^{\text{GFX}_{\text{Dollar}}} = (i_t^* - i_t, \text{Dollar}_{t+1})^\top$.

For X_{t+1}^{fp}, our *ex post* results are also an *ex ante* exercise comparable to our results in the next section, because the forward premium at time t is known in advance, even though it forms part of X_{t+1}^{GFX}. We find that the BTGP model conditional upon only the forward premium performs similarly to the RW benchmark, attaining a slightly higher median Sharpe ratio and a greater breadth of positive market timing statistics, whereas the RWWD and the LM exhibit inferior performance. These results suggest that models conditional on the forward premium may generate economic value for speculators, consistent with the results of [36] for the G3 currencies versus the US dollar, providing that instabilities in the relationship between the forward premium and excess returns are taken into account.

In Panel B, we find that perfect foresight of the Carry_{t+1} factor contains economic value, in the sense that directional strategies conditional only on $X_{t+1}^{\text{GFX}_{\text{Carry}}}$ are more profitable on average and have higher average Sharpe ratios than those based on the random walk. The improvement is modest, however, considering we are in perfect foresight setting: the average (median) Sharpe ratio across currencies is 0.53 (0.57) for the linear model and 0.51 (0.51) for the BTGP, compared to 0.37 (0.34) for the random walk and only 0.20 (0.18) for the random walk with drift. The Anatolyev-Gerko market timing statistic is positive and significant at the 5% level for 19 countries for the LM and 17 for the BTGP, versus only 9 countries for the random walk model, and 3 for the RWWD. The advantage of conditional models with respect to directional accuracy is positive but small, which leads us to infer that the relatively strong performance of models conditional on the carry factor fundamentals is due primarily to their ability to time large magnitude return events correctly.

Turning to Panel C, we see very clearly that perfect foresight of $X_{t+1}^{\text{GFX}_{\text{Dollar}}}$ alone can explain the high economic value of the *ex post* forecasts from the time series factor model in [2]. The LM and BTGP conditional on the forward premium and dollar factor achieve both mean and median Sharpe ratios of 2.1 for the LM, and 2.0 for the BTGP, respectively, with Anatolyev-Gerko statistics positive and significant at the 5% level or better for 40 out of 43 countries in both cases. Mean directional accuracy for both conditional models is 76%, compared to 58% for the RW and 56% for the RWWD.

Table 3. Performance statistics for *ex post* forecasting exercises: Subsets of GFX factors. The summary statistics and models are the same as for earlier tables. The linear model (LM) and BTGP models are conditional on perfect foresight of the following subsets of the four global forex (GFX) fundamentals discussed in the text: the forward premium (panel A), the forward premium plus the carry and conditional carry factors (panel B), and the forward premium plus the dollar factor (panel C).

GFX Factor Subset:		Benchmarks		Panel A: Forward Premium		Panel B: Carry		Panel C: Dollar	
	Model	RW	RWWD	LM	BTGP	LM	BTGP	LM	BTGP
Dir. accuracy (%)	Mean	0.5792	0.5601	0.5519	0.5698	0.5985	0.5929	0.763	0.762
	Median	0.5523	0.5414	0.5492	0.5667	0.5833	0.5848	0.7942	0.7909
	Min	0.3824	0.3636	0.303	0.3333	0.4242	0.3333	0.5758	0.5152
	Max	0.9083	0.9	0.825	0.875	0.8182	0.8182	0.9074	0.9074
EP test	EP > 0, $p < 0.05$	9	3	5	7	19	17	40	40
	EP insignificant	34	39	36	36	24	26	3	3
	EP < 0, $p < 0.05$	0	1	2	0	0	0	0	0
	Incidence of EP > 0	25	21	27	35	36	32	43	43
Cum. APR (pre-TC)	Mean	0.0279	0.0098	-0.0011	0.0217	0.053	0.0489	0.2332	0.2291
	Median	0.0285	0.0125	0.009	0.0246	0.0547	0.0474	0.266	0.2532
	Min	-0.0714	-0.1701	-0.2559	-0.1872	-0.1191	-0.1487	0.0036	0.0008
	Max	0.1139	0.1306	0.1139	0.1027	0.3262	0.3262	0.5007	0.5098
Sharpe ratio (pre-TC)	Mean	0.374	0.1958	0.1389	0.3524	0.529	0.5116	2.1233	2.0295
	Median	0.3369	0.1809	0.1484	0.3565	0.5711	0.5106	2.0975	1.9984
	Min	-0.33	-1.6726	-1.6781	-1.0858	-0.9747	-0.7985	0.6388	0.1512
	Max	1.3752	1.3879	1.3752	1.3752	2.5624	2.5624	4.5866	4.5866

GW tests of unconditional forecasting ability for the models in Panel B of Table 3 reveal that models conditional on perfect foresight of the carry factor fundamentals are dominated by the random walk for squared error and absolute error loss functions, but dominate the benchmark for directional accuracy and profitability-related loss functions (results available upon request). These findings are consistent with those in [34], suggesting that lower RMSEs do not always imply success with respect to economic criteria. In contrast, GW tests of unconditional predictive ability reveal that the models in Panel C dominate the benchmarks with respect to all four loss functions. These results suggest that [2]'s findings on the psuedo-predictability of individual exchange rate returns by his factor model relative to squared error criteria are driven entirely by the dollar factor. However, as we show in Table 3, perfect foresight of both the carry and dollar factors would clearly have economic value for speculators in individual currencies.

