# THE PREDICTIVE POWER OF BEHAVIORAL ECONOMICS IN THE FOREIGN EXCHANGE MARKET

VAJIHEH MAHDIYAN AMIRABADI

A DISSERTATION SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN ECONOMICS
YORK UNIVERSITY
TORONTO, ONTARIO

MARCH, 2022

# Abstract

This dissertation explores the out-of-sample forecastability of changes in exchange rates using behavioral economics and ensemble (combination) methods and contributes to the literature by introducing three approaches presented in three essays.

The first essay explores a new approach through behavioral heuristics, for example, overconfidence, to forecast changes in exchange rates. One key aspect of behavioral economics is that people do not use the realized distribution of historical data to predict variables in the traditional way. Instead, agents assume and assign subjective probabilities which depend on underlying heuristics in decision-making. These subjective probabilities are used as weights for the ten years of observations prior to time $t$ in linear models of changes in exchange rates used to estimate coefficients and form forecasts. This essay develops forecasting models for exchange rates using monthly data for the US dollar versus 37 (advanced and emerging/developing) currencies. In the first essay, monthly data are from Feb 1973 to Feb 2020 for all advanced and some emerging/developing countries. Data sets for some developing countries may start later.

The second essay incorporates behavioral economics by adding investor sentiment variables to macroeconomic models to form forecasts for exchange rates for the US dollar versus 37 (advanced and emerging/developing) currencies. An example of an investor sentiment index is a composite leading indicator index. In addition, I examine the predictive ability of a terms of trade index (in changes) both as a single predictor and as an added predictor in the uncovered interest rate parity (UIRP) model. Finally, I examine changes in commodity and oil prices (both in nominal and real terms) as predictors of changes in exchange rates.

In the second essay, all countries' monthly data run from Jan 2000 to Dec 2020.

Even though there are some differences between the first and second essays, including explanatory variables (predictors) and the length of data for changes in exchange rates, empirical findings are similar. In both essays, none of forecasting models outperforms the White Noise (WN) model according to statistics based on the ratio of Mean Squared Prediction Error ($\text{MSPE}_{ratio}$) or Diebold-Mariano (DM) statistics. Nevertheless, in both essays, some models provide promising results for some currencies using Pesaran-Timmermann (PT) statistics. In the first essay, models under the assumptions of anchoring-toward and optimism perform well in 1-month-ahead and 12-month-ahead forecasts for Mexico. In the second essay, the Taylor Rule model augmented by the composite leading indicator outperforms the WN model in predicting the direction of change in Mexico's exchange rates at 1-month-ahead. The Extended UIRP model included a business confidence index successfully outperforming the WN model in terms of PT statistics at 12-month-ahead forecasts for Mexico.

The third essay of the dissertation addresses model uncertainty. It focuses on combining forecasts from various individual models introduced in the second essay. I use standard (e.g., equal and MSPE weightings) and new forecast combination approaches to form combined forecasts. In particular, using a regularization technique (Ridge regression), I apply linear combination and convex combination estimates to combine forecasts. I also propose Directional Prediction as a new weighting approach. In addition, I contribute to the literature by combining forecasts from models that simultaneously include macroeconomic and investor sentiment variables. The results indicate that while combination approaches are not successful in predicting point forecasts in terms of $\text{MSPE}_{ratio}$ or DM statistics, these approaches perform well according to PT statistics. In emerging/developing countries, linear combination (Ridge regression) and Directional Prediction weighting approaches have proportions of correct direction of changes in exchange rates greater than 0.5, which means they perform better than the WN model.

# Acknowledgements

I am profoundly grateful to my supervisory committee members, Professor Joann Jasiak, Professor J. Barry Smith, and Professor Brenda Spotton Visano, for supporting and assisting me in finishing my study. I could not have completed my study without their help and insightful guidance, inputs, and comments.

I would like to express my gratitude to Professor Andrei Semenov, who contributed significantly to this dissertation's development and interpretation of models. I also appreciate Professor Henry Tam, Professor Tsvetanka Karagyozova, and Professor Andrey Stoyanov for their helpful suggestions during my study. In addition, I want to thank Professor Ahmet Akyol and Professor Aryn Martin, who made it possible to finish my study.

I am deeply indebted to my husband, Dr. Mahdi Lotfinezhad, for his endless love, professional support, and inspiration. I owe a huge thanks to my parents for their unconditional love and encouragement. I want to thank my sisters, who always encouraged me to stay strong and believe in myself. I dedicate this dissertation to my son, Farham, whose love has made me keep going and be resilient.

Finally, I wish to thank all the people whose assistance was instrumental in the completion of this dissertation.

# Contents

# List of Tables

# List of Figures

# Introduction

In this dissertation, I investigate (a) whether allowing the decision making of agents to diverge from strict rationality and follow bounded rationality or "irrationality" can lead to better forecasts of changes in exchange rates and (b) how to address model uncertainty. I show that accounting for "irrationality" and model uncertainty leads to improved out-of-sample forecasting of exchange rates for various currencies.

Forecasting changes in exchange rates has remained a fundamental economic challenge. Accuracy in predicting foreign exchange rate movements is crucial for almost any future international investment. Moreover, the volatility in currency values and increasing international transactions further necessitates studying and forecasting currency movements. Forecasting exchange rates provides opportunities for exporters and importers to make better decisions considering costs and revenues from global operations.

In addition, monetary policymakers must understand exchange rates movements. This understanding helps keep inflation stable at an appropriate moderate level and economic activity at a high level. Policymakers also monitor the foreign exchange market because currencies are financial assets that reveal important economic and financial conditions. For central banks to intervene efficiently in the foreign exchange market and occasionally restrict the free movement of exchange rates, they have to characterize the directions of exchange rate changes and their consequences.

Understanding and forecasting exchange rates movements has remained a difficult task. According to the study by Engel, Mark, and West (2015), a random walk without drift model works well compared to other models in forecasting exchange rates between countries with floating regimes and similar inflation rates. The random walk without drift prediction is when the exchange rate (log) level is predicted to stay at the current (log) level. In other words, the forecast is one of "no change" in exchange rates [Molodtsova and Papell (2009) and Engel et al. (2015)].

In the light of the above, the main question that researchers are trying to answer is whether exchange rates can be forecasted. Overall, the literature shows that the performance of existing models depends on the choice of predictors (e.g., macroeconomic and financial factors), estimation methods, evaluation tests, sample periods, benchmark models, and horizons. Yet, as mentioned above, existing models generally do not outperform the random walk model. In this context, a challenge is to develop models that forecast exchange rates more accurately than the random walk.

This dissertation proposes new approaches to examining the forecastability of exchange rates. It explores behavioral economics and takes into consideration behavioral aspects of decision-making. This approach considers the impacts of psychological, social, cognitive, and emotional factors on the economic decisions of individuals.

Behavioral economics is not a new concept. Economists in academia and financial practitioners have increasingly paid attention to this concept and incorporated this alternative framework in their modeling approaches. As an example, Richard Thaler of the University of Chicago was awarded the Nobel prize in Economic Sciences in 2017 for his work and contributions to behavioral economics. The Royal Swedish Academy of Sciences in Stockholm stated that Richard Thaler "has incorporated psychologically realistic assumptions into analyses of economic decision-making. By exploring the consequences of limited rationality, social preferences, and lack of self-control, he has shown how these human traits systematically affect individual decisions as well as market outcomes."

An excellent example highlighting the effects of people's perceptions on events is the coronavirus outbreak and its negative impacts on the global economy and financial markets. This pandemic has led to planes being grounded, cruise ships quarantined, and has stopped car production. The global economy has suffered significant hits as governments have tried to stop the virus's spread, help their citizens financially, and encourage them to get vaccines.

Given this backdrop, pessimism as an example of behavioral heuristics initially dominated the view toward the Covid virus and its potential vaccines; the uncertain long-term effects of both virus and vaccines made matters worse.

This event has impacted all countries. Many people compare the coronavirus outbreak to the SARS epidemic, which in 2003 caused financial losses. But economic losses by the coronavirus could be much more significant because China has expanded into a far larger economy in the past 17 years and is much more tightly connected to the rest of the world economy as the global economy has increasingly integrated. The interruption of supply chains in China and uncertainty about its financial market have directly impacted other countries' markets. While China was trying to control the spread of the virus, its economic growth tumbled, and its currency depreciated against the US dollar. When China had made progress preventing the virus from spreading, other countries, including the US and Canada, were impacted by the virus. When the virus first hits a country, the perceived and subjective chance of economic growth decreases because of pessimism and the uncertainty that parts of the economy might need to be closed. In addition, the lockdown has negatively impacted economic growth. As a country obtains better control of the virus and the number of positive cases and death tolls declines, optimism about the re-opening of the economy takes control of the market narrative.

Many companies and corporations use the behavioral economics framework through artificial intelligence, machine learning, and predictive behavioral analytics to minimize risk and fraud. As an example, ForMotive[1] provides services to companies and enterprises by different types of intelligence about the actions and behavior of users. ForMotive helps companies/businesses to analyze their users' behaviors and obtain information about causes for particular behaviors. This information allows companies/businesses to understand the intentions, incentives, and plans of customers/users and predict users' future behaviors.

Given the impact of people's behaviors and perceptions on the economy and financial

[1] https://www.formotiv.com

market, this dissertation uses behavioral economics in out-of-sample forecasting models of exchange rates. Moreover, individual investors, corporations, and governments face uncertainty about factors influencing exchange rate movements. They do not know which models are the best forecasting models and which forecasts to use. I use ensemble (i.e., combining forecasts) methods to form forecasts for exchange rates to address model uncertainty. Thereby, the dissertation contributes to the literature by introducing three essays as follows:

- *The first essay proposes a new approach through behavioral heuristics to develop forecasting models for monthly nominal exchange rates. In contrast to previous studies, which implicitly assume rationality, I construct forecasting models incorporating behavioral economics. Under behavioral economics assumptions, people assign different weights to the past historical data (in this essay, the previous ten years of observations). Underlying heuristics determine these weights. Examples of heuristics include optimism and pessimism. In this essay, I use weights to estimate linear models for changes in exchange rates and construct forecasts. I use heuristics both individually and in combinations of mutually exclusive heuristics to determine weights.*

- *The second essay investigates the out-of-sample predictability of monthly nominal exchange rates. It incorporates behavioral economics using investor sentiment indices (e.g., the composite leading indicator and consumer confidence indices) in combination with macroeconomic models. This is in contrast to the first essay, where only the historical exchange rate data are used for forecasting. Macroeconomic models include the purchasing power parity (PPP), the uncovered interest rate parity (UIRP), and the Taylor Rule model, among other models. I also evaluate the out-of-sample performance of models, including the oil price and commodity price (in changes and both real and nominal terms). Moreover, I examine the predictive ability of changes in terms of trade index and consider including changes in terms of trade index or the inflation rate differential (the difference between the US and a foreign country inflation rates) to the UIRP model to improve the forecast performance.*

- *The third essay takes a different perspective and focuses on combining forecasts from* *various individual models, introduced in the second essay. I use standard weighting approaches, including equal weights and Mean Squared Prediction Error (MSPE), and I propose Directional Prediction as a new weighting approach. Moreover, I use a regularization technique (Ridge regression) in the linear and convex combination methods to combine forecasts.*

The first essay develops forecasting models for nominal exchange rates using monthly data for the US dollar versus 37 (advanced and emerging/developing) countries. The data sets are from Feb 1973 to Feb 2020 for all advanced and some emerging/developing countries. For other emerging/developing countries, the data start later; for example, the data start from Jan 1979 for South Africa (see Table 1.2 for other countries). I construct forecasting models incorporating behavioral heuristics. I use behavioral heuristics as follows:

- Availability heuristic
- Conservatism
- Pessimism
- Optimism
- Overconfidence
- Underconfidence

I also use the median of historical data as an anchor and present two heuristics as follows:

- Anchoring-toward
- Anchoring-away

In addition, I consider the case of rationality, where people use the realized probability distribution of data and implicitly assign equal weights to the exchange rate historical data.

In the first step, I use heuristics individually to construct subjective probabilities. In the next step, I combine mutually exclusive heuristics to construct people's total subjective probability as follows:

- Rationality, Availability heuristic, Conservatism
- Rationality, Optimism, Pessimism
- Rationality, Overconfidence, Underconfidence
- Rationality, Anchoring-toward, Anchoring-away

To assess the out-of-sample forecasting performance of models and compare them with the White Noise (WN) model, I use various test statistics. In particular, I use common statistics in the literature, including the ratio of Mean Squared Prediction Error ($\text{MSPE}_{ratio}$), the Diebold-Mariano (DM) developed by Diebold and Mariano (1995), and the Pesaran-Timmermann (PT) developed by Pesaran and Timmermann (1992).

For financial practitioners and policymakers, the ability to precisely predicts the direction of changes in exchange rates could be as helpful as a correct point forecast. Therefore, it is essential to examine whether models successfully predict both point forecasts and the direction of changes in exchange rates. The statistics, $\text{MSPE}_{ratio}$ and DM, are used to evaluate models' ability to predict point forecasts accurately. However, PT statistics examine models' ability to predict the direction of changes in exchange rates correctly.

The empirical results show no model successfully outperforms the WN model for all countries in terms of out-of-sample forecast evaluated by statistics such as the $\text{MSPE}_{ratio}$ or DM. However, behavioral models provide promising forecasting results for some currencies by PT statistics. Particularly, models under the assumptions of anchoring-toward and optimism heuristics outperform the WN model in 1-month-ahead and 12-month-ahead forecasts, in some emerging/developing countries, for example, Mexico.

The second essay explores the out-of-sample predictability of monthly exchange rates for the US dollar versus 37 (advanced and emerging/developing) countries. I examine behavioral factors' predictive ability through investor sentiment indices in combinations with macroeconomic models. Investor sentiment indices are as follows:

- Consumer Confidence Index (CCI)

- Business Confidence Index (BCI)
- Composite Leading Indicator Index (CLI)
- Cboe Volatility Index (VIX)[2]
- Cboe SKEW Index (SKEW)[3]

Macroeconomic models include the purchasing power parity (PPP) model, the uncovered interest rate parity (UIRP) model, and the Taylor Rule model, among others. I also investigate the predictive power of oil and commodity prices (in changes) and changes in terms of trade index. Moreover, I examine the uncovered interest rate parity model's out-of-sample predictive ability when added to the inflation rate differential (i.e., the difference between the US inflation rate and a foreign country inflation rate, $\pi_t^{US} - \pi_t^{Foreign}$) or changes in terms of trade index.

For the assessment of forecasting models, I use the same statistics as in the first essay. The empirical results show that while forecasting models are not very successful in forecasting changes in exchange rates in terms of $\text{MSPE}_{ratio}$ or DM statistics, these models do well according to PT statistics. In particular, the results indicate that models containing behavioral factors outperform the WN model for several countries in 1-month-ahead and 12-months-ahead forecasts. A 12-months-ahead forecast means I use the ten years of observations, for example, from Jan 1973 to Dec 1982, to form a forecast for Dec 1983.

The empirical evidence shows forecasting models included changes in oil and commodity prices (in both real and nominal terms), are successful in predicting the correct direction of changes for Canada, Brazil, Germany, Sweden, and the United Kingdom's exchange rates.

The findings indicate that adding investor sentiment variables to macroeconomic models improves forecasting results in terms of PT statistics. Moreover, models that include macroeconomic and investor sentiment variables improve forecasting results for several countries,

[2] This index measures the market's expectation of future volatility. The VIX index is based on options of the $S\&P$ 500 index [Cboe (2020b)].

[3] This index indicates the skewness of $S\&P$ 500 returns at the end of a 30-day horizon [Cboe (2020a)].

especially for emerging/developing countries. Still, models do not outperform the WN model according to the ratio of MSPE and DM statistics.

The third essay addresses model uncertainty and combines forecasts from various individual models, explained in the second essay. There are two reasons which make combining forecasts an important task. First, individual investors, corporations, and governments, among other stakeholders, face uncertainty about the foreign exchange market. Second, individual models may have weak out-of-sample performance, yet they include some information.

I introduce a new weighting method and call it Directional Prediction.[4]. Furthermore, I explore linear and convex combinations by applying a regularization technique (Ridge regression) to form combined forecasts for exchange rates. Moreover, I use common weighting approaches in the literature, such as equal and MSPE. Monthly changes in exchange rates are for the US dollar versus 37 (advanced and emerging/developing) currencies using various individual models per currency.

My findings show weighting methods do not outperform the WN model in terms of $\text{MSPE}_{ratio}$ or DM statistics, but these methods do well in terms of PT statistics. The linear combination (Ridge regression) provides better results than the WN model in advanced countries in terms of $\hat{p}$ measures in both advanced and emerging/developing countries at 1-month-ahead forecasts. On average, in emerging/developing countries, the Directional Prediction weighting method improves forecasting results compared to the WN model at 12-months-ahead forecasts. In contrast to the results for emerging/developing countries, on average, no weighting method yields better forecasting results for advanced countries.

[4]This method sets the proportion of correct directions for a given model as a weight. Correct direction means both a forecast and its associated realized value have the same sign.

# Essay 1

# Behavioral Heuristics and Exchange Rate Forecasting

## 1.1 Introduction

The study of Meese and Rogoff (1983a) has become one of the most important papers in the exchange rate literature. Since their study, researchers have considered different models and approaches to forecast movements of exchange rates and do better than a random walk model, but the main problem is that these models cannot outperform the random walk model for all currencies. Forecasting from the random walk model has the same out-of-sample mean squared error and often smaller out-of-sample mean squared error than models that use fundamental variables including interest rate, output, money, and inflation [see, e.g., Meese and Rogoff (1983a), Wright (2008), Rossi (2013), Engel et al. (2015), Cheung, Chinn, Pascual, and Zhang (2019)].

Engel et al. (2015) develop factors from a cross-section of exchange rates and use the idiosyncratic deviation from factors to forecast. They forecast movements of exchange rates using fundamental variables suggested by the purchasing power parity, Taylor Rule, and monetary models. They choose a random walk without drift model as a benchmark model and compare their constructed models with this model. Engel et al. (2015) get mixed forecasting results. Except for a particular subsample (quarterly data from 1999 to 2007), their results suggest no improvement in forecasting movements of exchange rates.

S. Chen and Hsu (2019) evaluate whether stock returns could predict daily exchange rate movements. Using the uncovered equity parity condition, they find that stock return differentials have in-sample and out-of-sample predictive power for exchange rates in 1-day-ahead predictions.

Gourieroux, Jasiak, and Tong (2021) propose new techniques of predicting and filtering for the causal–noncausal convolution model. This model shows the dynamics of stationary processes with local explosions, such as bubbles and spikes, which represent the financial and macroeconomic variables time series, including commodity prices and cryptocurrency exchange rates. The new filtering and forecasting methods lead to favorable results in application to the WTI crude oil prices series.

Applying the combination of principal component analysis and linear regression, Jaworski (2021) predicts exchange rates for Central and Eastern European currencies using global and country-specific factors. This approach leads to better out-of-sample results compared to the random walk model at horizons of one month to over a year in.

The main question that researchers are trying to answer is whether exchange rates can be forecasted. Overall, the literature shows the performance of existing models depends on the choice of predictors (e.g., macroeconomic and financial factors), estimation methods, evaluation tests, sample periods, benchmark models, and horizons. A challenge is to develop forecasting models to perform better than the White Noise (WN) model in predicting point forecasts and the direction of change in exchange rates.

This essay proposes a new approach to investigate the forecastability of exchange rates. It explores a behavioral economics approach and takes into consideration behavioral aspects of decision making. This approach considers the impacts of psychological, social, cognitive, and emotional factors on individuals' economic decisions. The method is based on behavioral economics (finance), which has been developed to incorporate model uncertainty and behavioral heuristics into modeling financial markets. A fundamental assumption in most traditional models is that people are rational.

There are theoretical and empirical studies that have used behavioral heuristics, sentiment indices, and model uncertainty into their modeling procedures to resolve financial market anomalies [see, e.g., Abel (2002), Semenov (2009), Markiewicz (2012), Beckmann and Schüssler (2014), Kouwenberg, Markiewicz, Verhoeks, and Zwinkels (2017), Hauzenberger and Huber (2020), and Ito, Masuda, Naito, and Takeda (2021)].

Abel (2002) uses an asset pricing model to examine the implications of pessimism and doubt heuristics on the mean equity premium and the risk-free rate. He concludes that pessimism and doubt heuristics could help resolve the risk-free rate puzzle by explaining the mean equity premium and the risk-free rate. In another economic context, Baicker, Congdon, and Mullainathan (2012) consider health insurance coverage issues from a behavioral economics perspective. Their work provides evidence that the psychology of individual decision-making plays an important role in driving coverage outcomes. Baicker et al. (2012) discuss a set of findings from psychology and behavioral economics that indicate the possible role of deviations from perfectly rational behavior in health insurance take-up.

One central aspect of behavioral economics is that people assign subjective probabilities as opposed to objective probabilities in decision-making. I use historical exchange rate data to forecast changes in exchange rates, focus on the probability distribution of exchange rates, and consider subjective probabilities instead of objective probabilities. These subjective probabilities depend on heuristics used [see details in Section 1.2.1]. These subjective probabilities are used as weights in linear models that I estimate to forecast currencies [see details in Section 1.4].

I incorporate behavioral heuristics, including the availability heuristic, conservatism, pessimism, optimism, overconfidence, underconfidence to construct subjective probabilities [see details in Section 1.2.1]. This essay focuses on the conditional mean and the associated heuristics and leaves out the conditional volatility and the related heuristics for further research. I use the median of historical data as an anchor and present two heuristics, anchoring-toward and anchoring-away. Moreover, I consider the case of rationality. Under the rationality assumption, people believe the past historical data are equally important for predicting exchange rates (the past data should have equal weights).

In the first step, I use heuristics individually (extreme cases) to construct subjective probabilities. Next, I combine mutually exclusive heuristics to construct people's total subjective probabilities. For example, rationality, optimism, and pessimism are mutually exclusive events.

My goal is to form forecasting models and evaluate the out-of-sample performance of these models. On a rolling window basis, I split a given data set into an in-sample period to estimate a weighted linear model and an out-of-sample period to evaluate forecasting performance. Weights are determined according to individual or combinations of heuristics.

For assessments, I compare the out-of-sample forecasting performance of models with a benchmark model, the White Noise (WN) model. I use the ratio of Mean Squared Prediction Error ($\text{MSPE}_{ratio}$), Diebold-Mariano (DM), and Pesaran-Timmermann (PT) statistics. In addition to the statistical forecast evaluation measures, I use the Sharpe-ratio statistic to evaluate investment strategies, using the expected exchange rate change to calculate returns on investments.

In the next section, I present the modeling of behavioral heuristics. Section 1.3 describes data in this essay and Section 1.4 explains methodologies and implementations. Section 1.5 reports empirical results and Section 1.6 concludes.

## 1.2 Modeling Behavioral Heuristics

In this section, I provide a brief introduction to behavioral heuristics and how they are defined[5]. In particular, I explain how to use heuristics individually and in combination. Ultimately, these heuristics determine how people weight historical observations, which I use in Section 1.4 to estimate weighted linear forecasting models for currencies.

[5]I am grateful to Professor Andrei Semenov, who contributed significantly to this section's development and interpretation of models.

### 1.2.1 Individual Heuristics

In this section, I first consider the rationality assumption. Under what I define as the "rationality" assumption, agents believe past exchange rates are equally important to make investment decisions. In the next step, I consider behavioral heuristics individually (extreme cases) and focus on subjective probabilities instead of objective probabilities. I assign different weights to past historical data of exchange rates. Underlying heuristics determine these weights. I use weights to estimate linear models for changes in exchange rates and construct forecasts. Behavioral heuristics include the availability heuristic, conservatism, pessimism, optimism, overconfidence, underconfidence, anchoring-toward, and anchoring-away.

One well-known heuristic introduced by Tversky and Kahneman (1973) is the availability heuristic. The availability heuristic is a case where people believe the most recent information impacts their decisions. Therefore, they assign higher probabilities to the most recent information as in Semenov (2009).

In contrast to the availability heuristic, conservatism is when people maintain their prior view without explicitly acknowledging new information. They consider their original opinions and the information that formed them to be significant. However, new information learned after the view has been developed is considered less important. This heuristic is related to the mean-reverting strategy in the financial market. In this case, people assign higher probabilities to the past information. Conservatism introduced by Edwards (1968) and used in other studies, including Tversky and Kahneman (1974), Stracca (2004), and Wu, Wu, and Liu (2009).

Pessimism is a case in which people believe that small changes in data/information are more likely to happen in the future. So, they assign higher probabilities to the small changes as in Semenov (2009). By contrast, optimism is a case in which people believe large changes in data/information are more likely to occur in the future. Therefore, they assign higher probabilities to large changes.

In an overconfidence case, people assign higher probabilities to observations closer (have smaller distances) to the mean value of data. In contrast, in an underconfidence case, people assign lower probabilities to observations closer (have smaller distances) to the mean value of data as in Semenov (2009).

Anchoring is a heuristic considered in many situations where people do estimation by starting with an initial value. According to Tversky and Kahneman (1975), this heuristic involves considering a plausible number, the anchor, and shifting it up and down to reach a reasonable result.

To examine the impact of anchor, I set the median of exchange rate changes as the anchor because the median is the best indicator of central tendency in a skewed distribution. Besides, extreme values (outliers) do not affect the median strongly. I introduce two heuristics, anchoring-toward and anchoring-away. Under the assumption of anchoring-toward, people believe that changes in exchange rates close to the median are more likely to happen in the future. Therefore, they assign higher probabilities to changes closer (have smaller distances in absolute value) to the median value of changes in exchange rates for each country. Under the assumption of anchoring-away, people believe exchange rates changes close to the median are less likely to occur in the future. Therefore, they assign lower probabilities to changes closer (have smaller distances in absolute value) to the median value of exchange rates changes.

In the next section, I use mutually exclusive heuristics instead of only one heuristic to construct the subjective probabilities.

### 1.2.2 A Set of Mutually Exclusive Heuristics

In this section, suppose X is an observed variable such as changes in exchange rate ($\Delta s$) and R is the rationality assumption. $\mathrm{BH}_1$ and $\mathrm{BH}_2$ are heuristics that contain contrast beliefs. For instance, $\mathrm{BH}_1$ is optimism, and $\mathrm{BH}_2$ is pessimism. I follow the study by Semenov (2019) and assume investors are divided in three groups and assign a probability that an investor belongs to each given group. First group of investors uses the objective probability, $P((X = \Delta s)|R)$, the second group uses the subjective probability using heuristic $\mathrm{BH}_1$, $P((X = \Delta s)|BH_1)$, the third group uses the subjective probability using heuristic $\mathrm{BH}_2$, $P((X = \Delta s)|BH_2)$. The

R, $BH_1$, and $BH_2$ are mutually exclusive, by the formula of total probability,

$$\begin{aligned}
P(X = \Delta s) &= P((X = \Delta s) \cap R) + P((X = \Delta s) \cap BH_1) + P((X = \Delta s) \cap BH_2) \\
&= P((X = \Delta s)|R)P(R) + P((X = \Delta s)|BH_1)P(BH_1) + \\
&\quad P((X = \Delta s)|BH_2)[1 - P(R) - P(BH_1)], \quad (1.2.1)
\end{aligned}$$

where $\Delta s$ is calculated under the assumption that the level of exchange rate (S) process is a unit root process. There may be cases where more than three groups of mutually exclusive heuristics can be combined. I leave those cases for future research. In this essay, I limit my analysis to the case of a single pair of mutually exclusive heuristics plus rationality.

## 1.3 Data Description

This section describes monthly exchange rates data. Exchange rates are the end-of-month values of the US dollar versus a currency. The data source is IMF's International Financial Statistics (IFS). The data (i.e., changes in exchange rates) start from Feb 1973 to Feb 2020 for all advanced and some emerging/developing countries. For some emerging/developing countries, the start dates are different; for example, the data start from Jan 1999 for Brazil, Jan 1982 for Chile, Jan 1995 for Mexico, Jan 1992 for Peru, Jan 1979 for South Africa. The start dates are different for these countries because they either had pegged (fixed) [6] exchange rate regimes or the volatility of exchange rates was too extreme before the chosen start dates. Countries with small samples or with a fixed exchange rate system are excluded from the analysis.

In the empirical analysis, I use changes in exchange rates denoted by $\Delta s_t$. Therefore, I first take the logarithm of exchange rate price levels (i.e., $s_t = ln(S_t)$, where S is the US dollar price of a unit of foreign currency) and then calculate changes ($\Delta s_t$). I append the

[6] Market forces determine a floating exchange rate. Therefore, the exchange rate goes up and down depending on supply and demand. On the other hand, a pegged (fixed) exchange rate indicates a nominal exchange rate that policymakers restrict for a foreign currency or a basket of foreign currencies.

change of Euro exchange rate from January 1999 to the Euro countries (Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, and Spain) data sets as in Engel et al. (2015).

I divide many countries into advanced and emerging/developing countries using the World Economic Outlook (WEO) classifications. This classification has evolved over time [see more details in Appendix B]. Countries in a given group have similar economic conditions and financial markets, leading to similar investors' behaviors.

The descriptive statistics for changes in nominal exchange rates for advanced countries are reported in Table 1.1. Many changes occurred between 1973 to 2020. For example, some European Union (EU) members' exchange rates were pegged to the European Currency Unit (ECU) before introducing the Euro in 1999. As part of the European Monetary System (EMS), the European Exchange Rate Mechanism (ERM) was introduced in Jan 1999 to control exchange rates fluctuation and achieve monetary stability.

EU has created a single internal market with the free trade of goods, services, capital, and labor. Even though there are significant differences in per capita income among the Euro area countries and in national positions toward inflation, debt, and international trade, these countries have reached a high degree of coordination of monetary and fiscal policies [Washington, DC: Central Intelligence Agency (2019)]. This coordination could also explain similar descriptive statistics, for example, low kurtosis.

Table 1.1 shows for all advanced countries, the positive excess kurtosis. Positive excess kurtosis means the data distribution is not normal, and the distribution has a fat tail. Fat tails indicate a higher probability of extreme positive and negative observations compared to the normal distribution.

Table 1.1: Descriptive statistics for advanced countries

| Countries | Start Date | Last Date | Count | Mean | SD | 25% | Median(50%) | 75% | Min | Max | Skw | Exc-Kurtosis |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Australia** | 1973-02-28 | 2020-02-28 | 565 | -0.0012 | 0.0317 | -0.0163 | 0.0000 | 0.0160 | -0.1916 | 0.1056 | -1.0620 | 5.3699 |
| **Austria** | 1973-02-28 | 2020-02-28 | 565 | 0.0011 | 0.0311 | -0.0165 | 0.0003 | 0.0201 | -0.1213 | 0.1055 | -0.1048 | 1.1853 |
| **Belgium** | 1973-02-28 | 2020-02-28 | 565 | 0.0004 | 0.0314 | -0.0167 | 0.0013 | 0.0187 | -0.1211 | 0.1171 | -0.1572 | 1.3617 |
| **Canada** | 1973-02-28 | 2020-02-28 | 565 | -0.0005 | 0.0197 | -0.0100 | -0.0006 | 0.0094 | -0.1238 | 0.0887 | -0.4398 | 4.6526 |
| **Denmark** | 1973-02-28 | 2020-02-28 | 565 | 0.0000 | 0.0308 | -0.0173 | 0.0000 | 0.0183 | -0.1124 | 0.1105 | -0.0685 | 1.1147 |
| **Finland** | 1973-02-28 | 2020-02-28 | 565 | -0.0004 | 0.0294 | -0.0162 | 0.0008 | 0.0182 | -0.1511 | 0.0894 | -0.5285 | 2.2120 |
| **France** | 1973-02-28 | 2020-02-28 | 565 | -0.0002 | 0.0308 | -0.0169 | 0.0003 | 0.0185 | -0.1164 | 0.1054 | -0.2190 | 1.2388 |
| **Germany** | 1973-02-28 | 2020-02-28 | 565 | 0.0011 | 0.0315 | -0.0167 | 0.0010 | 0.0201 | -0.1217 | 0.1185 | -0.0558 | 1.3585 |
| **Greece** | 1973-02-28 | 2020-02-28 | 565 | -0.0041 | 0.0297 | -0.0185 | -0.0016 | 0.0106 | -0.1742 | 0.1054 | -0.8438 | 4.1745 |
| **Iceland** | 1973-02-28 | 2020-02-28 | 565 | -0.0086 | 0.0389 | -0.0215 | -0.0051 | 0.0108 | -0.2292 | 0.1601 | -1.7619 | 7.8676 |
| **Ireland** | 1973-02-28 | 2020-02-28 | 565 | -0.0009 | 0.0296 | -0.0178 | -0.0005 | 0.0179 | -0.1199 | 0.0896 | -0.2241 | 1.0188 |
| **Italy** | 1973-02-28 | 2020-02-28 | 565 | -0.0019 | 0.0296 | -0.0167 | -0.0002 | 0.0151 | -0.1343 | 0.0894 | -0.4333 | 1.4800 |
| **Japan** | 1973-02-28 | 2020-02-28 | 565 | 0.0018 | 0.0310 | -0.0163 | -0.0003 | 0.0168 | -0.1153 | 0.1501 | 0.3177 | 1.6680 |
| **Luxembourg** | 1973-02-28 | 2020-02-28 | 565 | 0.0004 | 0.0314 | -0.0167 | 0.0013 | 0.0187 | -0.1211 | 0.1171 | -0.1572 | 1.3617 |
| **Netherlands** | 1973-02-28 | 2020-02-28 | 565 | 0.0009 | 0.0314 | -0.0184 | 0.0005 | 0.0198 | -0.1221 | 0.1172 | -0.0864 | 1.2843 |
| **Norway** | 1973-02-28 | 2020-02-28 | 565 | -0.0007 | 0.0305 | -0.0186 | 0.0006 | 0.0165 | -0.1376 | 0.0966 | -0.2373 | 1.3464 |
| **Portugal** | 1973-02-28 | 2020-02-28 | 565 | -0.0034 | 0.0308 | -0.0197 | -0.0016 | 0.0150 | -0.1832 | 0.0894 | -0.9012 | 4.2119 |
| **Singapore** | 1973-02-28 | 2020-02-28 | 565 | 0.0012 | 0.0167 | -0.0062 | 0.0014 | 0.0089 | -0.0770 | 0.0939 | 0.1244 | 4.8727 |
| **Spain** | 1973-02-28 | 2020-02-28 | 565 | -0.0015 | 0.0302 | -0.0160 | -0.0003 | 0.0152 | -0.1991 | 0.0894 | -0.8314 | 4.2913 |
| **Sweden** | 1973-02-28 | 2020-02-28 | 565 | -0.0013 | 0.0317 | -0.0186 | 0.0005 | 0.0177 | -0.1721 | 0.1002 | -0.6630 | 3.2072 |
| **Switzerland** | 1973-02-28 | 2020-02-28 | 565 | 0.0023 | 0.0340 | -0.0184 | 0.0005 | 0.0230 | -0.1547 | 0.1469 | 0.0304 | 1.6027 |
| **United Kingdom** | 1973-02-28 | 2020-02-28 | 565 | -0.0011 | 0.0287 | -0.0170 | -0.0003 | 0.0157 | -0.1277 | 0.1314 | -0.1611 | 2.0116 |

Note: The descriptive statistics for the change of nominal exchange rates are computed based on available data for each country. The column "Countries" presents the names of countries, the columns "Start Date" and "Last Date" show the start date and the end date of the change of exchange rate in the empirical analysis for a given country. The column "Count" indicates the total number of observations for a given country. The columns "Mean" and "SD" present the mean and the standard deviation of total observations, respectively, for a given country. The columns "25%", "Median(50%)", and "75%" show the 25, the 50 (the median), and the 75 percentile of total observations, respectively, for a given country. The columns "Min" and "Max" show the minimum and the maximum values among all observations for a given country. The columns "Skw" indicates the skewness of data and the columns "Exc-Kurtosis" shows the excess kurtosis of data for a given country.

The descriptive statistics for changes in nominal exchange rates for emerging/developing countries are reported in Table 1.2. For all emerging/developing countries, skewness is negative; therefore, the data are skewed left. By skewed left, I mean that the left tail is long relative to the right tail. Based on skewness values in Table 1.2, the data are not symmetric. In addition, Table 1.2 shows that for all emerging/developing countries, the excess kurtosis is large and positive. Positive excess kurtosis means the data distribution is not normal and has a fat tail. There are extreme observations because of exchange rates regime-switching.

Table 1.2: Descriptive statistics for emerging/developing countries

| Countries | Start Date | Last Date | Count | Mean | SD | 25% | Median(50%) | 75% | Min | Max | Skw | Exc-Kurtosis |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Algeria** | 1973-02-28 | 2020-02-28 | 565 | -0.0059 | 0.0286 | -0.0116 | -0.0015 | 0.0062 | -0.3274 | 0.1042 | -4.9358 | 46.9167 |
| **Brazil** | 1999-01-28 | 2020-02-28 | 254 | -0.0051 | 0.0593 | -0.0286 | -0.0017 | 0.0276 | -0.4948 | 0.1816 | -2.7796 | 22.9084 |
| **Chile** | 1982-01-28 | 2020-02-28 | 455 | -0.0061 | 0.0362 | -0.0197 | -0.0053 | 0.0088 | -0.5265 | 0.0699 | -8.0606 | 88.8770 |
| **India** | 1973-02-28 | 2020-02-28 | 565 | -0.0039 | 0.0218 | -0.0126 | -0.0018 | 0.0044 | -0.1950 | 0.0695 | -2.1112 | 17.8450 |
| **Kenya** | 1973-02-28 | 2020-02-28 | 565 | -0.0047 | 0.0296 | -0.0122 | -0.0011 | 0.0057 | -0.2738 | 0.1513 | -2.5170 | 23.0081 |
| **Kuwait** | 1973-02-28 | 2020-02-28 | 556 | 0.0001 | 0.0090 | -0.0034 | 0.0000 | 0.0030 | -0.0469 | 0.1051 | 2.7774 | 36.5901 |
| **Mexico** | 1995-01-28 | 2020-02-28 | 300 | -0.0141 | 0.0571 | -0.0230 | -0.0044 | 0.0031 | -0.6573 | 0.1642 | -6.0672 | 55.1674 |
| **Morocco** | 1973-02-28 | 2020-02-28 | 565 | -0.0013 | 0.0241 | -0.0127 | -0.0001 | 0.0133 | -0.1177 | 0.1041 | -0.5304 | 2.6118 |
| **Niger** | 1973-02-28 | 2020-02-28 | 565 | -0.0015 | 0.0425 | -0.0174 | 0.0001 | 0.0185 | -0.6968 | 0.1041 | -7.8294 | 126.5600 |
| **Peru** | 1992-01-28 | 2020-02-28 | 338 | -0.0037 | 0.0182 | -0.0108 | -0.0007 | 0.0044 | -0.1306 | 0.0579 | -17.1035 | 310.6434 |
| **Philippines** | 1973-02-28 | 2020-02-28 | 565 | -0.0036 | 0.0259 | -0.0075 | -0.0001 | 0.0034 | -0.2513 | 0.0848 | -3.9337 | 31.4585 |
| **Senegal** | 1973-02-28 | 2020-02-28 | 565 | -0.0015 | 0.0425 | -0.0174 | 0.0001 | 0.0185 | -0.6968 | 0.1041 | -7.8294 | 126.5600 |
| **South Africa** | 1979-01-28 | 2020-02-28 | 492 | -0.0057 | 0.0436 | -0.0272 | -0.0024 | 0.0155 | -0.2426 | 0.1413 | -0.8815 | 4.6988 |
| **Tunisia** | 1973-02-28 | 2020-02-28 | 565 | -0.0032 | 0.0248 | -0.0165 | -0.0023 | 0.0113 | -0.1202 | 0.0914 | -0.4690 | 2.0140 |
| **Uruguay** | 1973-02-28 | 2020-02-28 | 565 | -0.0192 | 0.0432 | -0.0336 | -0.0143 | -0.0013 | -0.5804 | 0.2534 | -4.8273 | 60.1116 |

Note: The descriptive statistics for the change of nominal exchange rates are computed based on available data for each country. The column "Countries" presents the names of countries, the columns "Start Date" and "Last Date" show the start date and the end date of the change of exchange rate in the empirical analysis for a given country. The column "Count" indicates the total number of observations for a given country. The columns "Mean" and "SD" present the mean and the standard deviation of total observations, respectively, for a given country. The columns "25%", "Median(50%)", and "75%" show the 25, the 50 (the median), and the 75 percentile of total observations, respectively, for a given country. The columns "Min" and "Max" show the minimum and the maximum values among all observations for a given country. The columns "Skw" indicates the skewness of data and the columns "Exc-Kurtosis" shows the excess kurtosis of data for a given country.

As examples of monthly nominal exchange rates in advanced countries, I plot changes in exchange rates for Canada and Japan in Figures 1.1 and 1.2, respectively, over time from Feb 1973 to Feb 2020. The figures show that the foreign exchange market is volatile and that $\Delta s_t$ moves up and down around a constant level close to 0.