6.2. Ex Ante Forecasting Results

Given the implication of Table 2 that knowing next month's values of the carry and dollar factors would be economically highly valuable to currency speculators employing simple directional strategies, we turn now to the question of whether models conditional only upon time t information can generate forecasts whose economic value, measured by the performance of simple directional strategies, exceeds that of the random walk.

For this purpose, we adapt the linear model in [2], which conditions on the information set X_{t+1}^{GFX}, by replacing the realized values of this vector at time $t + 1$ with its predicted value at time t, $E_t[X_{t+1}^{\text{GFX}}]$. We form the forecasts $X_{t+1}^{\text{GFX},f} = E_t[X_{t+1}^{\text{GFX}}]$ using a set of factor-augmented linear models for Carry_{t+1} and Dollar_{t+1}, following [11]. The key difference between our approach (and that of [11]) is that we forecast the carry and dollar factors, whereas those authors forecast a set of three factors consisting of a subset of individual exchange rate returns (EUR/USD, CHF/USD, JPY/USD).

6.2.1. Forecasting the Carry and Dollar Factors

Our factor forecasting procedure consists of three stages. In stage 1, we estimate initial linear time series models for the factors:

$$\text{Carry}_{t+1}^{f,1} = \hat{\alpha}_C + \hat{\beta}_{\text{CC}}\text{Carry}_t + \hat{\beta}_{\text{CD}}\text{Dollar}_t$$
$$\text{Dollar}_{t+1}^{f,1} = \hat{\alpha}_D + \hat{\beta}_{\text{DC}}\text{Carry}_t + \hat{\beta}_{\text{DD}}\text{Dollar}_t + \hat{\kappa}\overline{\text{fp}}_t^{\text{Dev}}$$

Both equations are estimated using a rolling window of $W = 48$ months, and include the lagged values of both factors. Additionally, the equation for the Dollar_{t+1} factor includes the variable $\overline{\text{fp}}_t^{\text{Dev}}$, which is the average value of developed country forward premiums at time t. This innovation follows [10], which shows that the average value of developed country interest rate differentials has predictive value for the returns of the Dollar_{t+1} factor.

In stage 2, we estimate the models

$$\text{Carry}_{t+1}^{f,2} = \hat{\alpha}_C + \hat{\beta}_{\text{CC}}\text{Carry}_t + \hat{\beta}_{\text{CD}}\text{Dollar}_t + \hat{\lambda}_D\text{Dollar}_{t+1}^{f,1}$$
$$\text{Dollar}_{t+1}^{f,2} = \hat{\alpha}_D + \hat{\beta}_{\text{DC}}\text{Carry}_t + \hat{\beta}_{\text{DD}}\text{Dollar}_t + \hat{\kappa}\overline{\text{fp}}_t^{\text{Dev}} + \hat{\lambda}_C\text{Carry}_{t+1}^{f,1}$$

where $\text{Dollar}_{t+1}^{f,1}$ and $\text{Carry}_{t+1}^{f,1}$ are the forecasted values for the factors from the stage 1 estimation. Given values for $\text{Carry}_{t+1}^{f,2}$ and $\text{Dollar}_{t+1}^{f,2}$, we iteratively substitute these forecasts into the right hand side of the stage 2 equations until achieving convergence for Carry_{t+1}^f and Dollar_{t+1}^f, as in [11].

In the third stage, we replace the actual values of Carry_{t+1} and Dollar_{t+1}, respectively, in our *ex post* model equations with the factor forecasts $E_t(\text{Carry}_{t+t}) = \text{Carry}_{t+1}^f$ and $E_t(\text{Dollar}_{t+1}) = \text{Dollar}_{t+1}^f$

to form *ex ante* forecasts of individual exchange rate returns. Our *ex ante* factor augmented linear model forecast $\Delta s_{t+1}^{f,\text{LM}}$ is then given by

$$E_{t,W}^{\text{LM}}(\Delta s_{t+1}|E_t(X_{t+1}^{\text{GFX}})) = \hat{\alpha} + \hat{\beta}(i_t^* - i_t) + \hat{\gamma}(i_t^* - i_t)\text{Carry}_{t+1}^{f}$$
$$+ \hat{\delta}\text{Carry}_{t+1}^{f} + \hat{\tau}\text{Dollar}_{t+1}^{f}, \tag{14}$$

where $X_{t+1}^{\text{GFX},f} = (i_t^* - i_t, (i_t^* - i_t)\text{Carry}_{t+1}^{f}, \text{Carry}_{t+1}^{f}, \text{Dollar}_{t+1}^{f})^{\top}$ is the *ex ante* version the factors in [2]. The *ex ante* forecasts from the Bayesian treed Gaussian process model, $E_{t,W}^{\text{BTGP}}(\Delta s_{t+1}|E_t(X_{t+1}^{\text{GFX}}))$, are computed as the mean of the posterior distribution for $\Delta s_{\tau+1}|X_{\tau+1}^{\text{GFX}}$ evaluated at $E_t(X_{t+1}^{\text{GFX}})$, and we use CIP relationships to form *ex ante* forecasts of individual exchange rate excess returns, as usual.