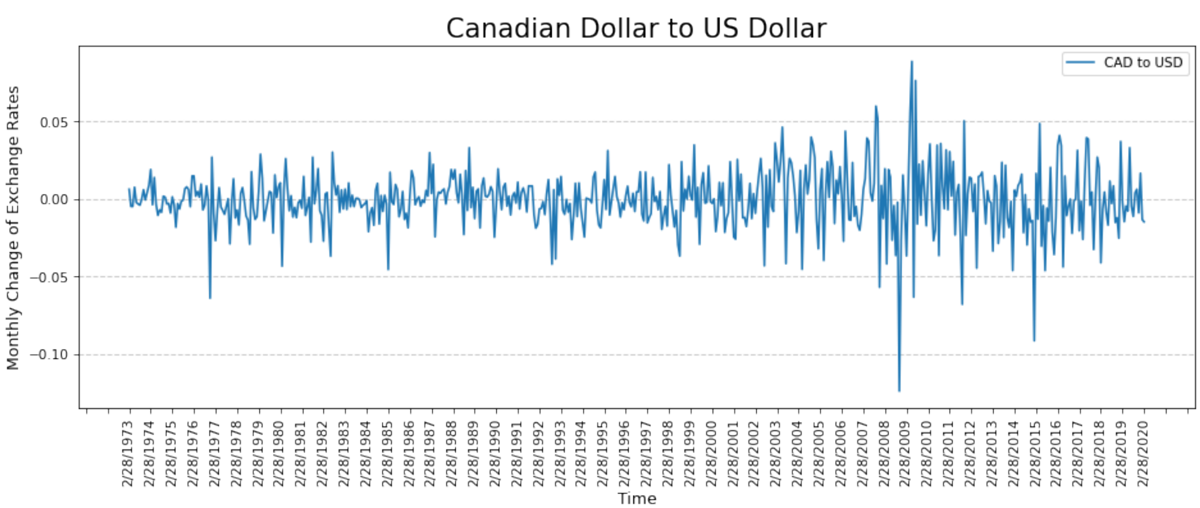


Figure 1.1: Monthly Change of Exchange Rate, CAD to USD. Figure shows the rate of change of exchange rate between the US dollar and the Canadian dollar. To calculate ($\Delta s_t$), I first take the logarithm of exchange rate price levels ($s_t = ln(S_t)$, where S is the US dollars per one unit of Canadian dollar). Data are from Feb 1973 to Feb 2020.

Figure 1.1 indicates the dramatic shifts in $\Delta s_t$ during the 2008-2009 financial crisis. The fluctuation has increased after the financial crisis compared to before the crisis. The US is the leading trading partner with Canada; therefore, the volatility in the US financial market could significantly influence Canada's economy and sustain growth.

Figure 1.2 shows changes in exchange rates for Japan. There were periods of notable shifts in the data from 1980 to 2000. A joint agreement (the Plaza Accord) was signed by countries, including Japan, the United Kingdom, and the US, to depreciate the US dollar against other currencies by the currency market interventions. This agreement increased the value of the Japanese yen versus the US dollar around 1985-1987, which increased speculative bubbles in asset prices. The Japanese economic growth declined at the beginning of the year 1990s. The Bank of Japan tried to minimize speculation and control inflation; therefore, the Bank of Japan increased the inter-bank lending rates at the beginning of the year 1990s. This sudden policy change caused the bubble's bursting and led to the crash of the Japanese stock market. There were considerable shifts around 1995-1996. There was high volatility around the 2008-2009 financial crisis.

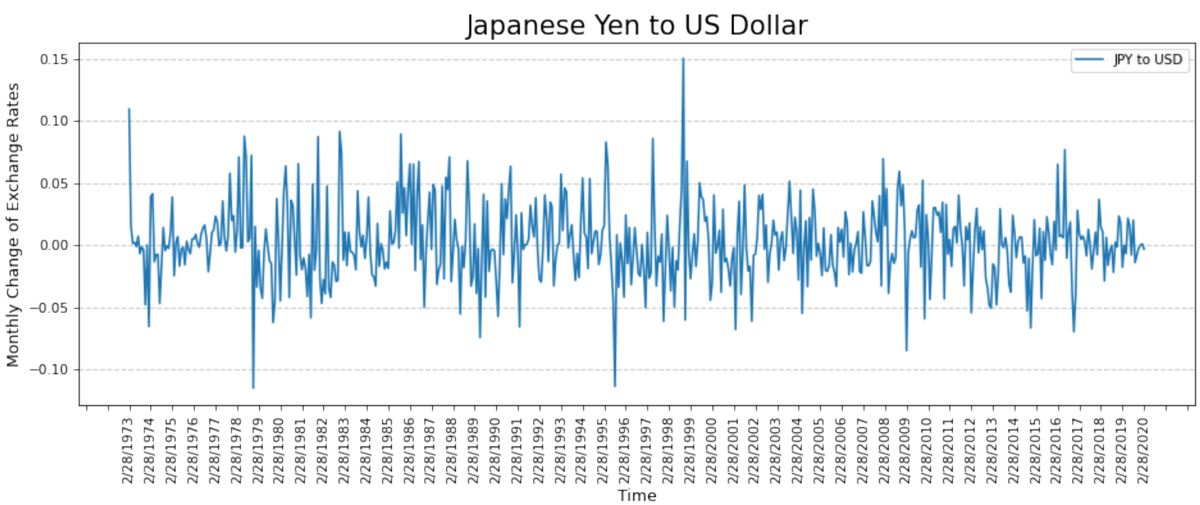


Figure 1.2: Monthly Change of Exchange Rate, JPY to USD. Figure shows the rate of change of exchange rate between the US dollar and the Japanese Yen. To calculate ($\Delta s_t$), I first take the logarithm of exchange rate price levels ($s_t = ln(S_t)$, where S is the US dollars per one unit of Japanese YEN). Data are from Feb 1973 to Feb 2020.

As examples of monthly nominal exchange rates in emerging/developing countries, I present changes in exchange rates for South Africa and Mexico. Figure 1.3 shows changes in exchange rates for South Africa over time period Jan 1979 to Feb 2020 and indicates that changes in exchange rates shifted down at the beginning of 1980s for few years. There was significant volatility around 1984-1988, also around 2008-2009. High inflation and political turmoil significantly impacted the trading relationship between South Africa and the rest of the World, particularly the US. These issues influenced the foreign exchange market as well.

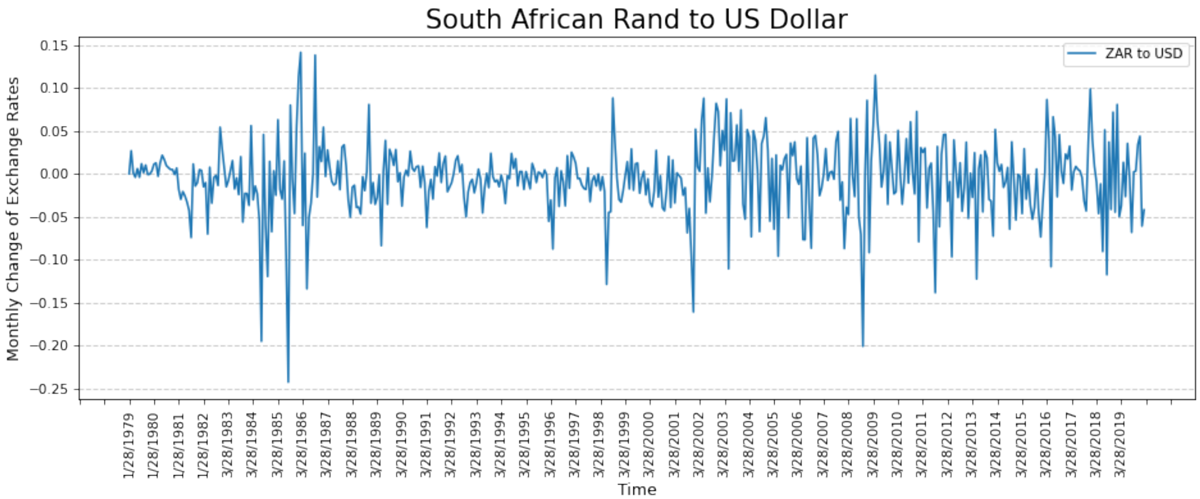


Figure 1.3: Monthly Change of Exchange Rate, ZAR to USD. Figure shows the rate of change of exchange rate between the US dollar and the South African Rand. To calculate ($\Delta s_t$), I first take the logarithm of exchange rate price levels ($s_t = ln(S_t)$, where S is the US dollars per one unit of South African Rand). Data are from Jan 1979 to Feb 2020.

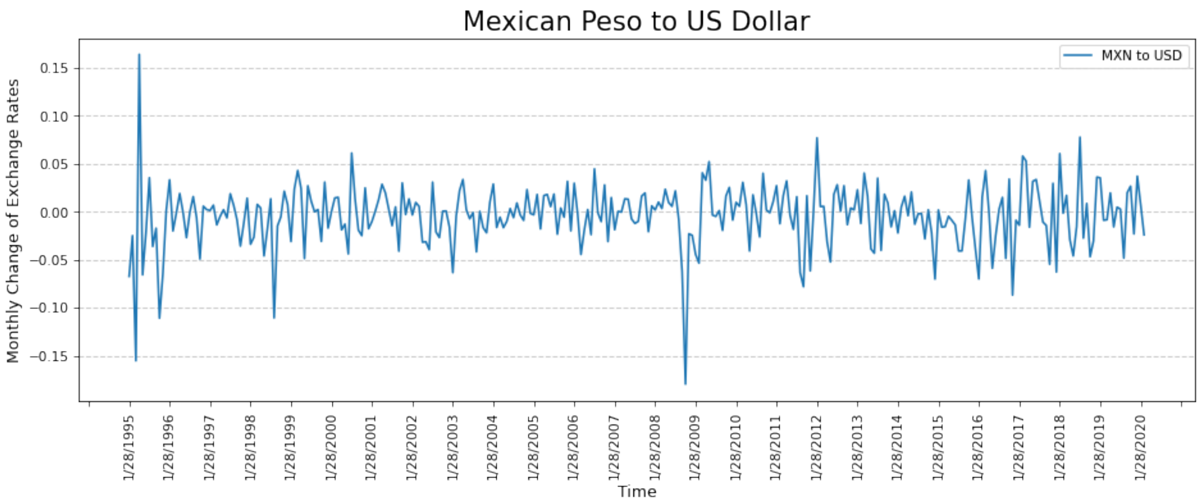


Figure 1.4: Monthly Change of Exchange Rate, MXN to USD. Figure shows the rate of change of exchange rate between the US dollar and the Mexican Peso. To calculate ($\Delta s_t$), I first take the logarithm of exchange rate price levels ($s_t = ln(S_t)$, where S is the US dollars per one unit of Mexican Peso). Data are from Jan 1995 to Feb 2020.

Figure 1.4 shows changes in exchange rates for Mexico between Jan 1995 to Feb 2020. Mexico is an emerging country with a high volume of foreign trade. There was significant volatility around 1995-1996. A reason could be that in Dec 1994, the Mexican government devalued the peso against the US dollar, which led to the financial crisis; it is called the "Mexican peso crisis." The was significant shift around 2008-2009. During the 2008-2009 global financial crisis, Mexico's economy was hit hard, mainly because it depended significantly on the US market. Mexico's gross domestic product decreased substantially in 2009. The US is the biggest trading partner with Mexico; therefore, fluctuations in the US market would broadly impact Mexico's economy and growth.

In the next section, I explain the methodology and its implementation in this essay.

## 1.4 Methodology and Implementation

In this section, I present methodologies to construct variables and models determination and estimations[7]. In particular, I describe the following:

[7] I am grateful to Professor Andrei Semenov, who contributed significantly to this section's development and interpretation of models.

- Construction of Monthly Changes in Exchange Rates
- Model Determination and Estimation
  - Analysis of Autocorrelation Coefficients for Changes in Exchange Rates
  - Determine a Linear Representation for Changes in Exchange Rates
  - Construction of Probabilities for Realized Changes in Exchange Rates
  - Extension to Mutually Exclusive Heuristics Sets
  - Determine Best Combination of $\lambda$/Heuristics/Probabilities
  - Estimation Setup and Forecasting
- Performance Measures and Statistics

### 1.4.1 Construction of Monthly Changes in Exchange Rates

I collect monthly data on exchange rates (price levels). I take the logarithm of price level for a given currency pair and denote it by $s_t$. The $h$-month-ahead rate of changes in exchange rates is calculated as $\Delta s_{t+h,h} = s_{t+h} - s_t$. Note that S is expressed as the log of US dollars for a unit of foreign currency. Several studies have used this definition of exchange rates [see, e.g., Molodtsova and Papell (2009), Ince (2014), and Alba, Park, and Xie (2015)].

### 1.4.2 Model Determination and Estimation

#### 1.4.2.1 Analysis of Autocorrelation Coefficients for Changes in Exchange Rates

Figures 1.1 to 1.4 indicate that $\Delta s_t$ do not display a global trend and are characterized by time-varying volatility (conditional heteroskedasticity). For a given currency pair, I calculate Pearson autocorrelation coefficients to check the number of lags that could be viewed as explanatory variables. These coefficients measure the strength of a linear relationship between the change of exchange rate and its historical data. I use the first original 120 observations to calculate partial autocorrelation (PACF) coefficients, $\rho(\Delta s_t, \Delta s_{t-k})$, where $k = 1, 2, 3$.

This step is taken only once. I choose to check $\rho(1), \rho(2), \rho(3)$ because the frequency of data is monthly and this choice covers three months. I calculate a test statistic and p-value to examine whether each autocorrelation coefficient is statistically significant. The significance level is 10%.

For a given country, I set the maximum number of lags ($p$) at which the partial autocorrelation coefficient is statistically significant as an order of an autoregressive, $AR(p)$, model. An autoregressive model relies on an assumption that observations at previous time steps are useful to forecast a value at the next time step.

#### 1.4.2.2 Determine a Linear Representation for Changes in Exchange Rates

In the previous section, I explained how an order of an autoregressive model is selected using the first 120 original observations. For a given country, I assume that $\Delta s_t$ is locally stationary and use the $AR(p)$ model with $p$ determined in Section 1.4.2.1:

$$\Delta s_{t+1,1} = \beta_0 + \sum_{i=1}^{p} \beta_i \Delta s_{t-i+1} + u_{t+1}, \tag{1.4.1}$$

where $E[u_{t+1}] = 0$, $Var[u_{t+1}] = \sigma^2 < \infty$, $E[u_{t+1}|\Delta s_{t-k}] = 0$ for all $0 \leq k \leq p$, and the autoregressive polynomial has roots outside the unit circle and the sum of autoregressive coefficients $\beta_i$ is not equal to 1.

To estimate coefficients in equation (1.4.1), I use the first original 120 observations and the generalized method of moments (GMM) method. The GMM estimation method does not require complete knowledge of the distribution of data. Only specified moments derived from an underlying model are needed for the GMM estimation.

In particular, I use the following GMM moments:

$$g_{T_1}(\theta) = \frac{1}{m} \sum_{t=p}^{T_1-1} Z_t u_{t+1}, \tag{1.4.2}$$

where $m = T_1 - p$, $T_1$ is the number of observations in the estimation period, $p$ is the number of lags in the $AR(p)$ model, $u_{t+1}$ is the error term given in equation (1.4.1), and $Z_t$ is a vector of instruments defined by

$$Z_t = [1, \Delta s_t, \ldots, \Delta s_{t-p+1}]\prime. \tag{1.4.3}$$

The GMM estimator of $\beta$ is a vector $\hat{\beta}$ that solves the problem

$$min\ [g_{T_1}(\theta)]'\ W^*\ [g_{T_1}(\theta)], \tag{1.4.4}$$

where $W^*$ is the optimal matrix which produces asymptotically efficient estimators and is defined as

$$W^* \equiv \Lambda^{-1}(\theta) = [Var(\sqrt{m} g_{T_1}(\theta))].^{-1} \tag{1.4.5}$$

To determine whether the linear representation for the change of exchange rate in equation (1.4.1) is adequate, I examine whether the $s_t$ process follows a random walk using the variance ratio test. This step is taken only once for the given country. The Lo and MacKinlay (1988) variance ratio test of random walk allows for the general form of conditional heteroskedasticity. $VR(q)$ is a statistic that under the null of random walk equals one. I consider the null hypothesis that all variance ratio statistics are jointly equal to one. I compute variance ratio statistics $VR(q)$ for $2 \leq q \leq p+1$ following Campbell, W.Lo, and MacKinlay (1997). The variance ratio statistic is defined as

$$VR(q) = 1 + \sum_{k=1}^{q-1} 2\left(1 - \frac{k}{q}\right)\hat{\beta}_k, \tag{1.4.6}$$

where the GMM estimators are denoted by $\hat{\beta}_k$.

I use the jointly estimated autoregressive coefficients ($\hat{\beta}_1$, $\hat{\beta}_2$, ..., $\hat{\beta}_k$) from equation (1.4.1) in equation (1.4.6) to calculate $VR(2), \ldots,\ VR(p{+}1)$, respectively. I use the sample variance-covariance matrix, $\hat{\Omega}$, of the autoregressive coefficients, $\hat{\beta}_k$, from the GMM estimation to calculate the variance of $VR(q)$:

$$\hat{\Omega} = \begin{pmatrix} var(\hat{\beta}_1) & \cdots & cov(\hat{\beta}_1, \hat{\beta}_{q-1}) \\ \vdots & & \vdots \\ cov(\hat{\beta}_{q-1}, \hat{\beta}_1) & \cdots & var(\hat{\beta}_{q-1}) \end{pmatrix}. \tag{1.4.7}$$

The variance of $VR(q)$ is calculated as follows

$$\hat{\sigma}^2_{VR}(q) = C \times \hat{\Omega}_{\beta_1,\ \ldots,\ \beta_{q-1}} \times C\prime, \tag{1.4.8}$$

where $C = (2[1 - \frac{1}{q}] \ \ \ldots \ \ 2[1 - \frac{k=q-1}{q}])$ and $C'$ is the transpose of C, i.e., $C' = (2[1 - \frac{1}{q}] \ \ \ldots \ \ 2[1 - \frac{k=q-1}{q}])'$.
Then the standardized test statistic is given by

$$\Psi(q) = \frac{VR(q) - 1}{\sqrt{\sigma^2_{VR}(q)}}, \tag{1.4.9}$$

and estimated by replacing the $VR(q)$ and $\sqrt{\sigma^2_{VR}(q)}$ by their estimators. Because the random walk hypothesis requires the variance ratios for all selected $q$ must be equal to one, a common method to test the null hypothesis is the multiple comparison of all selected variance ratios statistics with one (Chow and Denning (1993)). To conduct the multiple comparison, I proceed as in Chow and Denning (1993) and Fong, Koh, and Ouliaris (1997) where the authors show that

$$Pr\Big(max[|\Psi(q_1)|, \ldots, |\Psi(q_m)|] \leq SMM(\alpha; m; T_1)\Big) \geq (1 - \alpha), \tag{1.4.10}$$

where $SMM(\alpha; m; T_1)$ is the upper $\alpha$ point of the Studentized Maximum Modulus (SMM) distribution with parameter $m$ and $T_1$ (sample size) degrees of freedom [see Appendix A for more details].

Hence, if the maximum absolute value of $\Psi$ is greater than the SMM critical value at the significance level of 10%, then I reject the random walk hypothesis for the $s_t$ process; therefore, I reject the White Noise for the $\Delta s_t$ process and specify $\Delta s_t$ as $AR(p)$ process, equation (1.4.1). If the intercept ($\beta_0$) is statistically significant as well, I select the $AR(p)$ with a constant for the $\Delta s_t$ series. If the intercept is not statistically significant, $AR(p)$ without an intercept is selected.

The maximum absolute value of $\Psi$ less than the SMM critical value means the random walk hypothesis cannot be rejected for the $s_t$ process. If $s_t$ is RW, then $\Delta s_t$ is White Noise (WN) either with $constant = 0$ or $constant \neq 0$. To examine whether the model is WN with a constant, I estimate the intercept in equation (1.4.1) (when $p = 0$) using the ordinary least squares (OLS) and HAC approaches to correct the standard errors for autocorrelation and conditional heteroskedasticity. If the intercept is statistically significant, WN with a constant is selected. Otherwise, WN is selected for the given country.

I should emphasize that the chosen model in this section for the $\Delta s_t$ process (i.e., $AR(p)$ or WN with a constant or WN without a constant) is used for both forecasting horizons (1-month-ahead and 12-months-ahead) for the given country.

#### 1.4.2.3 Construction of Probabilities for Realized Changes in Exchange Rates

Probabilities for the historical change of exchange rates are constructed under behavioral heuristic assumptions. I assume all investors use the same heuristic to rank changes in exchange rates ($\Delta s_t$). Behavioral heuristics assume that investors assign probabilities to realized changes in exchange rates. To model this, given a time $T_1$ and behavioral heuristic $BH$, I take the following steps:

- Step 1: I order $\Delta s_t$ up to time $T_1$ based on the heuristic BH's specific definition (as described in Section 1.2.1).
- Step 2: I provide weights exponentially.

Following studies by Boudoukh, Richardson, and Whitelaw (1998) and Semenov (2008), I use the function:

$$P(X = \Delta s_t | BH) = \frac{\lambda^{i-1}(1-\lambda)}{1-\lambda^{T_1}}, \tag{1.4.11}$$

where $i$ shows that $\Delta s_t$ is the $i$th item in the ordered list of $\Delta s_t$, $\lambda$ is a decay factor, $T_1$ is the number of observations in the estimation period, and

$$\sum_{t=1}^{T_1} P(X = \Delta s_t | BH) = 1, \tag{1.4.12}$$

where $T_1$ is the number of observations in the estimation period, and the sum of probabilities is one. I choose this exponentially declining function to emphasize the fact that under behavioral assumptions, past exchange rates are not equally important. Some past exchange rates are more important, and the rest of exchange rates quickly become less important. This weighting methodology is based on a decay factor, $0 < \lambda < 1$, which determines the rate at which some $\Delta s_t$ are discounted based on the heuristic assumption. A lower $\lambda$ means the considered heuristic is more important. I use $\lambda = 0.95, 0.96, 0.97, 0.98, 0.99$. If I use very low ($\lambda$ below 0.95), there would be too few observations to use for estimations because weights would sharply decline toward zero. The few observations remaining in the data set would lead to unreliable results. A small sample would result in a large margin of error and a wide confidence interval. A wide confidence interval leads to less confidence in estimated parameters and less certainty about prediction outcomes.