Table 4 displays performance statistics for *ex ante* out-of-sample forecasts for the traded carry and dollar factors. For each factor, we compare the factor forecasting procedure of [11] against the random walk with drift, a linear model, and a BTGP model. The linear model and BTGP use the same set of individual predictors for each factor at the [11] (GM) procedure (lagged factors for the carry factor, and lagged factors plus $\overline{\text{fp}}_t^{\text{Dev}}$ for the dollar factor), but unlike the latter procedure, do not use information about the forecast of the other factor in forming predictions. We use the RWWD as our naïve benchmark, rather than the RW, because the factors are priced risk factors whose mean returns are known to be positive and significant (see [1,2,10]).

Table 4. Performance statistics for *ex ante* carry and dollar factor forecasts: [11] model versus competitors. Panel A displays results for the carry factor and Panel B displays results for the dollar factor. The statistics reported include the root mean squared error (RMSE) and the mean absolute error (MAE) of *ex ante* model forecasts, the directional accuracy percentage, the Anatolyev-Gerko excess profitability (EP) statistic, the cumulative strategy return, expressed as an annual percentage return (APR), and the annualized Sharpe ratio, corrected for auto-correlation of monthly returns as in [44], for strategies based on model forecasts. For the EP test, *** indicates significance at the 1% level. The forecasting models are the same as for the other tables. The GM procedure is described in the text.

Factor:	Panel A: Carry			
Model:	**RWWD**	**LM**	**BTGP**	**GM**
RMSE	0.0269	0.0277	0.028	0.0277
MAE	0.0202	0.0207	0.0207	0.0207
Dir. Accuracy	0.6976	0.6855	0.6855	0.6855
EP test	0.46	3.362 ***	2.617 ***	2.984 ***
Cum. APR	0.103	0.119	0.122	0.115
Sharpe ratio	0.762	1.109	1.232	1.088
Factor:	**Panel B: Dollar**			
Model:	**RWWD**	**LM**	**BTGP**	**GM**
RMSE	0.0204	0.0211	0.0211	0.0212
MAE	0.0151	0.0157	0.016	0.0159
Dir. Accuracy	0.5766	0.5565	0.5484	0.5565
EP test	0.097	1.477	1.338	1.372
Cum. APR	0.009	0.026	0.026	0.025
Sharpe ratio	0.135	0.315	0.38	0.309

A few important insights can be gleaned from Table 4, which reports out-of-sample RMSEs, MAEs, directional accuracy, the [4] EP statistic, and the APRs and Sharpe ratios of factor timing strategies that use model forecasts as inputs. First, conditional models obtain similar RMSEs and MAEs out-of-sample, and are unsurprisingly inferior to the RWWD benchmark with respect to these accuracy criteria. Second, we see that the RWWD benchmark also exhibits superior directional

accuracy versus conditional models. Third, we see that conditional models exhibit superior market timing ability versus the RWWD for both factors, and that their ability to time the carry factor is statistically highly significant, whereas their ability to time the dollar factor is not. Finally, we see that factor timing strategies conditional on LM, BTGP, and GM model forecasts exhibit superior profitability and Sharpe ratios versus the RWWD benchmark, with the BTGP model earning the highest Sharpe ratios for both factors. The Sharpe ratio for the carry factor, before transaction costs, for the BTGP is over 1.2, and for the LM and GM models are approximately 1.1. Our result, mirroring [12], indicates that timing carry trade portfolios using conditional models with non-trivial forecasting ability can be an effective strategy for foreign exchange market investors.

6.2.2. *Ex Ante* Exchange Rate Return Forecasts: Decomposing the Performance of the Bayesian Treed Gaussian Process

The performance statistics for *ex ante* forecasting exercises obtained using our carry and dollar factor forecasts, formed using the GM forecasting procedure, are reported in Table 5. Panel A of the table reports results for the benchmarks. Panel B reports results for the linear model. Panel C reports results for the following four subcases of the BTGP: the Bayesian linear model (BLM), the Bayesian treed linear model (BTLM), the Bayesian Gaussian process (BGP), and the full Bayesian treed Gaussian process (BTGP).