Note that rationality is a special case of the above in the limit of $\lambda$ approaching to one where the probability will be

$$P(X = \Delta s_{t+h,h} | R) = \frac{1}{T_1}, \tag{1.4.13}$$

where $R$ denotes the rationality assumption and $T_1$ is the number of observations in the estimation period. Changes in exchange rates have equal probabilities under rationality.

#### 1.4.2.4 Extension to Mutually Exclusive Heuristics Sets

As the definitions in Section 1.2 show, for each heuristic, there is an opposite (mutually exclusive) heuristic. The example of such pairs of heuristics are: anchoring-toward and anchoring-away, availability and conservatism, pessimism and optimism, etc.

I consider the following sets of mutually exclusive heuristics: (1) overconfidence, underconfidence, and rationality; (2) availability, conservatism, and rationality; (3) optimism, pessimism, and rationality; (4) anchoring-toward, anchoring-away, and rationality. Since heuristics are mutually exclusive, each realized change of exchange rate has a probability that is given by (as derived in Section 1.2.2):

$$
\begin{aligned}
P(X=\Delta s) &= P((X=\Delta s)\cap R)+P((X=\Delta s)\cap BH_1)+P((X=\Delta s)\cap BH_2)\\
&= P((X=\Delta s)|R)P(R)+P((X=\Delta s)|BH_1)P(BH_1)+\\
&\qquad P((X=\Delta s)|BH_2)[1-P(R)-P(BH_1)]. \quad (1.4.14)
\end{aligned}
$$

where X is a random variable, $\Delta s$ is the change of exchange rate, R is the rationality assumption, $BH_1$ and $BH_2$ are heuristics that have opposite assumptions. $\Delta s$ is calculated under the assumption that the level of exchange rate (S) process is a unit root process. To calculate the above total subjective probability, $P(R)$, $P(BH_1)$, and $P(BH_2)$ are required. I assume these variables can take discrete values with 10% step such that they are in $[0, 0.1, 0.2, ..., 1]$. For instance, one combination is $P(R) = 0.1$, $P(BH_1) = 0.1$, and $P(BH_2) = 0.8$; another combination is $P(R) = 0.2$, $P(BH_1) = 0.3$, and $P(BH_2) = 0.5$; and so on. There are 66 different combinations in total. The choice of 10% could be a contested choice, to which the results could be sensitive.

Recall that the rest of quantities in equation (1.4.14) are calculated as described in Section 1.4.2.3. In particular, recall that under rationality, $P((X = \Delta s)|R) = \frac{1}{T_1}$, where $T_1$ is the number of observation in the estimation period (as described in Section 1.2). For example, if $T_1 = 120$, $P((X = \Delta s)|R) = \frac{1}{120}$. To determine $P((X = \Delta s)|BH_1)$ and $P((X = \Delta s)|BH_2)$, I use equation (1.4.11) in Section 1.4.2.3.

#### 1.4.2.5 Determine Best Combination of $\lambda$/Heuristics/Probabilities

For a given country and forecast horizon, I use the first $T_1 = 120$ observations to estimate the linear representation (model) selected in Section 1.4.2.2 under different combinations of $\lambda$, heuristic sets, and heuristic probabilities. There are 1320 ($5\lambda \times 4$ heuristic sets $\times$ 66 probability combinations) different combinations. The AIC (Akaike information criterion) and BIC (Bayesian information criterion) statistics are valid under the stationarity assumption, which is satisfied in all processes $\Delta s_t$.

Using the AIC (Akaike information criterion) and BIC (Bayesian information criterion) statistics, I evaluate the selected model's in-sample performance under all 1320 combinations.

The combination of $\lambda$, heuristic set, and heuristic probabilities, which leads to the smallest AIC and BIC, is the 'best' combination for the given country and horizon.

For a given country and horizon, the selected model in Section 1.4.2.2 under the 'best' combination of $\lambda$, heuristic set, and heuristic probabilities chosen in this section denotes a forecasting model.

#### 1.4.2.6 Estimation Setup and Forecasting

To forecast the future movements of exchange rates for the US dollar against each currency. I follow the steps given below. Suppose the historical data set consists a total number of $T$ observations, and let $h$ denotes the forecasting horizon measured in months. I consider both short ($h = 1$) and long ($h = 12$) horizons.

I estimate the chosen forecasting model using a conventional approach in the exchange rate forecasting literature, i.e., a fixed-length rolling window approach [see, e.g., Clark and West (2006), Clark and West (2007), Molodtsova and Papell (2009), and Rossi (2013)]. The rationale for using the fixed-length rolling window approach is the sample period is very long with many regime changes. Using the rolling accommodates potential changes in the parameters values over time. Figures 1.1 to 1.4 show that $\Delta s_t$ may locally shift up and down around 0 due to sustained increases in the exchange rates.

The total sampling periods for most countries are from Feb 1973 to Feb 2020. Using the fixed-length rolling window approach and starting from Feb 1973, I take the window of $T_1 = 120$ weighted observations to estimate the chosen forecasting model and use the estimated coefficients to form a forecast for the change of exchange rates. The 120 observations are close to 20% of the total observations. Then, I roll the sample period forward one observation, re-estimate the model, and form a forecast using the estimated coefficients. I keep rolling and repeating the steps above until the end of the sample period[8].

If a chosen forecast model is $AR(p)$ with a constant, $h$-month-ahead forecast for changes

[8] Alternatively, a recursive (expanding) window approach can be used for estimation. The recursive method makes use of an increasing window to re-estimate coefficients, whereas the rolling approach makes use of a fixed-length window of data to re-estimate coefficients.

in exchange rates is as follows:

$$E_t(\Delta s_{t+h,h}) = \tilde{\hat{\beta}}_{0t} + \sum_{i=1}^{p} \tilde{\hat{\beta}}_{it} \Delta s_{t-i+1}, \tag{1.4.15}$$

where $h = 1$ (the monthly forecasts) or $h = 12$ (the annual forecasts). $\tilde{\hat{\beta}}_{0t}$ and $\tilde{\hat{\beta}}_{it}$ are estimated using weighted observations. I use the "$\tilde{}$" notation to emphasize this process. Weights come from heuristics assumptions.

If a chosen forecast model is $AR(p)$ without a constant, $h$-month-ahead forecast for changes in exchange rates is as follows:

$$E_t(\Delta s_{t+h,h}) = \sum_{i=1}^{p} \tilde{\hat{\beta}}_{it} \Delta s_{t-i+1}. \tag{1.4.16}$$

If a chosen forecast model is White Noise with a constant, $h$-month-ahead forecast for changes in exchange rates is as follows:

$$E_t(\Delta s_{t+h,h}) = \tilde{\hat{\beta}}_{0t}, \tag{1.4.17}$$

where $\tilde{\hat{\beta}}_{0t}$ is the weighted mean of dependent variable at time $t$.

If a chosen forecast model is White Noise, h-period-ahead forecast for exchange rates is as follows:

$$E_t(\Delta s_{t+h,h}) = 0. \tag{1.4.18}$$

Depending on $h$, the total number of forecasts ($T_2$) will vary according to data set sizes. More specifically, $T_2 = T - T_1$, where $h = 1$ and $T_2 = T - T_1 - 12$, where $h = 12$.

### 1.4.3 Performance Measures and Statistics

In this section, I use some well-known statistics to compare the out-of-sample forecasting performance of behavioral models with the benchmark model. I use the Mean Squared Prediction Error (MSPE) statistic to measure the forecasting accuracy and, in particular, the ratio of MSPE for a behavioral model over MSPE of the benchmark.

To have a view on statistically important outperformance, I use the Diebold-Mariano (DM) statistic [9].

One reason for using the White Noise (WN) model as the benchmark model is that this model follows from the random walk representation of the $S_t$ process common in the financial literature and used for the forecasting model. The forecast is one of "no change" in the exchange rate. It implies that past information does not help to predict the future movements of exchange rates.

To define the $MSPE_{ratio}$ statistic, first recall that $T$ is the total sample length, $T_1$ is the (initial) sample reserved for estimation, and $h$ denotes a horizon, hence $T_2 - h + 1$ is the number of forecasts. The sample MSPE of the benchmark and of the behavioral model for the $h$-month-ahead forecast is then calculated as

$$Benchmark: \quad MSPE_{WN} = \frac{1}{T_2 - h + 1} \sum_{t=T_1}^{T-h} \hat{e}^2_{(WN,t+h)}, \tag{1.4.19}$$

where $\hat{e}^2_{(WN,t+h)} = (\Delta s_{t+h,h})^2$ and $\Delta s_{t+h,h}$ are realized values.

$$Model: \quad MSPE_{M} = \frac{1}{T_2 - h + 1} \sum_{t=T_1}^{T-h} \hat{e}^2_{(M,t+h)}, \tag{1.4.20}$$

where $\hat{e}^2_{(M,t+h)} = (\Delta s_{t+h,h} - \Delta \hat{s}_{t+h,h})^2$, $\Delta s_{t+h,h}$ are realized values, and $\Delta \hat{s}_{t+h,h}$ are corresponding forecasts from the underlying behavioral model. I calculate $MSPE_{ratio}$ as follows

$$MSPE_{ratio} = \frac{MSPE_M}{MSPE_{WN}}. \tag{1.4.21}$$

$MSPE_{ratio} < 1$ implies that the behavioral model provides better forecasting performance compared to the benchmark model.

Similar to the $MSPE_{ratio}$ statistic, the DM statistic examines the null $H_0 : E(d_t) =$

[9] I should mention that there are other measures of forecasting performances that are used in other studies. Mean absolute percentage error (MAPE) is the mean of the sum of all of the percentage errors for a given data set taken without regard to sign. Mean absolute scaled error (MASE) is defined as the mean absolute error of a model divided by the mean absolute error of a naive random walk without drift model.

$E(e^2_{WN,t} - e^2_{M,t}) = 0$ against $H_1 : E(d_t) = E(e^2_{WN,t} - e^2_{M,t}) > 0$. Define $\hat{d}_{t+h}$ as

$$\hat{d}_{t+h} = \hat{e}^2_{WN,t+h} - \hat{e}^2_{M,t+h}, \tag{1.4.22}$$

where $\hat{e}_t$ are errors driven by forecasts at the given time $t$. The DM statistic can be calculated by regression of the loss differential ($\hat{d}_{t+h}$) on a constant using HAC approaches to correct the standard errors for autocorrelation and heteroskedasticity. I calculate the p-value associated with the DM statistic and use the 10% level of significance, which is a common practice in the literature and previous studies [see, e.g., Ince (2014), Engel et al. (2015), and Cheung et al. (2019)].

As an addition to the above, I use the Pesaran-Timmermann (PT) statistic as in Cerra and Saxena (2010), Kouwenberg et al. (2017), and Cheung et al. (2019) studies. For policy-makers and practitioners in financial markets, the ability to accurately forecast the direction of change of exchange rates is almost as useful as an accurate point forecast. Therefore, I calculate the PT test-statistic, which examines a behavioral model's ability to forecast the direction of change correctly, relative to the random walk. The PT test-statistic is as follows:

$$PT = \frac{\hat{p} - p^*}{\sqrt{p^*(1-p^*)/N}}, \tag{1.4.23}$$

where $p^* = 0.5$ is the White Noise (WN) proportion of correct direction forecasts, $\hat{p}$ is a behavioral model's proportion of correct direction, and $N$ is the total number of forecasts. If $\hat{p} = p^* = 0.5$ then the exchange rate' expected change is zero ($E_t(\Delta s_t) = 0$); hence $S_t$ is just as likely to rise as it is to fall. There is the 50% chance that the direction (sign) of forecast is correct. The PT statistic is asymptotically N(0,1) distributed. I calculate the p-value associated with the PT statistic and use the 10% level of significance. A value of $\hat{p}$ statistically larger than 0.5 indicates a better forecasting performance than the WN model.

In addition to the statistical evaluation discussed above, I examine behavioral models' economic value. I follow a simple investment strategy that buys (sells) the US dollar versus one unit of a foreign currency when a behavioral model forecasts an appreciation (depreciation) of foreign currency. I calculate returns of investment strategies as follows:

$$r_t = \frac{E_t(\Delta s_t)}{|E_t(\Delta s_t)|}\Delta s_t, \tag{1.4.24}$$

where $E_t(\Delta s_t) > 0$. I regress $r_t$ on a constant and use the HAC approach to correct the standard errors for autocorrelation and heteroskedasticity. Using the estimated constant and an adjusted standard error, I calculate an annualized Sharpe-ratio (Sharpe (1966)). I also report the annualized average return in percentage for each behavioral model for a given country and horizon.

## 1.5 Empirical Results

This section describes the out-of-sample forecasting performance of behavioral models from both statistical and economic perspectives. I particularly examine whether behavioral models can outperform the WN model to predict the direction of change of exchange rates. I report empirical results for forecasting models at both h=1 and h=12 for a given country. Tables 1.3 and 1.4 present the chosen forecasting models and the best combination of $\lambda$, heuristic sets, and probabilities for advanced and emerging/developing countries, respectively.

Table 1.3: Chosen forecasting models for advanced countries

| | | Horizon=1 | | Horizon=12 | |
|---|---|---|---|---|---|
| **Countries** | **Selected Models** | $\lambda$ | **The Best Heuristic Set with Probabilities** | $\lambda$ | **The Best Heuristic Set with Probabilities** |
| **Australia** | WN | - | - | - | - |
| **Austria** | WN | - | - | - | - |
| **Belgium** | WN | - | - | - | - |
| **Canada** | WN with a constant | 0.95 | R(prob=0)/O-C(prob=1)/U-C(prob=0) | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) |
| **Denmark** | WN | - | - | - | - |
| **Finland** | WN | - | - | - | - |
| **France** | WN | - | - | - | - |
| **Germany** | WN | - | - | - | - |
| **Greece** | WN with a constant | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) | 0.95 | R(prob=0)/Opt(prob=1)/Psm(prob=0) |
| **Iceland** | WN with a constant | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) |
| **Ireland** | WN with a constant | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) |
| **Italy** | WN with a constant | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) |
| **Japan** | WN | - | - | - | - |
| **Luxembourg** | WN | - | - | - | - |
| **Netherlands** | WN | - | - | - | - |
| **Norway** | WN | - | - | - | - |
| **Portugal** | WN with a constant | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) |
| **Singapore** | WN with a constant | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) |
| **Spain** | WN with a constant | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) |
| **Sweden** | WN with a constant | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) |
| **Switzerland** | WN with a constant | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) |
| **United Kingdom** | WN with a constant | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) |

Note: The column "Countries" presents the names of countries. The column "Selected Models" indicates the selected model based on the variance-ratio test for each country. The column "The Best Heuristic Set with Probabilities" indicates the best combination of heuristics based on the in-sample evaluation of selected models at the year 10th under the 1320 different combinations (5 lambdas * 66 probability combinations * 4 heuristic sets) for a given country and horizon. "WN" stands for the White Noise model. "R" denotes the rationality assumption, "A-T" denotes the anchoring-toward heuristic, and "A-A" denotes the anchoring-away heuristic. "O-C" and "U-C" represent the overconfidence and underconfidence heuristics, respectively. "Opt" and "Psm" suggest the optimism and pessimism heuristics, respectively. "$prob = 0$" means the probability of a given heuristic is zero, and "$prob = 1$" means the probability of a given heuristic is one. **" - "** means the selected model is WN; therefore, $E_t(\Delta s_{t+h,h}) = 0$ and no heuristic and $\lambda$ are used.

Table 1.3 shows that for 50% of advanced countries, the forecasting model is White Noise (WN) with a time-varying constant [see equation (1.4.17)]. The value of $\lambda = 0.95$ is common for all countries, and the anchoring-toward heuristic is the best for almost all countries. Recall that under the anchoring-toward assumption, people believe that changes in exchange rates close to the median of historical data is more likely to happen in the future. The results at both short and long horizons are similar. For the remaining advanced countries, the best forecasting model is the WN with mean 0. This implies that the best forecasting model of $S_t$ in these countries is a random walk model.

Table 1.4: Chosen forecasting models for emerging/developing countries

| | | Horizon=1 | | Horizon=12 | |
|---|---|---|---|---|---|
| **Countries** | **Selected Models** | $\lambda$ | **The Best Heuristic Set with Probabilities** | $\lambda$ | **The Best Heuristic Set with Probabilities** |
| **Algeria** | WN | - | - | - | - |
| **Brazil** | WN with a constant | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) |
| **Chile** | AR(3) with a constant | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) | 0.95 | R(prob=0)/Opt(prob=1)/Psm(prob=0) |
| **India** | WN | - | - | - | - |
| **Kenya** | WN with a constant | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) |
| **Kuwait** | WN with a constant | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) |
| **Mexico** | WN with a constant | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) | 0.95 | R(prob=0)/Opt(prob=1)/Psm(prob=0) |
| **Morocco** | WN | - | - | - | - |
| **Niger** | WN | - | - | - | - |
| **Peru** | AR(1) with a constant | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) |
| **Philippines** | WN with a constant | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) | 0.95 | R(prob=0)/Opt(prob=1)/Psm(prob=0) |
| **Senegal** | WN | - | - | - | - |
| **South Africa** | AR(3) without a constant | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) |
| **Tunisia** | AR(1) without a constant | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) |
| **Uruguay** | WN with a constant | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) | 0.95 | R(prob=0)/A-T(prob=1)/A-A(prob=0) |

Note: The column "Countries" presents the names of countries. The column "Selected Models" indicates the selected model based on the variance-ratio test for each country. The column "The Best Heuristic Set with Probabilities" indicates the best combination of heuristics based on the in-sample evaluation of selected models at the year 10th under the 1320 different combinations (5 lambdas * 66 probability combinations * 4 heuristic sets) for a given country and horizon. "WN" stands for the White Noise model. "R" denotes the rationality assumption, "A-T" denotes the anchoring-toward heuristic, and "A-A" denotes the anchoring-away heuristic. "O-C" and "U-C" represent the overconfidence and underconfidence heuristics, respectively. "Opt" and "Psm" suggest the optimism and pessimism heuristics, respectively. "$prob = 0$" means the probability of a given heuristic is zero, and "$prob = 1$" means the probability of a given heuristic is one. **" - "** means the selected model is WN; therefore, $E_t(\Delta s_{t+h,h}) = 0$ and no heuristic and $\lambda$ are used.

Table 1.4 shows that for 40% of emerging/developing countries, the forecasting model is the WN with a time-varying constant [see equation (1.4.17)]. In addition, for 27% of emerging/developing countries, the forecasting model is the $AR(p)$ with or without a time-varying constant [see equation (1.4.15) and equation (1.4.16)]. For remaining emerging/developing countries, the best forecasting model is the WN model with mean 0. This implies that the best forecasting model of $S_t$ is the random walk model.

Table 1.4 shows that the value of $\lambda = 0.95$ is common for all countries and the anchoring-toward heuristic is the best heuristic for almost all countries. The results in both short and long horizons are similar for most countries. The optimism heuristic is the best for Chile, Mexico, and Philippines in the 12-months-ahead forecasts. As a reminder, under the optimism assumption, people believe that large changes in exchange rates are more likely to happen in the future.

Table 1.5: Empirical results for advanced countries

| | Countries | MSPE_ratio | DM (t-stats) | DM (p-value) | $\hat{p}$ | PT (t-stats) | PT (p-value) | Avg. Return (%) | Sharpe_ratio |
|---|---|---|---|---|---|---|---|---|---|
| Horizon=1 | Canada | 1.0186 | -1.9501 | 0.0259 | 0.5034 | 0.1422 | 0.4435 | -0.8831 | -0.1275 |
| | Greece | 0.9903 | 0.7805 | 0.2177 | 0.5169 | 0.7110 | 0.2385 | 2.5225 | 0.2372 |
| | Iceland | 0.9810 | 0.8653 | 0.1937 | 0.5236 | 0.9955 | 0.1597 | 4.9248 | 0.3349 |
| | Ireland | 1.0299 | -2.5234 | 0.0060 | 0.4629 | -1.5643 | 0.9411 | -3.3899 | -0.3180 |
| | Italy | 1.0224 | -2.4157 | 0.0081 | 0.4562 | -1.8487 | 0.9678 | -3.0259 | -0.2807 |
| | Portugal | 1.0180 | -0.9562 | 0.1698 | 0.4876 | -0.5214 | 0.6990 | 0.3914 | 0.0347 |
| | Singapore | 1.0168 | -1.5955 | 0.0557 | 0.5056 | 0.2370 | 0.4063 | -0.0420 | -0.0079 |
| | Spain | 1.0248 | -3.1664 | 0.0008 | 0.4697 | -1.2799 | 0.8997 | -2.6479 | -0.2524 |
| | Sweden | 1.0123 | -1.4080 | 0.0799 | 0.5101 | 0.4266 | 0.3348 | -0.7569 | -0.0615 |
| | Switzerland | 1.0203 | -2.2141 | 0.0137 | 0.4562 | -1.8487 | 0.9678 | -4.8558 | -0.4540 |
| | United Kingdom | 1.0166 | -1.2372 | 0.1083 | 0.5124 | 0.5214 | 0.3010 | 0.3182 | 0.0293 |
| Horizon=12 | Canada | 1.2920 | -3.3113 | 0.0005 | 0.3802 | -4.9922 | 1.0000 | -1.9777 | -0.1303 |
| | Greece | 1.6151 | -4.2475 | 0.0000 | 0.4793 | -0.8640 | 0.8062 | -0.6929 | -0.0273 |
| | Iceland | 1.3376 | -1.8076 | 0.0357 | 0.4585 | -1.7281 | 0.9580 | -0.4125 | -0.0110 |
| | Ireland | 1.4705 | -3.2803 | 0.0006 | 0.4816 | -0.7680 | 0.7788 | -1.6326 | -0.0686 |
| | Italy | 1.4476 | -3.1177 | 0.0010 | 0.4862 | -0.5760 | 0.7177 | -1.4333 | -0.0587 |
| | Portugal | 1.4422 | -2.1496 | 0.0161 | 0.4885 | -0.4800 | 0.6844 | -0.0897 | -0.0036 |
| | Singapore | 1.1842 | -1.6310 | 0.0518 | 0.4977 | -0.0960 | 0.5382 | 0.2842 | 0.0232 |
| | Spain | 1.6483 | -4.0578 | 0.0000 | 0.4309 | -2.8801 | 0.9980 | -2.9208 | -0.1221 |
| | Sweden | 1.3132 | -3.9232 | 0.0001 | 0.4309 | -2.8801 | 0.9980 | -2.6762 | -0.1035 |
| | Switzerland | 1.0947 | -1.6867 | 0.0462 | 0.4562 | -1.8241 | 0.9659 | -2.1963 | -0.0881 |
| | United Kingdom | 1.1577 | -1.6267 | 0.0523 | 0.4977 | -0.0960 | 0.5382 | 0.0594 | 0.0028 |

Note: The column "Countries" shows the names of countries. The column "MSPE_ratio" indicates the ratio of the MSPE of the behavioral model to that of the WN model. The column "DM(t-stats)" shows the Diebold-Mariano (DM) statistics for the test of equal forecast errors and the column "DM(p-value)" shows the p-values associated with the DM statistics. The column "$\hat{p}$" indicates the proportion of correct predictions of direction by the behavioral model, the column "PT(t-stats)" shows the Pesaran-Timmermann (PT) statistics, and the the column "PT(p-value)" shows the p-values associated with the PT statistics. All reported test statistics are one-sided. The significance level is 10%. The column "Avg. Return (%)" shows the annualized average returns in percentage, and the column " Sharp_ratio " shows the annualized Sharp-ratios.

Table 1.5 presents the empirical results for advanced countries. The top part of the table shows the results for forecasts at h=1. The $\text{MSPE}_{ratio}$ is less than one for two countries: Greece and Iceland. Most of the DM statistics are negative; therefore, underlying behavioral models do not outperform the WN model for these countries. Even though the DM statistics are positive for Greece and Iceland, they are not statistically significant. The results show that for the 27% of countries, the proportion of correct predictions of direction ($\hat{p}$) for exchange rates is greater than 50%. However, no result is statistically significant at the 10% significance level. In addition, the results show that investment strategies based on the WN model with a constant under the anchoring-toward assumption have positive and notable average returns for Greece and Iceland. For the remaining advanced countries, the annualized average returns are either positive but small or negative.

The bottom part of Table 1.5 shows the results at h=12. The results show that no $\text{MSPE}_{ratio}$ is less than one, no DM statistic is positive, and no proportion is greater than

50% (i.e., no PT statistics is greater than 0). Therefore, underlying behavioral models do not provide better results than the WN model in predicting point forecasts and directions (signs). The findings show that the past information is not a good predictor of changes in exchange rates for these countries. Changing the probability distribution of the historical exchange rate data does not improve forecasting results. In addition, the results show that investment strategies based on the underlying model under heuristics' assumptions have either positive but small or negative annualized average returns for all advanced countries.

Table 1.6: Empirical results for emerging/developing countries

| | Countries | MSPE_ratio | DM (t-stats) | DM (p-value) | $\hat{p}$ | PT(t-stats) | PT (p-value) | Avg. Return (%) | Sharpe_ratio |
|---|---|---|---|---|---|---|---|---|---|
| Horizon=1 | Brazil | 1.0637 | -2.9994 | 0.0015 | 0.4451 | -1.4825 | 0.9309 | -7.2597 | -0.4404 |
| | Chile | 0.9762 | 0.9064 | 0.1826 | 0.6045 | 4.4086 | 0.0000 | 4.5235 | 0.3240 |
| | Kenya | 0.9689 | 1.2468 | 0.1066 | 0.5258 | 1.0903 | 0.1378 | 4.6083 | 0.3353 |
| | Kuwait | 1.0040 | -0.4505 | 0.3263 | 0.4381 | -2.5861 | 0.9951 | -0.6619 | -0.2378 |
| | Mexico | 0.9538 | 1.3093 | 0.0956 | 0.5423 | 1.6958 | 0.0450 | 5.9490 | 0.3428 |
| | Peru | 1.0323 | -2.1548 | 0.0161 | 0.4894 | -0.3261 | 0.6278 | -1.2022 | -0.2261 |
| | Philippines | 0.9995 | 0.1877 | 0.4256 | 0.5101 | 0.4266 | 0.3348 | 4.0161 | 0.3671 |
| | South Africa | 1.0027 | -0.3575 | 0.3604 | 0.5326 | 1.3750 | 0.0846 | 0.8039 | 0.0487 |
| | Tunisia | 0.9963 | 0.4792 | 0.3160 | 0.5393 | 1.6591 | 0.0485 | 1.9747 | 0.2248 |
| | Uruguay | 0.8264 | 3.1194 | 0.0000 | 0.7027 | 4.5424 | 0.0000 | 16.8092 | 0.8557 |
| Horizon=12 | Brazil | 1.3801 | -4.1159 | 0.0000 | 0.3392 | -4.2060 | 1.0000 | -7.0788 | -0.2125 |
| | Chile | 1.1079 | -0.7979 | 0.2127 | 0.6290 | 5.3761 | 0.0000 | 4.2189 | 0.1191 |
| | Kenya | 0.8696 | 1.0011 | 0.1587 | 0.5760 | 3.1681 | 0.0008 | 3.8975 | 0.1133 |
| | Kuwait | 1.1029 | -1.8151 | 0.0351 | 0.4800 | -0.8246 | 0.7952 | -0.1737 | -0.0272 |
| | Mexico | 0.9895 | 0.1922 | 0.4238 | 0.6189 | 4.7032 | 0.0000 | 4.6941 | 0.1067 |
| | Peru | 1.7708 | -4.3734 | 0.0000 | 0.4063 | -2.8062 | 0.9975 | -1.6391 | -0.1526 |
| | Philippines | 1.4649 | -4.9001 | 0.0000 | 0.4677 | -1.3440 | 0.9105 | -2.5506 | -0.0925 |
| | South Africa | 1.0276 | -0.6220 | 0.2671 | 0.5645 | 2.6880 | 0.0036 | 2.5352 | 0.1055 |
| | Tunisia | 1.0448 | -1.6755 | 0.0473 | 0.4977 | -0.0960 | 0.5382 | 0.0074 | 0.0007 |
| | Uruguay | 0.8912 | 2.2862 | 0.0000 | 0.7206 | 5.1788 | 0.0000 | 16.7672 | 0.8502 |

Note: The column "Countries" shows the names of countries. The column "MSPE_ratio" indicates the ratio of the MSPE of the behavioral model to that of the WN model. The column "DM(t-stats)" shows the Diebold-Mariano (DM) statistics for the test of equal forecast errors and the column "DM(p-value)" shows the p-values associated with DM statistics. The column "$\hat{p}$" indicates the proportion of correct predictions of direction by the behavioral model, the column "PT(t-stats)" shows the Pesaran-Timmermann (PT) statistics, and the column "PT(p-value)" shows the p-values associated with the PT statistics. All reported test statistics are one-sided. The significance level is 10%. The column "Avg. Return (%)" shows the annualized average returns in percentage, and the column " Sharp_ratio " shows the annualized Sharp-ratios.

Table 1.6 presents the empirical results for emerging/developing countries. The top part of the table shows the results for forecasts at h=1. The $\text{MSPE}_{ratio}$ is less than one for six countries, including Chile, Mexico, and Uruguay. The DM statistics are positive and statistically significant for two countries, Mexico and Uruguay. The proportion of correct directions ($\hat{p}$) is greater than 50% and statistically significant for five countries, including Chile, Mexico, South Africa, Tunisia, and Uruguay. Therefore, underlying behavioral models outperform the WN model in predicting the direction of changes in exchange rates for these countries. In addition, the results show that investment strategies based on underlying models under the anchoring-toward assumption have positive and notable annualized average returns for some countries: Chile, Mexico, Kenya, Philippines, and Uruguay. For the remaining emerging/developing countries, the annualized average returns are either positive but small or negative.

The bottom part of Table 1.6 shows the results for forecasts at h=12. The results show that the $\text{MSPE}_{ratio}$ is less than one for three countries, including Kenya, Mexico, and Uruguay. The DM statistic is positive and statistically significant for Uruguay. However, DM statistics are either negative or statistically insignificant for other countries. Therefore, underlying behavioral models do not outperform the WN model in terms of point forecasts for these countries. The findings show proportions of correct direction ($\hat{p}$) are greater than 0.5 and statistically significant for five countries: Chile, Kenya, and Mexico. Hence, underlying behavioral models provide better results than the WN model in predicting these countries' directions (signs). The findings highlight the advantage of using historical data and changing their probability distributions for some emerging/developing countries such as Mexico.

Overall, the findings are consistent with the literature that predicting exchange rates is a difficult task. The results depend on different factors, including sample sizes, test statistics, and horizons. The empirical results are different within emerging/developing countries. A plausible explanation could be that the number of observations is less for some emerging/developing countries, including Chile, Mexico, and South Africa, than for other emerging/developing countries. This could also explain the differences between the results for the advanced and emerging/developing countries.

In the cases when both advanced and emerging/developing countries have the same number of observations but the empirical results are different, a plausible reason could be the lack of deep foreign exchange markets in emerging/developing countries. In addition, there are other differences between the advanced and emerging/developing countries, including economic policies, trades volume, and market speculations, among other factors. These factors could impact people's financial decisions. Economic policies are more transparent, and economic conditions are more stable in the advanced countries than in the emerging/developing countries.

The important point is that currency prices result from supply and demand in the foreign exchange market. Supply and demand factors are constantly shifting, and there are various traders who may react to these shifts differently. Various traders in the foreign exchange market include governments, central banks, financial institutions, large corporations, small businesses, currency speculators, and individuals. If I had complete information about all traders' supply and demand of currency at each time $t$, I would have been able to precisely forecast the exchange rate movements in principle. Therefore, the significant challenge is that economists and financial analysts, including this study, do not have a complete picture to correctly predict the point forecasts and the direction of exchange rates for all countries, models, horizons, etc.

## 1.6 Conclusion

This essay investigated the forecastability of exchange rates using behavioral economics and the historical exchange rates data. The purpose of this essay was to forecast changes in exchange rates for the US dollar versus 37 currencies. As explained before, the out-of-sample performance of existing models depends on various specifications, including estimation methods, horizons, and test statistics. Therefore, in the literature, there is no single forecasting model with consistent performance across all currencies.

I focused on the probability distribution of exchange rates changes and considered subjective probabilities. These subjective probabilities depend on underlying assumptions of individual or combined heuristics. I used these subjective probabilities as weights in autoregressive models, and I estimated coefficients and formed forecasts for changes in exchange rates.

I used behavioral heuristics, including availability heuristic, conservatism, pessimism, optimism, overconfidence, and underconfidence. I set the median of historical data for each currency pair as an anchor and presented two heuristics, anchoring-toward and anchoring-away.

I incorporated heuristics both individually and as a set of mutually exclusive heuristics. The set of mutually exclusive heuristics are as follows:

- Rationality, Availability heuristic, Conservatism
- Rationality, Optimism, Pessimism
- Rationality, Overconfidence, Underconfidence
- Rationality, Anchoring-toward, Anchoring-away

I evaluated the out-of-sample predictive ability of models in two ways: their ability to predict point forecasts and their ability to forecast the direction of changes in exchange rates. Therefore, in addition to the $\text{MSPE}_{ratio}$ and DM statistics, I used the PT statistic.

The empirical results show no model successfully outperforms the WN model for all countries in terms of diagnostics such as the $\text{MSPE}_{ratio}$ or DM statistics. However, behavioral models provide promising results for some countries forecasting by the direction of change measure (the PT statistic). Particularly, models under assumptions of anchoring-toward and optimism heuristics outperform the WN model in 1-month-ahead and 12-month-ahead forecasts, respectively, in some emerging/developing countries.

The empirical results are different across emerging/developing countries. A reasonable explanation could be that the data spans are different for some emerging/developing countries, including Chile, Mexico, and South Africa, from other emerging/developing countries. This could also be a reason why the empirical results for these countries are different from the advanced countries. There are differences in the empirical results between advanced and emerging/developing with the same data spans. There are plausible reasons, including the lack of deep foreign exchange markets in emerging/developing countries, economic policies, trades volume, and market speculations, etc. Clearly, the economic conditions have significant effects on people's financial decisions.

Consistent with the literature, the findings highlight the challenge of predicting exchange rates. The empirical evidence confirms that the choice of data spans, test statistics, horizons, etc., is important in predicting exchange rates. The results indicate the advantage of using the behavioral heuristics in forecasting procedures. This essay used weights (the subjective probabilities) in the autoregressive model of changes in exchange rates. Concerning the forecasting of exchange rates, $s_t$, I show that in a few cases it is possible to improve upon the random walk model forecast $\hat{s}_{t+h} = s_t$ by replacing it by $\hat{s}_{t+h} = s_t + \widehat{\Delta} s_{t+h}$. For all other cases, the random walk model is the best forecasting model for exchange rates.

Note the analysis with a shorter rolling window, like 40 or 60, might have given better results. Farmer, Schmidt, and Timmermann (2019) explained the "pockets of predictability." They found the predictability of stock returns in short intervals (locally) but did not find predictability globally. The future directions will be to use weights in non-linear models and consider a shorter rolling window.

# Essay 2

# Investor Sentiment Indicators and Exchange Rate Forecasting

## 2.1 Introduction

Since the Meese and Rogoff (1983a) and Meese and Rogoff (1983b) studies, it has been well known that forecasting exchange rates is a challenging task. It has been shown that a simple model such as a random walk without drift model often forecasts exchange rates better than multi-factor models. This finding is known as "the Meese and Rogoff puzzle" Rossi (2013). Overall, the literature shows that existing models' performance depends on predictor choice (e.g., macroeconomic and financial factors), estimation methods, evaluation measures, sample periods, benchmark models, and horizons. Nevertheless, as mentioned before, the random walk without drift model generally provides better out-of-sample performance than existing models. The task is to develop models that forecast exchange rates more accurately than existing models.

Cerra and Saxena (2010) use exchange rates (in levels) and fundamental variables (in levels). They obtain forecasts across different horizons using the following set of models: different monetary models (in levels, in growth rates using estimated coefficients, and in growth rates using theoretical coefficients) and the error correction model. In monetary models in the levels and growth rates, they use the actual future values of fundamental variables. In error correction models, they use current variables as regressors. They conclude that the fundamental-based models beat the random walk without drift and the random walk with drift models in terms of root mean squared error (RMSE).

Rossi (2013) uses the most common predictors and successful methodologies in the forecasting exchange rate literature, including interest rate differentials, price differentials, money and output differentials, the Taylor Rule model, among other factors. She then considers two multivariate models: Bayesian Model Averaging (BMA) and Vector Error Correction Model (VECM). Rossi (2013) compares the forecasting models' performance with the random walk without drift model using the root mean squared forecast error (RMSFE). She concludes that the Taylor Rule model and the net foreign assets have some predictive abilities at short horizons and the monetary models show some predictive ability at longer horizons. However, none of these models could outperform the random walk model for all countries and periods.

Alba et al. (2015) consider Taylor Rule models based on different assumptions. 1) They examine both symmetric and asymmetric Taylor Rule models; the foreign country central bank does not target exchange rates in the symmetric model. 2) They test smoothing and no-smoothing Taylor Rule models; interest rate adjustment happen instantaneously in the no-smoothing model. 3) They study both homogeneous and heterogeneous coefficients models, the coefficients of inflation, the output gap, and the interest rate smoothing are the same in the US and the foreign country in the homogeneous coefficients model. 4) They examine Taylor Rule models with and without a constant; the constant is excluded when the coefficients on inflation and interest rate smoothing, the inflation targets, and the equilibrium real interest rates are equal in the US and the foreign country. Alba et al. (2015) use Clark and West (2006) and Clark and West (2007) (CW) metric to evaluate the out-of-sample forecasting performance of considered models. The benchmark model is the random walk model. They conclude that there is evidence of out-of-sample exchange rate predictability using Taylor Rule models for some emerging countries, for example, Brazil.

This essay contributes to the literature by investigating whether adding investor sentiment (behavioral) variables to macroeconomic models could better forecast changes in exchange rates. In the first step, I examine the exchange rate predictability using well-known macroeconomic variables for the US dollar versus 37 countries, both advanced and emerging/developing countries. To be clear, I use realized macroeconomic variables as explanatory variables (predictors) in linear models to forecast changes in exchange rates. Macroeconomic models include the purchasing power parity (PPP), uncovered interest rate parity (UIRP), and Taylor Rule, among others.

In the second step, I investigate the predictive ability of changes in a terms of trade index both as a single predictor and as an added predictor in the UIRP model. Also, I examine changes in commodity and oil prices (both in nominal and real terms) as predictors of changes in exchange rates. Moreover, in contrast to Engel, Lee, Liu, Liu, and Wu (2019) that examines the in-sample performance of the Extended UIRP model, I look at the model's out-of-sample performance.

In the third and most significant step, I add changes in investor sentiment indices, including business confidence index (BCI), consumer confidence index (CCI), composite leading indicator (CLI), Cboe Volatility Index (VIX), and Cboe SKEW Index (SKEW) to macroeconomic models to forecast changes in exchange rates. I examine whether this approach improves the predictive ability of macroeconomic models and also whether using changes in investor sentiment indices generally leads to better forecasting results than a benchmark model, the White Noise (WN) model of $\Delta s_t$, assuming that $S_t$ (the level of exchange rate) follows a random walk model.

In contrast to Morales-Arias and Moura (2013), I use changes in BCI and changes in CCI as predictors and add these variables to macroeconomic models. As an extension to the Cheung et al. (2019) study, which adds VIX to the monetary model, I add this variable to other macroeconomic models, such as the PPP and Taylor Rule models. As far as I am aware, this essay is the first to use composite leading indicator and SKEW indices as the predictors of exchange rates. VIX and SKEW indices provide risk measures in financial markets.

My objective is to form forecasting models by adding investor sentiment variables to macro models and evaluate the out-of-sample performance of these models. On a rolling window basis, I split a given data set into an in-sample period to estimate a linear model and an out-of-sample period to evaluate forecasting performance.

For evaluations, I compare the out-of-sample forecasting performance of models with the benchmark model. I use various statistics, including the ratio of Mean Squared Prediction Error ($MSPE_{ratio}$), Diebold-Mariano (DM) [developed by Diebold and Mariano (1995)], and Pesaran-Timmermann (PT) statistics [developed by Pesaran and Timmermann (1992)]. In addition to statistical forecast evaluation measures, I use the Sharpe-ratio statistic to evaluate investment strategies, using the expected exchange rate changes to calculate investment returns.

The rest of the essay is organized as follows: Section 2.2 presents modeling approaches. Section 2.3 describes data and Section 2.4 explains methodologies and implementations. Section 2.5 reports empirical results and Section 2.6 concludes.

## 2.2 Modeling Approaches

In this section, I explain factor models used in this essay to forecast the future movements of exchange rates. In the first step, I use common macroeconomic factor models in the literature to forecast exchange rates. As well, I extend existing macroeconomic models by adding predictors to models and/or analyzing the out-of-sample performance of models. I also introduce new predictors. In the second step, I add behavioral factors to macroeconomic models in the first step. To demonstrate and distinguish steps, I divide models into the following groups :

- Non-Behavioral Models (Group A): Existing macroeconomic models, the extensions of existing macroeconomic models, and new macroeconomic predictors belong to this group.

- Behavioral Models (Group B): Group A models, which are extended by investor sentiment variables (behavioral factors), belong to this group.

### 2.2.1 Non-Behavioral Models (Group A)

In this section, I first examine the predictability of exchange rates using the set of macroeconomic models based on the following criteria: (1) These models are prominent in economic and policy literature. (2) These models can readily be implemented. In addition and most importantly, the benefit of examination from the perspective of practitioners in financial markets is that such evaluations have been done rarely. For policymakers and market participants, predicting the correct direction of changes in exchange rates is as essential as predicting accurate point forecasts.