The main results of Table 5 are as follows. First, conditional models outperform the benchmarks with respect to market timing ability, as measured by the broadly positive values by country for the [4] excess profitability statistic, which are positive and significant at the 5% level for 13 countries for the full BTGP model, *versus* only 9 countries for the RW model. The directional accuracy results are mixed: conditional models outperform benchmarks for the median country, but the RW achieves a slightly higher average directional accuracy percentage across countries due to very high directional accuracy percentages for the handful of countries in which it performs best. Second, mean and median APRs and Sharpe ratios for directional strategies based on model forecasts generally follow the patterns of performance with respect to the directional accuracy and market timing results, as expected. The RW strategy median Sharpe ratio is 0.358, *versus* 0.424 for the full BTGP model.

Turning to panel C, all subcases of the BTGP outperform the benchmarks and the LM with respect to profitability for the median country, and all but the BTLM outperform the benchmarks and the LM with respect to median Sharpe ratios. However, neither paired t-tests for differences in means, nor paired Wilcoxon signed rank tests for differences in medians, of APRs or Sharpe ratios for the BTGP or its subcases versus the RW are significant at the 10% level (not shown). This is due to large variation across countries in strategy performance for each strategy. Thus outperformance of the BTGP is broad, but not deep. Finally, among the subcases of the BTGP model, the highest median directional accuracy is attained by the BLM and the BTLM, whereas the highest median cumulative profitability and Sharpe ratios are attained by the BGP, which appears to have a slight edge in market timing ability. Taking into account non-stationarity, via trees, seems to improve performance in directional accuracy, whereas taking into account non-linearity, via Gaussian processes, seems to improve performance with respect to market timing.

Table 5. Performance statistics for *ex ante* forecasting exercises: Benchmarks, linear model, and subcases of the BTGP. The *ex ante* GFX factor forecasts are formed using an adaptation of the GM model, as described in the text.

		Panel A: Benchmarks		Panel B: LM		Panel C: Subcases of BTGP		
	Model	RW	RWWD	LM	BLM	BTLM	BGP	BTGP
Dir.	Mean	0.5798	0.559	0.5626	0.5675	0.5714	0.5707	0.5756
accuracy	Median	0.5517	0.5379	0.5667	0.5688	0.566	0.5606	0.5682
(%)	Min	0.3824	0.3636	0.3333	0.3333	0.3824	0.3276	0.3636
	Max	0.9083	0.9	0.8485	0.8485	0.85	0.875	0.8833
EP test	EP > 0, $p < 0.05$	9	3	9	11	6	11	13
	EP insignificant	34	39	32	29	35	29	29
	EP < 0, $p < 0.05$	0	1	2	3	2	3	1
	Incidence of $EP > 0$	25	21	32	32	31	33	34
Cum. APR	Mean	0.0285	0.0087	0.0115	0.0147	0.0186	0.0204	0.0309
(pre-TC)	Median	0.03	0.0112	0.0297	0.0322	0.0303	0.0356	0.0354
	Min	-0.0714	-0.1701	-0.264	-0.264	-0.1965	-0.2167	-0.1564
	Max	0.1156	0.1306	0.2553	0.2553	0.163	0.1782	0.2391
Sharpe ratio	Mean	0.3778	0.1857	0.2538	0.2758	0.3214	0.3625	0.4483
(pre-TC)	Median	0.3582	0.1695	0.3787	0.4081	0.3413	0.4309	0.4239
	Min	-0.33	-1.6726	-1.7623	-1.7623	-0.8017	-1.4444	-0.9104
	Max	1.3752	1.3879	2.109	2.109	1.4546	1.4546	2.1321

6.3. GW Tests of Unconditional and Conditional out-of-Sample Predictive Ability

To get an idea of how directional accuracy and profitability results compare to the results of our models with respect to traditional squared and absolute error criteria, we present results from GW tests of unconditional and conditional forecasting ability, for both *ex post* and *ex ante* out-of-sample forecasting exercises.

Our results for *ex post* forecasting exercises are reported in Table 6, and our results for *ex ante* forecasting exercises are reported in Table 7. Results for the unconditional GW statistics are displayed in Panel A, with results for conditional GW statistics displayed in Panel B, of each table respectively. For a given forecasting exercise, loss function, and pair of methods under comparison, we evaluate the relevant two-sided GW statistic for each country time series in our sample of 43 countries. The three numbers displayed in each such entry of the tables are of the form $a/b/c$, where a is the number of countries for which the model in the column of each sub-table dominates the model in the row at the 10% level or better, b is the number of countries for which neither model significantly dominates the other, and c is the number of countries for which the model in the row dominates the model in the column at the 10% level or better. In all cases we have $a + b + c = 43$.

Table 6. Giacomini–White (GW) tests of unconditional (panel A) and conditional (panel B) predictive ability: *Ex post* out-of-sample forecasting exercises with GFX factors or each of the six pairs of distinct models in the set {RW, RWWD, LM, BTGP}, for four loss functions described in the text: squared forecast error, absolute forecast error, directional (in)accuracy, and (minus) profits. Forecasts from LM and BTGP models are *ex post* forecasts of individual exchange rate excess returns conditional on perfect foresight of the global foreign exchange (GFX) market factors described in the text. Unconditional predictive ability tests are computed using heteroskedasticity and auto-correlation consistent covariance matrix estimators, as in [48], following the procedures laid out in [14]. The triad of numbers displayed for each model comparison are of the form $a/b/c$, where a is the number of countries for which the model in the column of each sub-table dominates the model in the row at the 10% level or better, b is the number of countries for which neither model significantly dominates the other, and c is the number of countries for which the model in the row dominates the model in the column at the 10% level or better. In all cases we have $a + b + c = 43$.