Second, I examine models/predictors that are either the extensions of existing models (e.g., Extended UIRP) or introduced as a predictor of exchange rates (e.g., the real commodity price and the real oil price) in this essay. Adding one or more variables to existing models is motivated by the fact that added variables (predictors) may have information that could improve these models' overall out-of-sample forecasting performance.

Because my objective is to forecast changes in exchange rates, which is the stationary process, I examine whether explanatory variables series are stationary processes. Suppose a data set has trends and therefore is a non-stationary process. In that case, I transform the data set into a stationary process by computing the differences between the logarithm of consecutive observations.

Non-behavioral models (Group A) are as follows:

- Purchasing Power Parity (PPP)
- Oil Price and Commodity Price
- Real Oil Price and Real Commodity Price
- Terms of Trade
- Uncovered Interest Rate Parity (UIRP) and UIRP augmented by Terms of Trade
- Extended UIRP

- Taylor Rule Model
- Monetary Model with Flexible Prices

In the following subsections, I provide a brief overview of each model.

#### 2.2.1.1 Purchasing Power Parity (PPP)

Based on purchasing power parity (PPP), the real price of comparable commodity baskets in two countries should be the same. This means the home country's price level, converted to the foreign currency using the exchange rate, should be the same as the price level in the foreign country. This model has been used in previous studies, for example, Rossi (2013). Note that if I define $S$ as the nominal exchange rate and $R$ the real exchange rate, then by definition $R = S(\frac{P^*}{P})$, where $P$ is the US price level and $P^*$ is the foreign price level. The idea of PPP is that $R = 1$, so that $S = \frac{P}{P^*}$. If I use $s = log(S)$, $p = log(P)$ and $p^* = log(P^*)$, then I have in log, $s = p - p^*$. I use the rate of change of variables; therefore, the PPP model is as follows:

$$\Delta s_{t+h,h} = \alpha_0 + \alpha_1(\pi_t - \pi_t^*) + u_{t+h,h}, \tag{3.1.1}$$

where $\Delta s_{t+h,h} = s_{t+h} - s_t$ is the $h$-month-ahead change of nominal exchange rates, with $s_t$ being the log of US dollar per unit of foreign currency. For brevity, I use 'changes in exchange rates', instead of 'the $h$-month-ahead changes in nominal exchange rates' in the rest of this essay. $\pi_t$ is the inflation rate for the US and $\pi_t^*$ is the inflation rate for a foreign country.

#### 2.2.1.2 Oil Price and Commodity Price

Chen and Rogoff (2003) examine commodity prices as new macroeconomic factors for exchange rates. They examine the relationship between real exchange rates and commodity prices and focus on the in-sample empirical analysis. They explain that from a policy standpoint, examining and understanding the impacts of commodity price shocks on the exchange rates should be of great importance to developing commodity-exporting countries, particularly as they open capital markets and adopt flexible exchange rate regimes. Their in-sample empirical results show evidence in favor of commodity prices as predictors of real exchange rates for three advanced countries.

Chen, Rogoff, and Rossi (2010) use commodity prices to forecast the future movements of nominal exchange rates. Their empirical results indicate that commodity prices are not successful out-of-sample predictors for exchange rates. The empirical results in the Ferraro, Rogoff, and Rossi (2015) study suggest that commodity prices (particularly, the oil price) can predict exchange rates at a daily frequency. However, the predictive ability is not significant at quarterly and monthly frequencies.

In contrast to the study by Ferraro et al. (2015), which only focuses on a few countries, I use changes in commodity and oil prices as predictors of exchange rates for a large number of countries, both advanced and emerging/developing countries. Also, I use the Pesaran-Timmermann (PT) to evaluate the out-of-sample performance of models including these variables; none of the above studies uses this statistic.

Even though the oil price is included in the aggregate commodity price, I examine its predictive ability separately. Because the oil price represents the energy sector, and its volatility alone could significantly impact the foreign exchange market for all countries. In addition, according to International Monetary Fund (IMF) definitions, the weight of energy prices in the aggregate commodity price index is close to 0.4 (40%). Therefore, the aggregate commodity price index mostly represents the non-fuel commodity prices. Models with the oil price and aggregate commodity price are as follows:

$$\Delta s_{t+h,h} = \alpha_0 + \alpha_1 \Delta Oil_t + u_{t+h,h}, \tag{3.1.2}$$

$$\Delta s_{t+h,h} = \alpha_0 + \alpha_1 \Delta CP_t + u_{t+h,h}, \tag{3.1.3}$$

where $\Delta CP_t$ is the first difference of the logarithm of (world) aggregate commodity price and $\Delta Oil_t$ is the first difference of the logarithm of oil price.

#### 2.2.1.3 Real Oil Price and Real Commodity Price

This section presents the motivation to examine exchange rates' predictability using real commodity and real oil prices. The literature [see, e.g., Chen et al. (2010) and Ferraro et al. (2015)] uses nominal commodity and oil prices to forecast movements of exchange rates. Essentially, the commodity price and oil price are intended to capture the terms of trade effect.

Suppose the world demand for primary products (natural resources) rises (relative to world demand for manufacturing products such as electronics and automobiles). Such an increase in the relative demand for primary products will be reflected in the world relative price of primary products vs. that of manufacturing products. In such a case, real commodity prices will rise.

An example of an event that leads to a change in demand is the rise of some emerging/developing countries (e.g., China and India), which forms a strong need to enhance their social infrastructure, which leads to a massive demand for primary products. Examples of social infrastructure are schools and other educational centers, transport services such as sidewalks and highways, and health care services such as hospitals, medical clinics, and long-term care facilities. Examples of primary products are energy (e.g., oil and natural gas), wood, and metals.

An increase in the world relative demand for primary products will, therefore, exogenously increase the demand for the exported products from countries that are rich in natural resources. Thus, the relative demand for those countries' currencies increases, hence driving up those countries' currency values.

To check whether there are differences between the predictive ability of the nominal oil price and real oil price to forecast changes in exchange rates, I use real oil price as a predictor. The same logic goes for the commodity price index. Using US inflation rate as a deflator is a common practice in the literature. Following the Chen et al. (2010) study, since the nominal commodity price and oil price are in the US dollars, I use US inflation as the deflator. Models with real commodity and real oil prices are as follows:

$$\begin{aligned} \Delta s_{t+h,h} &= \alpha_0 + \alpha_1(\Delta CP_t - \pi_t) + u_{t+h,h}, \\ \Delta s_{t+h,h} &= \alpha_0 + \alpha_1(\Delta Oil_t - \pi_t) + u_{t+h,h}, \end{aligned} \tag{3.1.4}$$

where $\Delta CP_t$ is the first difference of the logarithm of (world) aggregate commodity price, $\pi_t$ is the US inflation rate, and $\Delta Oil_t$ is the first difference of the logarithm of oil price.

#### 2.2.1.4 Terms of Trade

In this section, I explain the terms of trade index and the motivation to use it as a single predictor and as an added predictor to the UIRP model. Terms of trade is the ratio of export prices to import prices for a given country. In the global market, countries use capital flows to balance out trade deficits. Improving terms of trade generally means that the currency appreciates and there is more demand for it. In contrast, deteriorating terms of trade means depreciation of the currency since the given country must spend more money to import the same products. Therefore, fluctuations in export and import prices are reflected in the currency pairs. The terms of trade effect, which essentially captures the relative demand (and/or supply) for a country's exports (relative to a country's imports), will cause movement in real exchange rates (and hence nominal exchange rates).

Broda (2004) examines the impact of terms of trade shocks on the movements of exchange rates. The empirical results show after a negative shock to terms of trade, real exchange rates depreciate. Based on this result, the terms of trade may have information to predict changes in exchange rates.

The terms of trade index has been used as a predictor in the Cheung et al. (2019) study in the behavioral equilibrium exchange rate model for only five advanced countries. In contrast to the Cheung et al. (2019) study, I examine whether the terms of trade index is a good predictor of changes in exchange rates, both as a single factor and as an added factor to the UIRP model for a large number of countries. In this essay, the terms of trade model is as follows:

$$\Delta s_{t+h,h} = \alpha_0 + \alpha_1 \Delta tot_t^* + u_{t+h,h}, \tag{3.1.5}$$

where $\Delta tot_t^*$ is the first difference of the logarithm of terms of trade index for a foreign country.

#### 2.2.1.5 Uncovered Interest Rate Parity (UIRP) and UIRP augmented by Terms of Trade

Fisher (1896) explains the analysis of the relation between interest rates and expected changes in the relative value of account or commodities. He uses international currencies as an example and this is known as the uncovered interest rate parity (UIRP) model. Based on UIRP, the expected change of nominal exchange rate is equal to the nominal interest rate differential. If UIRP holds, it can be used as a forecasting equation. Meese and Rogoff (1988) use the uncovered interest rate parity to forecast real exchange rates.

According to the Molodtsova and Papell (2009) study, while empirical results show that UIPR can explain movements of exchange rates in the long-run, the model does not hold in the short-run. Therefore, Molodtsova and Papell (2009) use a flexible specification of the model. They do not put any restriction on coefficients and their signs. This model has been used in other studies including Rossi (2013). The UIRP model is as follows:

$$\Delta s_{t+h,h} = \alpha_0 + \alpha_1(i_t - i_t^*) + u_{t+h,h}, \tag{3.1.6}$$

where the US interest rate is $i_t$ and a foreign country interest rate is $i_t^*$. The motivation to include terms of trade in the UIRP is that this predictor contains information about a given country's economic position in international markets that interest rate differentials may not provide. Therefore, I add changes in terms of trade index to equation (3.1.6) as follows:

$$\Delta s_{t+h,h} = \alpha_0 + \alpha_1(i_t - i_t^*) + \alpha_2 \Delta tot_t^* + u_{t+h,h}, \tag{3.1.7}$$

where the US interest rate is $i_t$, a foreign country interest rate is $i_t^*$, and $\Delta tot_t^*$ is the first difference of the logarithm of terms of trade index for a foreign country.

#### 2.2.1.6 Extended Uncovered Interest Rate Parity (Extended UIRP)

I use the uncovered interest parity (UIRP) model to forecast the future movements of exchange rates. I extend the standard UIRP model and add inflation rates as in Engel et al.

(2019). As they explain, this specification is based on the fact that many central banks have adopted inflation targeting monetary policy rules. The inflation rate may have information that is not included in the interest rate differential. Engel et al. (2019) focus on the in-sample forecasting power of the Extended UIRP model.

My work is different from the Engel et al. (2019) in the following ways. First, in this essay, I evaluate the out-of-sample forecasting performance of the Extended UIRP model. Second, I use the same coefficients on inflation rates for the US and a foreign country. Using the same coefficients is consistent with the assumption of homogeneous coefficients across all models in this essay. Third, I examine a large number of countries, both advanced and emerging/developing countries. The Extended UIRP model is as follows:

$$\Delta s_{t+h,h} = \alpha_0 + \alpha_1(i_t - i_t^*) + \alpha_2(\pi_t - \pi_t^*) + u_{t+h,h}, \tag{3.1.8}$$

where the US interest rate is $i_t$ and a foreign country interest rate is $i_t^*$. $\pi_t$ is the inflation rate for the US and $\pi_t^*$ is the inflation rate for a foreign country.

#### 2.2.1.7 Taylor Rule Model

The Taylor Rule indicates that central banks adjust the nominal interest rate in response to inflation rates and output gap changes. Using the Taylor Rule for two countries (home and foreign) and subtracting the foreign country from the home country, an equation is derived with the interest rate differential on the left-hand-side and the inflation and output gap differential on the right-hand-side. Molodtsova and Papell (2009) modify the original Taylor Rule model. Therefore, they replace policy rates with the Taylor Rule objects, use UIRP, and re-define coefficients.

Several studies including Molodtsova and Papell (2009) and Rossi (2013) have used the Taylor Rule model to forecast the future movements of exchange rates. The Taylor Rule model is as follows:

$$\Delta s_{t+h,h} = \alpha_0 + \alpha_1(\pi_t - \pi_t^*) + \alpha_2(y_t^{gap} - y_t^{gap*}) + u_{t+h,h}, \tag{3.1.9}$$

where $\pi_t$ is the US inflation rate, $\pi_t^*$ is a foreign country inflation rate, $y_t^{gap}$ is the US output gap, and $y_t^{gap*}$ is a foreign country output gap.

#### 2.2.1.8 Monetary Model with Flexible Prices

Based on the monetary model's definition, movements of exchange rates should reflect changes in countries' relative money, output, interest rates, and prices. Real money demand is considered as a function of interest rate and income. To obtain the relationship between money and output differential and exchange rates, UIRP and PPP are used to substitute relative interest rates and prices as a function of exchange rates. One approach assumes that PPP holds at every point in time and replaces it in relative money demand to get the monetary model's flexible price version (see, e.g., Molodtsova and Papell (2009), Cerra and Saxena (2010), Rossi (2013), Morales-Arias and Moura (2013), and Kouwenberg et al. (2017)). In this essay, I use the flexible price monetary model as follows:

$$\Delta s_{t+h,h} = \alpha_0 + \alpha_1(i_t - i_t^*) + \alpha_2(\Delta y_t - \Delta y_t^*) + \alpha_3(\Delta m_t - \Delta m_t^*) + u_{t+h,h}, \qquad (3.1.10)$$

where $i_t$ is the US interest rate, $i_t^*$ is a foreign country interest rate, $\Delta y_t$ is the growth output rate for the US, $\Delta y_t^*$ is the growth output rate for a foreign country, $\Delta m_t$ is the growth money rate for the US, and $\Delta m_t^*$ is the growth money rate for a foreign country.

### 2.2.2 Behavioral Models (Group B)

Behavioral models (Group B) are extensions of non-behavioral models (Group A) by adding investor sentiment variables (behavioral factors)[10]. In the first step, I examine whether adding investor sentiment indices to the 'best' macroeconomic model could improve the 'best' macro model's ability to correctly forecast the direction of change in exchange rates. Therefore, among all non-behavioral (Group A) models, the 'best' model is chosen using

[10] I am grateful to Professor Andrei Semenov for his comments and inputs in this section, particularly the extension of 'best' macroeconomic models by adding behavioral factors and choosing the 'best' model among up to 32 available forecasting models for a given country and horizon.

out-of-sample statistics, particularly, the PT statistic. Based on data availability for a given country, I use up to 5 investor sentiment indices; therefore, I add up to 31 different combinations of investor sentiment indices to the 'best' model in Group A. The full list of 31 combinations are in Table 2.4.

In the next step, I evaluate the out-of-sample performance of all behavioral (i.e., macroeconomic plus investor sentiment variables) models. Based on data availability for a given country, there are up to 11 macroeconomic (non-behavioral) models and up to 5 investor sentiment indices, thus there are up to 352 behavioral models. I provide more details about procedures and estimation of models in Section 2.4.

Here, I present investor sentiment indices (behavioral factors) and motivations to add them to the set of macroeconomic models in Group A. I focus on the following:

- Overviews for Investor Sentiment Indices (Behavioral Factors)
- Motivations for Investor Sentiment Indices (Behavioral Factors)

#### 2.2.2.1 Overviews for Investor Sentiment Indices (Behavioral Factors)

Here, I provide definitions and descriptions of investor sentiment indices. As I explained in Section 2.1, there are 5 investor sentiment indices used in this essay, including the composite leading indicator (CLI), consumer confidence index (CCI), business confidence index (BCI), VIX, and SKEW indices.

CLI is developed to provide an early indication of turning points in business cycles showing the shift of the economic activity around its potential long-term level. CLI shows short-term economic movements in qualitative rather than quantitative terms (OECD (2020b)). CCI is based on households' plans for major purchases and their economic situation, both currently and their expectations for the near future. Opinions are collected and compared to a normal state, and the difference between positive and negative answers provides the index on economic conditions (OECD (2020c)). BCI provides information about people's opinions about future developments on orders, stocks, sales, and productions. The value of the index shows that whether people (investors) are optimistic or pessimistic toward the future performance of businesses (OECD (2020a)).

The VIX index provides a measure of constant, 30-day expected volatility of the US stock market, derived from real-time, mid-quote prices of $S\&P$ 500 index. The VIX index is one of the well-known global volatility indicators, and it is widely reported by financial media and different financial market participants closely follow its movements (Cboe (2020b)). The SKEW index is an index derived from the price of $S\&P$ 500 tail risk. The price of $S\&P$ 500 tail risk is calculated from the price of $S\&P$ 500 out-of-the-money options. SKEW values generally are between 100 and 150. A value higher than 100 means the left tail of the $S\&P$ 500 distribution gains more weight, and the probabilities of outlier returns become more significant (Cboe (2020a)). In other words, the risk of a sudden move in the stock market increases. High VIX means a high degree of worry/concern and uncertainty.

#### 2.2.2.2 Motivations for Investor Sentiment Indices (Behavioral Factors)

Here, I explain why adding investor sentiment indices could improves the predictive ability of macroeconomic models. According to Akerlof and Shiller (2010), the word "confidence" is used often in the business and economic literature. Most economists focus on the predictive definition of confidence, which means an expectation of a promising future. Akerlof and Shiller (2010) argue the common usage among people emphasizes the word's implication of trust and belief. In the 2008 financial crisis, the lack of confidence created uncertain environments in credit markets.

In this essay, motivated by Akerlof and Shiller (2010) and Morales-Arias and Moura (2013), among other studies, I take into account impacts of emotional and cognitive factors on forecasting decisions; therefore, I use investor sentiment indices, which are proxies for consumer and businesse' trust and beliefs about economic conditions.

I note that Morales-Arias and Moura (2013) examine BCI and CCI as single predictors in panel data for 12 advanced countries. My work is different from the Morales-Arias and Moura (2013) study in the following ways. First, in this essay, I use time-series data for individual countries. Second, the data cover 26 countries. Third, I incorporate changes in these variables (instead of their levels) as single predictors and combine them with other investor sentiment indices. Fourth, I use the composite leading indicator (CLI) index, which is not used in the Morales-Arias and Moura (2013) study.

The United States' efficient fiscal and monetary policies and a stable economy make the US dollar the safe-haven currency, especially in times of crisis and uncertainty. The US dollar is the default and strong currency for international investors facing any weak and uncertain domestic currency since it is the world's reserve currency and the denomination for many international business deals.

In this essay, motivated by the significance of the US dollar as the stable currency, I use VIX and SKEW as predictors of exchange rates. Investors use VIX and SKEW to measure the stress and fear in the financial market when making investment decisions. VIX and SKEW indicators provide early warning of significant drops in the stock market.

As an extension to the Cheung et al. (2019) study that uses VIX in the monetary model for only five advanced countries, I add VIX as both a single predictor and simultaneously with other investor sentiment indices to Group A's macroeconomic models for 37 countries, both advanced and emerging/developing.

In this essay, SKEW is, for the first time, introduced and examined as a predictor of exchange rates. I add SKEW as a single predictor and simultaneously with other investor sentiment indices to Group A's macroeconomic models. Note that as predictors, I use the first difference of the logarithm of each investor sentiment index, i.e., $\Delta BIC, \Delta CCI, \Delta CLI, \Delta VIX$, and $\Delta SKEW$ to ensure that all processes are stationary.

## 2.3 Data Description

This section describes monthly data and methodologies used to prepare the data for empirical analysis [see more details about variables' definitions and their sources in Appendix C]. The exchange rates are the end-of-month values of the US dollar versus a currency. The data source is IMF's International Financial Statistics (IFS). The data are from Jan 2000 to Dec

2020. Countries with a fixed-exchange-rate system or a small sample of macroeconomic variables are excluded from the analysis.

To compare the empirical results in this essay with the previous discussed studies, including Molodtsova and Papell (2009) and Rossi (2013), and where possible, I choose macroeconomic variables similar to variables used in these studies. In addition, I use seasonally adjusted data as in earlier studies. All data for macroeconomic and investor sentiment indices are from Jan 2000 to Dec 2020. I choose data starting from Jan 2000 because some macroeconomic variables start later for some emerging/developing countries than advanced countries. I want to include as many countries as possible in the analysis.

The price level is the consumer price index (CPI). The monthly inflation rate is measured as the 1-month difference of logarithm of CPI, and the annual inflation rate is measured as the 12-month difference of logarithm of CPI. In the PPP, Extended UIRP, and Taylor Rule models, the monthly and annual inflation rates are used for the 1-month-ahead and 12-months-ahead forecasts, respectively. The seasonally adjusted industrial production (IP) index is used for output because the gross domestic product (GDP) is unavailable monthly. The data source is the IMF's International Financial Statistics (IFS) for both CPI and IP. The data source for seasonally adjusted money (M1) and non-seasonally adjusted terms of trade index is Refinitive (formerly Thomson Reuters Financial & Risk Business). After collecting terms of trade data sets, they are seasonally adjusted [11].

Interest rates are Call Money (Interbank Rates), and the data source is Refinitive. The commodity price (CP) is the world commodity price index, including fuel and non-fuel price indices. The oil price is Crude Oil (petroleum), West Texas Intermediate 40 API, Midland Texas, US$ per barrel. The data source is IMF's International Financial Statistics (IFS) for these variables.

To introduce behavioral aspects of consumers and businesses in the economy, I use monthly data on CCI, BCI, and CLI and take the logarithm of these variables and cal-

[11] Terms of trade data are seasonally adjusted using a one-sided moving average with backward, equal weights, following the approach used in Rossi (2013). Therefore, the seasonal adjustment for monthly data is $\frac{1}{12} + \frac{1}{12}L + ... + \frac{1}{12}L^{11}$.

culate the first differences. I should emphasize that data for these variables are not available for some countries. The data source for these variables is the Organisation for Economic Co-operation and Development (OECD). The first difference of the logarithm of VIX and SKEW indices are used as predictors. The data source for these variables is Global Markets, Inc. (Cboe). I use the same data sets for the commodity price, oil price, VIX, and SKEW variables for all countries.

I consider two groups of countries, advanced and emerging/developing, using the World Economic Outlook (WEO) classifications. This classification has evolved [see more details in Appendix B]. Countries in a given group have similar economic conditions and financial markets, leading to similar forecasting results. The descriptive statistics for changes in nominal exchange rates for advanced countries are reported in Table 2.1. The table shows for all countries, the excess kurtosis is positive. Positive excess kurtosis means the data distribution is not normal, and the distribution has a fat tail. Fat tails indicate a higher probability of larger positive and negative observations compared to the normal distribution.

Table 2.1: Descriptive statistics for advanced countries

| **Countries** | **Start Date** | **End Date** | **Count** | **Mean** | **SD** | **25%** | **Median(50%)** | **75%** | **Min** | **Max** | **Skw** | **Exc-Kurtosis** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Canada | 2000-01-28 | 2020-12-28 | 252 | 0.0001 | 0.0261 | -0.0143 | 0.0001 | 0.0157 | -0.1238 | 0.0887 | -0.4229 | 2.4615 |
| Czech Republic | 2000-01-28 | 2020-12-28 | 252 | 0.0021 | 0.0346 | -0.0179 | 0.0026 | 0.0250 | -0.1169 | 0.0993 | -0.2892 | 0.8512 |
| Denmark | 2000-01-28 | 2020-12-28 | 252 | 0.0008 | 0.0286 | -0.0157 | 0.0000 | 0.0197 | -0.1124 | 0.1024 | -0.2075 | 1.6007 |
| Estonia* | 2000-01-28 | 2020-12-28 | 252 | 0.0007 | 0.0285 | -0.0157 | 0.0010 | 0.0201 | -0.1197 | 0.0910 | -0.3346 | 1.6219 |
| Euro Area | 2000-01-28 | 2020-12-28 | 252 | 0.0008 | 0.0282 | -0.0160 | 0.0000 | 0.0198 | -0.1144 | 0.0894 | -0.2548 | 1.4112 |
| Greece* | 2000-01-28 | 2020-12-28 | 252 | 0.0007 | 0.0282 | -0.0160 | 0.0000 | 0.0198 | -0.1144 | 0.0894 | -0.2391 | 1.4085 |
| Iceland | 2000-01-28 | 2020-12-28 | 252 | -0.0022 | 0.0404 | -0.0226 | -0.0011 | 0.0215 | -0.2019 | 0.1601 | -0.9694 | 4.9983 |
| Israel | 2000-01-28 | 2020-12-28 | 252 | 0.0010 | 0.0219 | -0.0106 | 0.0019 | 0.0149 | -0.1006 | 0.0592 | -0.5485 | 1.6771 |
| Japan | 2000-01-28 | 2020-12-28 | 252 | -0.0001 | 0.0257 | -0.0149 | 0.0006 | 0.0146 | -0.0850 | 0.0766 | -0.1969 | 0.6600 |
| Korea | 2000-01-28 | 2020-12-28 | 252 | 0.0002 | 0.0303 | -0.0126 | 0.0024 | 0.0160 | -0.1292 | 0.1539 | -0.2369 | 5.1808 |
| Norway | 2000-01-28 | 2020-12-28 | 252 | -0.0002 | 0.0341 | -0.0210 | 0.0000 | 0.0193 | -0.1376 | 0.0806 | -0.2562 | 0.9611 |
| Slovakia* | 2000-01-28 | 2020-12-28 | 252 | 0.0022 | 0.0286 | -0.0152 | 0.0016 | 0.0214 | -0.0989 | 0.0950 | -0.1582 | 0.8440 |
| Slovenia* | 2000-01-28 | 2020-12-28 | 252 | 0.0000 | 0.0282 | -0.0167 | 0.0004 | 0.0198 | -0.1144 | 0.0894 | -0.2728 | 1.3653 |
| Sweden | 2000-01-28 | 2020-12-28 | 252 | 0.0002 | 0.0328 | -0.0193 | -0.0002 | 0.0212 | -0.1257 | 0.1002 | -0.0479 | 0.9074 |
| Switzerland | 2000-01-28 | 2020-12-28 | 252 | 0.0024 | 0.0285 | -0.0141 | -0.0000 | 0.0177 | -0.1075 | 0.1287 | 0.1584 | 1.9139 |
| United Kingdom | 2000-01-28 | 2020-12-28 | 252 | -0.0009 | 0.0250 | -0.0141 | 0.0001 | 0.0145 | -0.1069 | 0.0859 | -0.4006 | 1.7937 |

Note: The descriptive statistics for the change of nominal exchange rates are computed based on available data for each country. The column "Countries" presents the names of countries. "*" denotes the countries which joined the Euro Area after 1 January 1999. Estonia joined the Euro Area on 1 January 2011, Greece joined the Euro Area on 1 January 2001, Slovakia joined the Euro Area on 1 January 2009, and Slovenia joined the Euro Area on 1 January 2007. The columns "Start Date" and "Last Date" show the start date and the end date of the change of exchange rates in the empirical analysis for a given country. The column "Count" indicates the total number of observations for a given country. The columns "Mean" and "SD" present the mean and the standard deviation of total observations, respectively, for a given country. The columns "25%", "Median(50%)", and "75%" show the 25, the 50 (the median), and the 75 percentile of total observations, respectively, for a given country. The columns "Min" and "Max" show the minimum and the maximum values among all observations for a given country. The columns "Skw" indicates the skewness of data and the columns "Exc-Kurtosis" shows the excess kurtosis of data for a given country.

The descriptive statistics for changes in nominal exchange rates for emerging/developing countries are reported in Table 2.2. The table indicates that for almost all countries, skewness is negative; therefore, the data are skewed left. By skewed left, I mean that the left tail is long relative to the right tail. Based on skewness values, exchange rates data are not symmetric. In addition, Table 2.2 shows that for all countries, excess kurtosis is positive. For some countries, excess kurtosis is large because of extreme observations. A regime-switching process could lead to such extreme observations.

Table 2.2: Descriptive statistics for emerging/developing countries

| Countries | Start Date | Last Date | Count | Mean | SD | 25% | Median(50%) | 75% | Min | Max | Skw | Exc-Kurtosis |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Brazil | 2000-01-28 | 2020-12-28 | 252 | -0.0044 | 0.0510 | -0.0284 | 0.0021 | 0.0280 | -0.2537 | 0.1487 | -0.9046 | 2.8079 |
| Chile | 2000-01-28 | 2020-12-28 | 252 | -0.0013 | 0.0331 | -0.0211 | 0.0000 | 0.0208 | -0.1853 | 0.0789 | -0.7856 | 3.2056 |
| Colombia | 2000-01-28 | 2020-12-28 | 252 | -0.0024 | 0.0375 | -0.0231 | 0.0011 | 0.0186 | -0.1476 | 0.1120 | -0.3929 | 1.4649 |
| Hungary | 2000-01-28 | 2020-12-28 | 252 | -0.0006 | 0.0402 | -0.0244 | 0.0047 | 0.0247 | -0.2008 | 0.1075 | -1.1652 | 4.3284 |
| India | 2000-01-28 | 2020-12-28 | 252 | -0.0021 | 0.0213 | -0.0115 | -0.0006 | 0.0075 | -0.0856 | 0.0695 | -0.2895 | 2.6413 |
| Mexico | 2000-01-28 | 2020-12-28 | 252 | -0.0039 | 0.0326 | -0.0186 | -0.0007 | 0.0164 | -0.2048 | 0.0779 | -1.6526 | 8.4629 |
| Peru | 2000-01-28 | 2020-12-28 | 252 | 0.0001 | 0.0142 | -0.0081 | 0.0009 | 0.0060 | -0.0479 | 0.0580 | 0.5394 | 3.0855 |
| Poland | 2000-01-28 | 2020-12-28 | 252 | 0.0004 | 0.0387 | -0.0208 | 0.0016 | 0.0257 | -0.1831 | 0.0995 | -0.9460 | 3.0396 |
| South Africa | 2000-01-28 | 2020-12-28 | 252 | -0.0035 | 0.0486 | -0.0343 | 0.0018 | 0.0326 | -0.2009 | 0.1148 | -0.6166 | 0.9959 |
| Tunisia | 2000-01-28 | 2020-12-28 | 252 | -0.0034 | 0.0215 | -0.0159 | -0.0032 | 0.0106 | -0.0954 | 0.0550 | -0.4491 | 1.6039 |
| Turkey | 2000-01-28 | 2020-12-28 | 252 | -0.0106 | 0.0499 | -0.0346 | -0.0032 | 0.0163 | -0.3083 | 0.1028 | -1.8522 | 8.5996 |

Note: The descriptive statistics for the change of nominal exchange rates are computed based on available data for each country. The column "Countries" presents the names of countries, the columns "Start Date" and "Last Date" show the start date and the end date of the change of exchange rates in the empirical analysis for a given country. The column "Count" indicates the total number of observations for a given country. The columns "Mean" and "SD" present the mean and the standard deviation of total observations, respectively, for a given country. The columns "25%", "Median(50%)", and "75%" show the 25, the 50 (the median), and the 75 percentile of total observations, respectively, for a given country. The columns "Min" and "Max" show the minimum and the maximum values among all observations for a given country. The columns "Skw" indicates the skewness of data and the columns "Exc-Kurtosis" shows the excess kurtosis of data for a given country.

In the next section, I explain the methodologies to construct variables and models estimations process.

## 2.4 Methodology and Implementation

In this section, I explain how I construct variables and estimate models. In particular, I describe the following:

- Construction of Monthly Variables
- Overall Estimation Setup and Forecasting

- Performance Measures and Statistics

### 2.4.1 Construction of Monthly Variables

As a reminder, the goal is to forecast changes in exchange rates. To calculate $\Delta s_{t+h,h}$, I first take the logarithm of the exchange rate price level, $S_t$, for a given currency pair and denote it by $s_t$, where $S_t$ is expressed as the US dollars for a unit of foreign currency. Second, I compute the h-month-ahead changes in exchange rates, $\Delta s_{t+h,h} = s_{t+h} - s_t$.

The output gap ($y_t^{gap}$) is constructed by (1) using the seasonally adjusted output ($y_t$) and (2) applying Hodrick-Prescott (HP) detrending, calculated recursively, using data from periods prior to the forecast period as in Molodtsova and Papell (2009)[12].The VIX and SKEW indices are reported as daily series. Therefore, I convert daily frequency to monthly frequency. I do this by setting the last day of month price as the price of the given month. For example, the SKEW price on 31 January 1999 was 128.22; therefore, the SKEW price for January 1999 would be 128.22. Complete data descriptions are presented in Section 2.3.

### 2.4.2 Overall Estimation Setup and Forecasting

To forecast the $h$-month-ahead of change of exchange rates, I follow the steps below for each currency against the US dollar. Suppose the historical data set consists of a total number of $T$ observations, and let $h$ denotes forecasting horizons measured in months. I consider both short (horizon$=1$) and long (horizon$=12$) horizons.

**The First Step:** I estimate the non-behavioral models (Group A), described in Table 2.3, using a conventional approach in exchange rates forecasting literature, a fixed-length rolling window approach [see, e.g., Clark and West (2006), Clark and West (2007), Molodtsova and Papell (2009), and Rossi (2013)]. The rationale for using a fixed-length rolling window approach is there are changes over time, such as regime-switching or financial crises. The

[12] The smoothness parameter for HP filter is 14400 as in Molodtsova and Papell (2009).

fixed window allows me to adjust locally for those changes. I use ordinary least squares (OLS) regressions to estimate the linear models.[13]

Table 2.3: Non-behavioral models (Group A)

| **Models/Predictors** | **Specification** |
|---|---|
| PPP | $\Delta s_{t+h,h} = \alpha_0 + \alpha_1(\pi_t - \pi_t^*) + u_{t+h,h}$ |
| Oil Price | $\Delta s_{t+h,h} = \alpha_0 + \alpha_1 \Delta Oil_t + u_{t+h,h}$ |
| Commodity Price ($CP$) | $\Delta s_{t+h,h} = \alpha_0 + \alpha_1 \Delta CP_t + u_{t+h,h}$ |
| Real Commodity Price | $\Delta s_{t+h,h} = \alpha_0 + \alpha_1(\Delta CP_t - \pi_t) + u_{t+h,h}$ |
| Real Oil Price | $\Delta s_{t+h,h} = \alpha_0 + \alpha_1(\Delta Oil_t - \pi_t) + u_{t+h,h}$ |
| Terms of Trade | $\Delta s_{t+h,h} = \alpha_0 + \alpha_1 \Delta tot_t^* + u_{t+h,h}$ |
| UIRP | $\Delta s_{t+h,h} = \alpha_0 + \alpha_1(i_t - i_t^*) + u_{t+h,h}$ |
| UIRP augmented by Terms of Trade | $\Delta s_{t+h,h} = \alpha_0 + \alpha_1(i_t - i_t^*) + \alpha_2 \Delta tot_t^* + u_{t+h,h}$ |
| Extended UIRP | $\Delta s_{t+h,h} = \alpha_0 + \alpha_1(i_t - i_t^*) + \alpha_2(\pi_t - \pi_t^*) + u_{t+h,h}$ |
| Taylor Rule Model | $\Delta s_{t+h,h} = \alpha_0 + \alpha_1(\pi_t - \pi_t^*) + \alpha_2(y_t^{gap} - y_t^{*gap}) + u_{t+h,h}$ |
| Monetary Model with Flexible Prices | $\Delta s_{t+h,h} = \alpha_0 + \alpha_1(i_t - i_t^*) + \alpha_2(\Delta y_t - \Delta y_t^*) + \alpha_3(\Delta m_t - \Delta m_t^*) + u_{t+h,h}$ |

Notes: Table reports the non-behavioral models (Group A). $\Delta s_{t+h,h}$ is the $h$-month-ahead changes in exchange rates. PPP denotes the purchasing power parity model. $\pi_t^*$ is a foreign country inflation rate and $\pi_t$ is the US inflation rate. $\Delta Oil_t$ is the first difference of the logarithm of oil price and $\Delta CP_t$ is the first difference of the logarithm of the aggregate commodity price. $\Delta tot_t^*$ is the first difference of the logarithm of terms of trade index for a foreign country. $i_t^*$ is the interest rate of a foreign country and $i_t$ is the interest rate of the US. $y_t^{*gap}$ is the output gap of a foreign country and $y_t^{gap}$ is the output gap of the US. $\Delta y_t^*$ is the growth output rate for a foreign country and $\Delta y_t$ is the growth output rate for the US. $\Delta m_t^*$ is the growth money rate for a foreign country and $\Delta m_t$ is the growth money rate for the US. UIRP denotes the uncovered interest rate parity model.

The total sampling period for all countries is from Jan 2000 to Dec 2020. Using the fixed-length rolling window approach and starting from Jan 2000, I take the window of $T_1 = 60$ observations to estimate the forecasting model and use the estimated coefficients to form a forecast for the change of exchange rates. The 60 observations are close to 20% of the total observations. Then, I roll the sample period forward one observation, re-estimate the model, and form a forecast using the estimated coefficients. I keep rolling and repeating the steps above until the end of the sample period[14]. Depending on $h$, the total number of forecasts ($T_2$) will vary. More specifically, $T_2 = T - T_1$, where $h = 1$ and $T_2 = T - T_1 - 12$, where $h = 12$.

[13] I am grateful to Professor Andrei Semenov for his comments and inputs in this section, specifically 'The First Step and 'The Second Step' parts.

[14] Alternatively, a recursive (expanding) window approach can be used for estimation. The recursive method makes use of an increasing window to re-estimate coefficients, whereas the rolling approach makes use of a fixed-length window of data to re-estimate coefficients.

Using standard notation, I denote a single-equation, lagged factor model as follows:

$$\Delta s_{t+h,h} = \alpha_{0t} + \alpha_t' f_t + u_{t+h,h}, \tag{3.1.1}$$

where $\alpha_{0t}$ is a time-varying constant, $\alpha_t'$ is a vector of time-varying coefficients, and $f_t$ is a vector of predictors. Using equation (3.1.1), the $h$-month-ahead forecast for the change of exchange rates is as follows:

$$E_t(\Delta s_{t+h,h}) = \hat{\alpha}_{0t} + \hat{\alpha}_t' f_t. \tag{3.1.2}$$

As mentioned in Section 2.1, I use White Noise (WN) as the benchmark model. For the WN model, the $h$-month-ahead forecast for changes in exchange rates is as follows:

$$E_t(\Delta s_{t+h,h}) = 0. \tag{3.1.3}$$

Using out-of-sample statistics described in Section 2.4.3, I choose the 'best' forecasting model among non-behavioral models (Group A) for a given country and horizon and call it 'Model A'.

**The Second Step:** I extend the 'Model A' by adding different combinations of investor sentiment variables (behavioral factors), presented in Table 2.4, and create behavioral models for each country and horizon. As I explained before, there are up to 5 behavioral factors for each country based on data availability. I start by adding one factor, then two factors,..., and finally, all the five factors together (where data are available). This approach creates up to 31 behavioral models, each of which is an extension of Model A for the given country and horizon.

Table 2.4: Different combinations of behavioral factors

| **Combination's Names** | **Behavioral Factors** |
|---|---|
| VIX | $\Delta VIX_t$ |
| SKEW | $\Delta SKEW_t$ |
| Business Confidence | $\Delta BCI_t - \Delta BCI_t^*$ |
| Consumer Confidence | $\Delta CCI_t - \Delta CCI_t^*$ |
| Composite Leading Indicator | $\Delta CLI_t - \Delta CLI_t^*$ |
| Investor Sentiment (1) | $\Delta SKEW_t$ & $\Delta VIX_t$ |
| Investor Sentiment (2) | $\Delta VIX_t$ & $(\Delta CLI_t - \Delta CLI_t^*)$ |
| Investor Sentiment (3) | $\Delta VIX_t$ & $(\Delta CCI_t - \Delta CCI_t^*)$ |
| Investor Sentiment (4) | $\Delta VIX_t$ & $(\Delta BCI_t - \Delta BCI_t^*)$ |
| Investor Sentiment (5) | $\Delta SKEW_t$ & $(\Delta CLI_t - \Delta CLI_t^*)$ |
| Investor Sentiment (6) | $\Delta SKEW_t$ & $(\Delta CCI_t - \Delta CCI_t^*)$ |
| Investor Sentiment (7) | $\Delta SKEW_t$ & $(\Delta BCI_t - \Delta BCI_t^*)$ |
| Investor Sentiment (8) | $(\Delta CLI_t - \Delta CLI_t^*)$ & $(\Delta CCI_t - \Delta CCI_t^*)$ |
| Investor Sentiment (9) | $(\Delta CLI_t - \Delta CLI_t^*)$ & $(\Delta BCI_t - \Delta BCI_t^*)$ |
| Investor Sentiment (10) | $(\Delta BCI_t - \Delta BCI_t^*)$ & $(\Delta CCI_t - \Delta CCI_t^*)$ |
| Investor Sentiment (11) | $(\Delta CLI_t - \Delta CLI_t^*)$ & $\Delta VIX_t$ & $\Delta SKEW_t$ |
| Investor Sentiment (12) | $(\Delta CCI_t - \Delta CCI_t^*)$ & $\Delta VIX_t$ & $\Delta SKEW_t$ |
| Investor Sentiment (12) | $(\Delta CCI_t - \Delta CCI_t^*)$ & $\Delta VIX_t$ & $\Delta SKEW_t$ |
| Investor Sentiment (13) | $(\Delta BCI_t - \Delta BCI_t^*)$ & $\Delta VIX_t$ & $\Delta SKEW_t$ |
| Investor Sentiment (14) | $(\Delta BCI_t - \Delta BCI_t^*)$ & $(\Delta CLI_t - \Delta CLI_t^*)$ & $\Delta VIX_t$ |
| Investor Sentiment (15) | $(\Delta CCI_t - \Delta CCI_t^*)$ & $(\Delta BCI_t - \Delta BCI_t^*)$ & $\Delta VIX_t$ |
| Investor Sentiment (16) | $(\Delta CCI_t - \Delta CCI_t^*)$ & $(\Delta CLI_t - \Delta CLI_t^*)$ & $\Delta VIX_t$ |
| Investor Sentiment (17) | $(\Delta CLI_t - \Delta CLI_t^*)$ & $(\Delta BCI_t - \Delta BCI_t^*)$ & $\Delta SKEW_t$ |
| Investor Sentiment (18) | $(\Delta CLI_t - \Delta CLI_t^*)$ & $(\Delta CCI_t - \Delta CCI_t^*)$ & $\Delta SKEW_t$ |
| Investor Sentiment (19) | $(\Delta CCI_t - \Delta CCI_t^*)$ & $(\Delta BCI_t - \Delta BCI_t^*)$ & $\Delta SKEW_t$ |
| Investor Sentiment (20) | $(\Delta BCI_t - \Delta BCI_t^*)$ & $(\Delta CCI_t - \Delta CCI_t^*)$ & $(\Delta CLI_t - \Delta CLI_t^*)$ |
| Investor Sentiment (21) | $(\Delta CCI_t - \Delta CCI_t^*)$ & $(\Delta CLI_t - \Delta CLI_t^*)$ & $\Delta VIX_t$ & $\Delta SKEW_t$ |
| Investor Sentiment (22) | $(\Delta CLI_t - \Delta CLI_t^*)$ & $(\Delta BCI_t - \Delta BCI_t^*)$ & $\Delta VIX_t$ & $\Delta SKEW_t$ |
| Investor Sentiment (23) | $(\Delta CCI_t - \Delta CCI_t^*)$ & $(\Delta BCI_t - \Delta BCI_t^*)$ & $\Delta VIX_t$ & $\Delta SKEW_t$ |
| Investor Sentiment (24) | $(\Delta CLI_t - \Delta CLI_t^*)$ & $(\Delta BCI_t - \Delta BCI_t^*)$ & $(\Delta CCI_t - \Delta CCI_t^*)$ & $\Delta VIX_t$ |
| Investor Sentiment (25) | $(\Delta CCI_t - \Delta CCI_t^*)$ & $(\Delta CLI_t - \Delta CLI_t^*)$ & $(\Delta BCI_t - \Delta BCI_t^*)$ & $\Delta SKEW_t$ |
| Investor Sentiment (26) | $(\Delta BCI_t - \Delta BCI_t^*)$ & $(\Delta CCI_t - \Delta CCI_t^*)$ & $(\Delta CLI_t - \Delta CLI_t^*)$ & $\Delta VIX_t$ & $\Delta SKEW_t$ |

Notes: The table reports 31 different combinations of the investor sentiment variables. The column "Combination's Names" presents each combination's name and the column "Behavioral Factors" indicates what investor sentiment indices are included in each specification. $\Delta VIX_t$ is the first difference of the logarithm of Cboe Volatility Index at time $t$ and $\Delta SKEW_t$ is the first difference of logarithm of Cboe SKEW Index at time $t$. $\Delta BCI_t^*$ is the first difference of the logarithm of business confidence index for a foreign country and $\Delta BCI_t$ is the first difference of the logarithm of business confidence index for the US at time $t$. $\Delta CLI_t^*$ is the first difference of the logarithm of composite leading indicator index for a foreign country and $\Delta CLI_t$ is the first difference of the logarithm of composite leading indicator index for the US at time $t$. $\Delta CCI_t^*$ is the first difference of the logarithm of consumer confidence index for a foreign country and $\Delta CCI_t$ is the first difference of the logarithm of consumer confidence index for the US at time $.t$.