	RW	RWWD	LM	RW	RWWD	LM
Panel A: Tests of Unconditional *ex Post* Forecasting Ability						
		Squared Error			*Abs. Error*	
RWWD	22/21/0			12/31/0		
LM	1/16/26	1/13/29		2/12/29	2/13/28	
BTGP	1/16/26	1/17/25	6/36/1	2/16/25	1/14/28	9/33/1
		Dir. Acc.			*Profits*	
RWWD	5/37/1			7/36/0		
LM	1/12/30	1/12/30		1/10/32	0/9/34	
BTGP	2/9/32	2/9/32	6/36/1	0/11/32	0/10/33	5/37/1
Panel B: Tests of conditional *ex post* forecasting ability						
		Squared Error			*Abs. Error*	
RWWD	12/30/1			9/34/0		
LM	0/17/26	0/16/27		3/14/26	2/13/28	
BTGP	0/19/24	0/19/24	3/40/0	2/17/24	1/16/26	6/37/0
		Dir. Acc.			*Profits*	
RWWD	3/39/1			5/38/0		
LM	1/16/26	1/13/29		0/13/30	0/10/33	
BTGP	2/12/29	2/11/30	1/42/0	0/13/30	0/11/32	1/41/1

Table 7. GW tests of unconditional (panel A) and conditional (panel B) predictive ability: *Ex ante* out-of-sample forecasting exercises with GFX factors. Forecasts from LM and BTGP models are *ex ante* forecasts of individual exchange rate excess returns conditional on *ex ante* forecasts of the global foreign exchange (GFX) market factors formed using the GM model, as described in the text. See Table 6 caption for more details.

	RW	RWWD	LM	RW	RWWD	LM
Panel A: Tests of Unconditional *ex Ante* Forecasting Ability						
	Squared Error			*Abs. Error*		
RWWD	22/21/0			12/31/0		
LM	20/23/0	9/33/1		22/21/0	14/28/1	
BTGP	21/21/1	12/31/0	2/38/3	22/20/1	15/28/0	5/37/1
	Dir. Acc.			*Profits*		
RWWD	5/37/1			7/36/0		
LM	6/35/2	5/36/2		4/39/0	3/37/3	
BTGP	5/36/2	4/37/2	2/38/3	5/36/2	3/35/5	2/38/3
Panel B: Tests of conditional *ex ante* forecasting ability						
	Squared Error			*Abs. Error*		
RWWD	10/32/1			9/34/0		
LM	15/28/0	8/35/0		15/28/0	12/31/0	
BTGP	15/28/0	8/35/0	3/40/0	18/25/0	12/31/0	3/38/2
	Dir. Acc.			*Profits*		
RWWD	3/39/1			5/38/0		
LM	7/35/1	3/35/5		1/42/0	1/39/3	
BTGP	5/38/0	2/37/4	0/40/3	1/41/1	1/39/3	1/41/1

The results, which are largely consistent with our previous findings, can be summarized as follows. First, in Table 6, we see that *ex post* out-of-sample forecasts by LM and BTGP models dominate the benchmarks with respect to all four loss functions, but that the dominance is more pronounced for directional (in)accuracy and (minus) profit loss functions than for squared and absolute error. For the unconditional tests shown in Panel A, the BTGP significantly outperforms the RW at the 10% level for 32 countries for the profitability loss function, *versus* only 26 countries for squared error. Second, we see that the results for conditional tests in Panel B are very similar to those for unconditional tests, but slightly weaker in the sense that fewer strategies display dominance at the 10% level across loss functions.

Turning to Table 7, we see from Panel A that the RW is clearly superior to conditional models with respect to the accuracy-related loss functions, squared and absolute error, as expected. With respect to directional accuracy and profitability-related loss functions, however, the conditional models and benchmarks perform roughly equally. Turning to Panel B, we see that the results of the conditional predictive ability tests are generally less significant than for tests of unconditional predictive ability, similar to what we found for the *ex post* exercises displayed in Table 6. Overall, results from GW tests validate our emphasis on profitability-related criteria as opposed to measures of forecasting accuracy in this paper, consistent with the message in [34]. The results also confirm that perfect foresight of carry and dollar factors confers significant economic value on investors, whereas good *ex ante* forecasts of the factors lead to *ex ante* forecasts for individual exchange rate excess returns whose economic value, as measured in this paper, is quite comparable to that of the random walk benchmark, and for the median country, better.

7. Conclusions

We assess the out-of-sample forecasting performance of linear and nonparametric Bayesian treed Gaussian process (BTGP) models conditional on a set of novel, global foreign exchange carry and dollar factors recently brought to attention in the finance literature. In exercises styled after [3], our conditional models soundly beat the random walk across various measures of forecasting performance. Simple carry-trade timing strategies in individual currencies based on perfect foresight of factors realize large Sharpe ratios, with profitability and market timing ability driven by the dollar factor in particular. Forming *ex ante* forecasts of the traded factors themselves using conditional models, in particular the BTGP, leads to large Sharpe ratios for the carry factor timing strategy, but less so for the dollar factor timing strategy. *Ex ante* currency forecasts based on factor forecasts modestly outperform the random walk benchmark for the median currency with respect to directional accuracy, market timing ability, profitability, and Sharpe ratios. Future research can build upon our results by focusing on the gains from optimal diversification of conditional carry trade timing strategies, among other topics.

Appendix: Robustness Checks

This Online Appendix presents robustness checks on the core results of Tables 2 and 5, pertaining to the RW and RWWD benchmark models, as well as the linear model and BTGP model conditional on the forward premium, the conditional carry factor, the carry factor, and the dollar factor, as described in Section 6.1.2 of the paper. In particular, we present those results by country group, for all countries, developed countries, and emerging market countries, as well as net of transaction costs.

In the tables that follow, the set of "Developed" country currencies consists of those of: Australia, Belgium, Canada, Denmark, the euro area, France, Germany, Italy, Japan, Netherlands, New Zealand, Norway, Sweden, Switzerland, and the United Kingdom, while the set of "Emerging Market" currencies consists of the remaining 28 countries in the dataset. For the five Eurozone countries, we use the cross-rates of their currencies versus the US dollar pre-euro as one set of time series, and the euro versus US dollar time series after the inception of the euro in January 1999 as a separate, non-overlapping currency time series.

For the purpose of computing returns net of transaction costs, in what follows we compute the net log currency excess return for an investor who goes long in a foreign currency as $rx^l_{t+1} = f^b_t - s^a_{t+1}$, whereas the next log currency excess return for an investor who goes short in the foreign currency, and long the dollar, is $rx^s_{t+1} = -f^a_t + s^b_{t+1}$. Here f^b_t denotes the bid price for selling the dollar forward in period t, s^a_{t+1} the ask price for buying dollars at time $t+1$, f^a_t the ask price for buying the dollar forward at time t, and s^b_{t+1} the bid price for selling dollars in the spot market at time $t+1$, respectively.

Table A1 presents robustness exercises for Table 2, and Table A2 presents robustness exercises for Table 5. The relative ranking of models is maintained in these exercises. We find that results for both *ex post* and *ex ante* forecasting exercises are more variable within the set of emerging market countries than the set of developed countries, in terms of the range of realized profitabilities and Sharpe ratios of directional strategies, as is to be expected.

Table A1. Performance statistics for *ex post* forecasting exercises with GFX factors: Results by country group and net of transaction costs.

	Model	RW	RWWD	LM	BTGP
All Countries					
Cum. APR	Mean	0.0101	−0.0076	0.2074	0.2079
(post-TC)	Median	0.0166	−0.0026	0.2244	0.2192
	Min	−0.1189	−0.1826	−0.0027	−0.0018
	Max	0.0986	0.1097	0.4802	0.5077
Sharpe ratio	Mean	0.1643	−0.0044	1.877	1.896
(post-TC)	Median	0.2284	0.0217	1.445	1.5026
	Min	−0.6559	−1.8427	−0.3023	−0.2542
	Max	0.7252	1.1974	4.3518	4.4182
Developed Countries					
Cum. APR	Mean	0.0313	−0.0026	0.237	0.2361
(post-TC)	Median	0.0283	0.0084	0.2829	0.2755
	Min	−0.0095	−0.0521	0.0793	0.08
	Max	0.0836	0.0174	0.3646	0.353
Sharpe ratio	Mean	0.3239	0.0392	2.2891	2.2497
(post-TC)	Median	0.3243	0.1392	2.3287	2.3807
	Min	−0.0334	−0.4391	0.7119	0.8279
	Max	0.653	0.2524	3.9934	3.674
Emerging Market Countries					
Cum. APR	Mean	−0.0012	−0.0103	0.1915	0.1927
(post-TC)	Median	0.0017	−0.0077	0.1804	0.1757
	Min	−0.1189	−0.1826	−0.0027	−0.0018
	Max	0.0986	0.1097	0.4802	0.5077
Sharpe ratio	Mean	0.0788	−0.0277	1.6562	1.7064
(post-TC)	Median	0.1041	−0.0401	1.3278	1.2661
	Min	−0.6559	−1.8427	−0.3023	−0.2542
	Max	0.7252	1.1974	4.3518	4.4182

Table A1 presents robustness exercises for Table 2.

Table A2. Performance statistics for *ex ante* forecasting exercises with GFX factors: Results by country group and net of transaction costs.

	Model	RW	RWWD	LM	BTGP
All Countries					
Cum. APR	Mean	0.0108	−0.0087	−0.003	0.0153
(post-TC)	Median	0.0166	−0.0026	0.015	0.0245
	Min	−0.1189	−0.1826	−0.228	−0.1618
	Max	0.0986	0.1097	0.2432	0.2271
Sharpe ratio	Mean	0.1681	−0.0133	0.0638	0.237
(post-TC)	Median	0.2365	0.0217	0.2136	0.2919
	Min	−0.6559	−1.8427	−1.6912	−0.963
	Max	0.7252	1.1974	2.0263	2.0458
Developed Countries					
Cum. APR	Mean	0.033	−0.0048	0.033	0.0364
(post-TC)	Median	0.0323	0.0053	0.0366	0.0417
	Min	−0.0132	−0.0559	−0.0426	−0.0389
	Max	0.0957	0.0178	0.0906	0.0906
Sharpe ratio	Mean	0.3333	0.0196	0.3316	0.3631
(post-TC)	Median	0.3258	0.1042	0.375	0.3865
	Min	−0.0687	−0.4656	−0.238	−0.2345
	Max	0.6569	0.2561	0.5525	0.6695
Emerging Market Countries					
Cum. APR	Mean	−0.0011	−0.0108	−0.0223	0.004
(post-TC)	Median	0.0017	−0.0085	0.0008	0.0058
	Min	−0.1189	−0.1826	−0.228	−0.1618
	Max	0.0986	0.1097	0.2432	0.2271
Sharpe ratio	Mean	0.0796	−0.0309	−0.0796	0.1694
(post-TC)	Median	0.1212	−0.0419	0.0648	0.2131
	Min	−0.6559	−1.8427	−1.6912	−0.963
	Max	0.7252	1.1974	2.0263	2.0458

Table A2 presents robustness exercises for Table 5.

References

1. Lustig, H.; Roussanov, N.; Verdelhan, A. Common risk factors in currency markets. *Rev. Financ. Stud.* **2011**, *24*, 3731–3777.
2. Verdelhan, A. The Share of Systematic Variation in Bilateral Exchange Rates. *J. Finance*, forthcoming.
3. Meese, R.A.; Rogoff, K. Empirical exchange rate models of the seventies: Do they fit out of sample? *J. Int. Econ.* **1983**, *14*, 3–24.
4. Anatolyev, S.; Gerko, A. A trading approach to testing for predictability. *J. Bus. Econ. Stat.* **2005**, *23*, 455–461.
5. Burnside, C.; Eichenbaum, M.; Rebelo, S. Carry Trade: The Gains of Diversification. *J. Eur. Econ. Assoc.* **2008**, *6*, 581–588.
6. Gramacy, R.; Malone, S.W.; Horst, E.T. Exchange Rate Fundamentals, Forecasting, and Speculation: Bayesian Models in Black Markets. *J. Appl. Econom.* **2014**, *29*, 22–41.
7. Sarno, L.; Valente, G. Exchange rates and fundamentals: Footloose or evolving relationship? *J. Eur. Econ. Assoc.* **2009**, *7*, 786–830.
8. Gramacy, R.B.; Lee, H.K. Bayesian treed Gaussian process models with an application to computer modeling. *J. Am. Stat. Assoc.* **2008**, *103*, 1119–1130.
9. Gramacy, R.B. tgp: An R package for Bayesian Nonstationary, Semiparametric nonlinear Regression and Design by treed Gaussian process models. *J. Stat. Softw.* **2007**, *19*, 6.
10. Lustig, H.; Roussanov, N.; Verdelhan, A. Countercyclical currency risk premia. *J. Financ. Econ.* **2014**, *111*, 527–553.
11. Greenaway-McGrevy, R.; Mark, N.C.; Sul, D.; Wu, J.L. Exchange Rates as Exchange Rate Common Factors; Technical Report; 2012. Available online: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2138713 (accessed on 2 July 2015).
12. Bakshi, G.; Panayotov, G. Predictability of currency carry trades and asset pricing implications. *J. Financ. Econ.* **2013**, *110*, 139–163.
13. Johannes, M.; Korteweg, A.; Polson, N. Sequential learning, predictability, and optimal portfolio returns. *J. Financ.* **2014**, *69*, 611–644.
14. Giacomini, R.; White, H. Tests of Conditional Predictive Ability. *Econometrica* **2006**, *74*, 1545–1578.
15. Nag, A.K.; Mitra, A. Forecasting daily foreign exchange rates using genetically optimized neural networks. *J. Forecast.* **2002**, *21*, 501–511.
16. Preminger, A.; Franck, R. Forecasting exchange rates: A robust regression approach. *Int. J. Forecast.* **2007**, *23*, 71–84.
17. Liu, T.R.; Gerlow, M.E.; Irwin, S.H. The performance of alternative VAR models in forecasting exchange rates. *Int. J. Forecast.* **1994**, *10*, 419–433.
18. Yang, J.; Su, X.; Kolari, J.W. Do Euro exchange rates follow a martingale? Some out-of-sample evidence. *J. Bank. Financ.* **2008**, *32*, 729–740.
19. Cerra, V.; Saxena, S.C. The monetary model strikes back: Evidence from the world. *J. Int. Econ.* **2010**, *81*, 184–196.
20. Alexander, D.; Thomas, L.R., III. Monetary/Asset models of Exchange Rate Determination: How well have they Performed in the 1980's? *Int. J. Forecast.* **1987**, *3*, 53–64.
21. Wolff, C.C. Time-varying parameters and the out-of-sample forecasting performance of structural exchange rate models. *J. Bus. Econ. Stat.* **1987**, *5*, 87–97.
22. Wolff, C.C. Models of exchange rates: A comparison of forecasting results. *Int. J. Forecast.* **1988**, *4*, 605–607.
23. Diebold, F.X.; Nason, J.A. Nonparametric exchange rate prediction? *J. Int. Econ.* **1990**, *28*, 315–332.
24. Meese, R.A.; Rose, A.K. Nonlinear, nonparametric, nonessential exchange rate estimation. *Am. Econ. Rev.* **1990**, *80*, 192–196.
25. Engel, C.; Hamilton, J.D. Long Swings in the Dollar: Are They in the Data and Do Markets Know It? *Am. Econ. Rev.* **1990**, *80*, 689–713.
26. Engel, C. Can the Markov Switching Model forecast exchange rates? *J. Int. Econ.* **1994**, *36*, 151–165.
27. Kuan, C.M.; Liu, T. Forecasting exchange rates using feedforward and recurrent neural networks. *J. Appl. Econom.* **1995**, *10*, 347–364.
28. Brooks, C. Linear and Non-linear (Non-) Forecastability of High-frequency Exchange Rates. *J. Forecast.* **1997**, *16*, 125–145.

29. Gencay, R. Linear, non-linear and essential foreign exchange rate prediction with simple technical trading rules. *J. Int. Econ.* **1999**, *47*, 91–107.

30. Qi, M.; Wu, Y. Nonlinear prediction of exchange rates with monetary fundamentals. *J. Empir. Financ.* **2003**, *10*, 623–640.

31. Rogoff, K. The Failure of Empirical Exchange Rate Models: No Longer New but Still True. *Econ. Policy Web Essay* **2001**, *1*(1).

32. Engel, C.; Mark, N.C.; West, K.D. Factor Model Forecasts of Exchange Rates; Technical Report; 2012. Available online: http://www.nber.org/papers/w18382 (accessed on 2 July 2015).

33. Rossi, B. Exchange Rate Predictability. *J. Econ. Lit.* **2013**, *51*, 1063–1119.

34. Leitch, G.; Tanner, J.E. Economic forecast evaluation: Profits versus the conventional error measures. *Am. Econ. Rev.* **1991**, *81*, 580–590.

35. Menkhoff, L.; Sarno, L.; Schmeling, M.; Schrimpf, A. Carry trades and global foreign exchange volatility. *J. Financ.* **2012**, *67*, 681–718.

36. Della Corte, P.; Sarno, L.; Tsiakas, I. An Economic Evaluation of Empirical Exchange Rate Models. *Rev. Financ. Stud.* **2009**, *22*, 3491–3530.

37. Cochrane, J. *Asset Pricing*; Princeton University Press: Princeton, NJ, USA, 2001.

38. Fama, E.F. Forward and spot exchange rates. *J. Monet. Econ.* **1984**, *14*, 319–338.

39. Akram, Q.F.; Rime, D.; Sarno, L. Arbitrage in the foreign exchange market: Turning on the microscope. *J. Int. Econ.* **2008**, *76*, 237–253.

40. Chipman, H.A.; George, E.I.; McCulloch, R.E. Bayesian CART Model Search (with discussion). *J. Am. Stat. Assoc.* **1998**, *93*, 935–960.

41. Chipman, H.A.; George, E.I.; McCulloch, R.E. Bayesian Treed Models. *Mach. Learn.* **2002**, *48*, 303–324.

42. Breiman, L.; Friedman, J.H.; Olshen, R.; Stone, C. *Classification and Regression Trees*; Wadsworth: Belmont, CA, USA, 1984.

43. Abhyankar, A.; Sarno, L.; Valente, G. Exchange rates and fundamentals: Evidence on the economic value of predictability. *J. Int. Econ.* **2005**, *66*, 325–348.

44. Lo, A.W. The Statistics of Sharpe Ratios. *Financ. Anal. J.* **2002**, *58*, 36–52.

45. Diebold, F.X.; Mariano, R.S. Comparing Predictive Accuracy. *J. Bus. Econ. Stat.* **1995**, *13*, 253–263.

46. Dornbusch, R. Expectations and Exchange Rate Dynamics. *J. Political Econ.* **1976**, *84*, 1161–1176.

47. Frankel, J.A. On the Mark: A Theory of Floating Exchange Rates Based on Real Interest Rate Differentials. *Am. Econ. Rev.* **1979**, *69*, 601–622.

48. Andrews, D.W.K. Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. *Econometrica* **1991**, *59*, 817–858.