I follow the same approach as in the 'The First Step' section to estimate behavioral models and form forecasts. There are up to 32 forecasting models for each country and each horizon: one non-behavioral model (Model A) plus up to 31 behavioral models. Using out-of-sample statistics, I select the 'best' forecasting model among all 32 possible models and present it for each country and horizon in Section 2.5.

**The Third Step:** I extend all macroeconomic (Group A) models, and create a large number of behavioral models for each country and horizon. I follow the same approach as in the 'The First Step' section to estimate the behavioral models and form forecasts. There are up to 352 forecasting models for each country and each horizon based on data availability. Using the out-of-sample statistics, I select the 'best' forecasting model among all possible models and present it for each country and horizon in Section 2.5.

In the next section, I explain out-of-sample statistics used for forecast evaluation.

### 2.4.3 Performance Measures and Statistics

To evaluate the out-of-sample performance of forecasting models with the White Noise (WN) (benchmark) model, I use various statistics, including the ratio of Mean Squared Prediction Error ($\text{MSPE}_{ratio}$), Diebold-Mariano (DM), and Pesaran-Timmermann (PT) statistic. In addition to statistical forecast evaluation measures, I use the Sharpe-ratio statistic to evaluate investment strategies, using the expected exchange rate changes to calculate investment returns, discussed below. Explanations and motivations for using these statistics are provided in Section 1.4.3.

I use equations (1.4.19), (1.4.20), and (1.4.21) to calculate $\text{MSPE}_{WN}$ (the MSPE statistic for the WN model), $\text{MSPE}_{M}$ (the MSPE statistic for a forecasting model), and $\text{MSPE}_{ratio}$ (the ratio of MSPE for a forecasting model to MSPE of the benchmark). $\text{MSPE}_{ratio} < 1$ implies that a forecasting model provides better out-of-sample performance compared to the benchmark model.

I use equation (1.4.22) to calculate $\hat{d}_{t+h}$. The DM statistic can be calculated by regression of the loss differential ($\hat{d}_{t+h}$) on a constant using HAC approaches to correct standard errors for autocorrelation and heteroskedasticity. I calculate the p-value for the DM statistic. I use equation (1.4.23) to calculate the PT statistic. A value of $\hat{p}$ statistically larger than 0.5 shows a better forecasting performance than the benchmark model. I calculate the p-value associated with the PT statistic. I use the 10% level of significance.

In addition to statistical evaluation statistics discussed above, I examine the forecasting models' economic value. I follow a simple investment strategy that buys (sells) the US dollar against one unit of a foreign currency when a factor model forecasts the foreign currency appreciation (depreciation). I calculate returns, $r_t$, for an investment strategy using equation (1.4.24) and I regress $r_t$ on a constant and use the HAC approach to correct the standard errors for autocorrelation and heteroskedasticity. Using the estimated constant and an adjusted standard error, I calculate an annualized Sharpe-ratio. I also report the annualized average return in percentage for each factor model for a given country.

In the next section, I explain the empirical results in this essay.

## 2.5 Empirical Results

This section describes the out-of-sample forecasting performance of factor models from both statistical and economic perspectives. I examine whether factor models can outperform the benchmark model to predict the direction of changes in exchange rates. I report the empirical results in both horizon=1 and horizon=12.

Table 2.5 reports the summary results for horizon=1 and the median of statistics for the best macroeconomic model, the best behavioral model (i.e., the best macroeconomic plus behavioral factors), and the best model among all available (macroeconomic and behavioral) models within each country group (advanced and emerging/developing). The results indicate that, on average, macroeconomic models have proportions of correct direction of changes greater than 0.5, and the results are statistically significant at the 10% significance level at horizon=1 for the emerging/developing group. A value of $\hat{p}$ above 0.5 and the p-value less than 0.1 indicates a better forecasting performance of the underlying model than the WN model which predicts changes in exchange rates have an equal chance to go up or down. However, on average, no macroeconomic model in either group has a $\text{MSPE}_{ratio}$ less than one and positive and statistically significant DM statistics.

Table 2.5: Summary results in horizon=1

| WEO Country Group | Definitions | MSPE_ratio | DM (t-stats) | DM (p-value) | $\hat{p}$ | PT(t-stats) | PT (p-value) | Avg. Return (%) | Sharpe_ratio |
|---|---|---|---|---|---|---|---|---|---|
| | Best Macro Model | 1.0356 | -0.9795 | 0.1646 | 0.5340 | 0.9406 | 0.1734 | 1.7242 | 0.1920 |
| Advanced | Best Behavioral Model | 1.0870 | -1.6481 | 0.0508 | 0.5681 | 1.8813 | 0.0303 | 2.6898 | 0.2793 |
| | Best Model | 1.1264 | -1.6416 | 0.0515 | 0.5759 | 2.0984 | 0.0179 | 3.0032 | 0.3438 |
| Emerging/Developing | Best Macro Model | 1.0114 | -0.3184 | 0.2170 | 0.5707 | 1.9537 | 0.0254 | 2.1977 | 0.2820 |
| | Best Behavioral Model | 1.0362 | -1.0381 | 0.1503 | 0.5654 | 1.8089 | 0.0352 | 4.0571 | 0.4432 |
| | Best Model | 1.0358 | -1.0591 | 0.1454 | 0.5707 | 1.9537 | 0.0254 | 5.4322 | 0.4786 |

Note: The column “WEO Country Group” shows the World Economic Outlook (WEO) countries’ classifications. The column “Definitions” defines the underlying best forecasting model. The column “MSPE_ratio” indicates the median of the $MSPE_{ratio}$ for the best macro models, the best behavioral models (i.e., the best macro model plus behavioral factors), and the best models among all macro and behavioral models within each country group. The column “DM(t-stats)” shows the median of DM statistics for the best macro models, the best behavioral models, and the best models among all macro and behavioral models within each country group. The column “DM(p-value)” shows the p-values associated with the DM statistics. The column “$\hat{p}$” indicates the median of proportions of correct predictions of direction for the best macro models, the best behavioral models, and the best models among all macro and behavioral models within each country group. The column “PT(t-stats)” indicates the median of the Pesaran-Timmermann (PT) statistics for the best macro models, the best behavioral models, and the best models among all macro and behavioral models within each country group and the column “PT(p-value)” shows the p-values associated with the PT statistics. All reported test statistics are one-sided. The significance level is 10%. The column “Avg. Return (%)” shows the annualized average returns in percentage, and the column “ Sharp_ratio ” shows the annualized Sharp-ratio.

Table 2.5 indicates that, on average, adding investor sentiment variables (behavioral factors) improves the out-of-sample performance of the best macroeconomic models in both groups. Moreover, the findings show, on average, the ‘best model’ among all macroeconomic and behavioral models can correctly predict the direction of change in exchange rates in both advanced and emerging/developing countries at horizon=1.

Tables 2.6 and 2.7 report the detailed empirical results for advanced countries at horizon=1. The results show the best macroeconomic model, the best behavioral model, and best model among all available models for a given country. For example, for Canada, the model that included changes in the real oil price is the best macroeconomic forecasting model. The model that included changes in the real oil price, changes in SKEW, and the difference between the Canada and US business confidence indices (in changes) is the best behavioral model. Moreover, the Taylor Rule model that includes changes in SKEW and the difference between the Canada and US business confidence indices (in changes) is the best forecasting model among all available models in this essay.

Table 2.6: Empirical results for advanced countries in horizon=1

| Countries | Models | MSPE_ratio | DM (t-stats) | DM (p-value) | $\hat{p}$ | PT(t-stats) | PT (p-value) | Avg. Return (%) | Sharpe_ratio |
|---|---|---|---|---|---|---|---|---|---|
| Austria | PPP | 1.0217 | -0.8267 | 0.2047 | 0.5340 | 0.9406 | 0.1734 | 2.0069 | 0.2153 |
| Austria | PPP+Investor Sentiment (20) | 1.2015 | -2.2310 | 0.0134 | 0.5864 | 2.3878 | 0.0085 | 3.2902 | 0.3731 |
| Austria | UIRP+Composite Leading Indicator | 1.0882 | -1.4724 | 0.0713 | 0.5916 | 2.5325 | 0.0057 | 2.8699 | 0.3463 |
| Belgium | Terms of Trade | 1.0174 | -0.6667 | 0.2529 | 0.5236 | 0.6512 | 0.2575 | 1.8554 | 0.1973 |
| Belgium | Terms of Trade+SKEW | 1.0225 | -0.6705 | 0.2517 | 0.5707 | 1.9537 | 0.0254 | 4.8040 | 0.4474 |
| Belgium | Terms of Trade+SKEW | 1.0225 | -0.6705 | 0.2517 | 0.5707 | 1.9537 | 0.0254 | 4.8040 | 0.4474 |
| Canada | Real Oil Price | 1.0186 | -0.6238 | 0.2668 | 0.5769 | 2.0755 | 0.0190 | 4.5250 | 0.5546 |
| Canada | Real Oil Price+Investor Sentiment (7) | 1.0488 | -1.1075 | 0.1348 | 0.5934 | 2.5202 | 0.0059 | 4.4299 | 0.4695 |
| Canada | Taylor Rule Model+Investor Sentiment (7) | 1.1295 | -2.4779 | 0.0071 | 0.6209 | 3.2615 | 0.0006 | 6.3567 | 0.7857 |
| Czech Republic | Extended UIRP | 1.0563 | -1.8200 | 0.0352 | 0.5602 | 1.6642 | 0.0480 | 1.5994 | 0.1392 |
| Czech Republic | Extended UIRP+Composite Leading Indicator | 1.0843 | -2.4304 | 0.0080 | 0.5497 | 1.3748 | 0.0846 | 2.5209 | 0.2247 |
| Czech Republic | UIRP+SKEW | 1.0516 | -1.5511 | 0.0613 | 0.5602 | 1.6642 | 0.0480 | 1.4628 | 0.1147 |
| Denmark | UIRP | 1.0771 | -1.9832 | 0.0244 | 0.5183 | 0.5065 | 0.3063 | -1.4867 | -0.1660 |
| Denmark | UIRP+Consumer Confidence | 1.0866 | -2.2663 | 0.0123 | 0.5288 | 0.7959 | 0.2130 | 0.8745 | 0.0984 |
| Denmark | Real Oil Price+Investor Sentiment (15) | 1.1717 | -2.1239 | 0.0175 | 0.5497 | 1.3748 | 0.0846 | 2.9801 | 0.3304 |
| Estonia | Terms of Trade | 0.9826 | 0.4706 | 0.3192 | 0.5812 | 2.2431 | 0.0124 | 4.6169 | 0.4931 |
| Estonia | Terms of Trade+Investor Sentiment (8) | 0.9761 | 0.5010 | 0.3085 | 0.5916 | 2.5325 | 0.0057 | 7.0074 | 0.7684 |
| Estonia | Terms of Trade+Investor Sentiment (8) | 0.9761 | 0.5010 | 0.3085 | 0.5916 | 2.5325 | 0.0057 | 7.0074 | 0.7684 |
| Finland | PPP | 1.0367 | -1.0275 | 0.1527 | 0.5288 | 0.7959 | 0.2130 | 1.7976 | 0.1970 |
| Finland | PPP+Business Confidence | 1.0504 | -1.2151 | 0.1129 | 0.5393 | 1.0854 | 0.1389 | 1.7918 | 0.1935 |
| Finland | Terms of Trade+Investor Sentiment (2) | 1.2361 | -1.7705 | 0.0391 | 0.5759 | 2.0984 | 0.0179 | 1.3373 | 0.1684 |
| France | UIRP augmented by Terms of Trade | 1.0505 | -1.3246 | 0.0934 | 0.5393 | 1.0854 | 0.1389 | 1.4743 | 0.1698 |
| France | UIRP augmented by Terms of Trade+Investor Sentiment (7) | 1.1128 | -2.4965 | 0.0067 | 0.5550 | 1.5195 | 0.0643 | 0.3844 | 0.0398 |
| France | Real Commodity Price+Investor Sentiment (9) | 1.1746 | -1.5565 | 0.0606 | 0.5759 | 2.0984 | 0.0179 | 3.0262 | 0.3459 |
| Germany | Terms of Trade | 1.0203 | -0.7580 | 0.2247 | 0.5550 | 1.5195 | 0.0643 | 4.1219 | 0.4856 |
| Germany | Terms of Trade+Composite Leading Indicator | 1.0368 | -0.6865 | 0.2466 | 0.5916 | 2.5325 | 0.0057 | 5.7222 | 0.5999 |
| Germany | Real Commodity Price+Consumer Confidence | 1.0671 | -1.7412 | 0.0416 | 0.6126 | 3.1114 | 0.0009 | 3.8064 | 0.4927 |
| Greece | UIRP | 1.0508 | -1.8886 | 0.0302 | 0.5183 | 0.5065 | 0.3063 | 1.8340 | 0.2064 |
| Greece | UIRP+Investor Sentiment (26) | 1.2281 | -2.4837 | 0.0069 | 0.5445 | 1.2301 | 0.1093 | 1.1581 | 0.1393 |
| Greece | Real Commodity Price+Investor Sentiment (14) | 1.2015 | -1.8833 | 0.0306 | 0.5707 | 1.9537 | 0.0254 | 2.5933 | 0.2834 |
| Iceland | Extended UIRP | 1.1198 | -0.8973 | 0.1853 | 0.5340 | 0.9406 | 0.1734 | 0.6781 | 0.0434 |
| Iceland | Extended UIRP+Composite Leading Indicator | 1.2077 | -1.5896 | 0.0568 | 0.5916 | 2.5325 | 0.0057 | 6.4058 | 0.4147 |
| Iceland | Commodity Price+Investor Sentiment (11) | 1.0602 | -0.4365 | 0.3315 | 0.5916 | 2.5325 | 0.0057 | 8.5818 | 0.6109 |
| Ireland | Terms of Trade | 1.0275 | -0.9314 | 0.1764 | 0.5288 | 0.7959 | 0.2130 | 0.9604 | 0.0967 |
| Ireland | Terms of Trade+Investor Sentiment (9) | 1.0817 | -2.0418 | 0.0213 | 0.5497 | 1.3748 | 0.0846 | 0.0812 | 0.0086 |
| Ireland | Real Commodity Price+Investor Sentiment (8) | 1.1353 | -2.4784 | 0.0070 | 0.5497 | 1.3748 | 0.0846 | 0.4299 | 0.0456 |
| Israel | Commodity Price | 1.0274 | -0.7314 | 0.2327 | 0.5707 | 1.9537 | 0.0254 | 2.7311 | 0.3810 |
| Israel | Commodity Price+Investor Sentiment (9) | 1.2564 | -1.7041 | 0.0450 | 0.5812 | 2.2431 | 0.0124 | 2.0726 | 0.2792 |
| Israel | Commodity Price+Investor Sentiment (9) | 1.2564 | -1.7041 | 0.0450 | 0.5812 | 2.2431 | 0.0124 | 2.0726 | 0.2792 |

Note: The column "Countries" shows the names of countries. The column "Models" defines the underlying model. The column "MSPE_ratio" indicates the $\text{MSPE}_{ratio}$ for the best macro models, the best behavioral models (i.e., the best macro model plus the behavioral factors), and the best models among all macro and behavioral models for a given country. The column "DM(t-stats)" shows the DM statistics for the best macro models, the best behavioral models, and the best models among all macro and behavioral models for each country. The column "DM(p-value)" shows the p-values associated with the DM statistics. The column "$\hat{p}$" indicates the proportions of correct predictions of direction for the best macro models, the best behavioral models, and the best models among all macro and behavioral models for each country. The column "PT(t-stats)" shows the Pesaran-Timmermann (PT) statistics for the best macro models, the best behavioral models, and the best models among all macro and behavioral models for each country and the column "PT(p-value)" shows the p-values associated with the PT statistics. All reported test statistics are one-sided. The significance level is 10%. The column "Avg. Return (%)" shows the annualized average returns in percentage, and the column " Sharp_ratio " shows the annualized Sharp-ratio.

Table 2.7: Empirical results for advanced countries in horizon=1, continued

| Countries | Models | MSPE_ratio | DM (t-stats) | DM (p-value) | $\hat{p}$ | PT(t-stats) | PT (p-value) | Avg. Return (%) | Sharpe_ratio |
|---|---|---|---|---|---|---|---|---|---|
| Italy | Terms of Trade | 1.0176 | -0.6326 | 0.2639 | 0.5759 | 2.0984 | 0.0179 | 3.7708 | 0.4183 |
| Italy | Terms of Trade+SKEW | 1.0232 | -0.6204 | 0.2679 | 0.5969 | 2.6772 | 0.0037 | 5.0365 | 0.5452 |
| Italy | Terms of Trade+SKEW | 1.0232 | -0.6204 | 0.2679 | 0.5969 | 2.6772 | 0.0037 | 5.0365 | 0.5452 |
| Japan | UIRP | 1.0862 | -1.8427 | 0.0335 | 0.5288 | 0.7959 | 0.2130 | 1.2232 | 0.1278 |
| Japan | UIRP+VIX | 1.0875 | -1.5920 | 0.0565 | 0.5340 | 0.9406 | 0.1734 | 1.7830 | 0.2153 |
| Japan | Extended UIRP+Investor Sentiment (21) | 1.1841 | -2.0256 | 0.0221 | 0.5602 | 1.6642 | 0.0480 | 1.8581 | 0.2052 |
| Korea | Terms of Trade | 1.0321 | -0.6779 | 0.2493 | 0.5445 | 1.2301 | 0.1093 | 0.9041 | 0.0889 |
| Korea | Terms of Trade+Investor Sentiment (21) | 1.2015 | -1.8212 | 0.0351 | 0.5864 | 2.3878 | 0.0085 | 4.9513 | 0.5559 |
| Korea | Terms of Trade+Investor Sentiment (21) | 1.2015 | -1.8212 | 0.0351 | 0.5864 | 2.3878 | 0.0085 | 4.9513 | 0.5559 |
| Luxembourg | PPP | 1.0345 | -1.0402 | 0.1498 | 0.5393 | 1.0854 | 0.1389 | 2.2651 | 0.2589 |
| Luxembourg | PPP+SKEW | 1.0447 | -1.0026 | 0.1587 | 0.5131 | 0.3618 | 0.3588 | 2.1911 | 0.2228 |
| Luxembourg | PPP | 1.0345 | -1.0402 | 0.1498 | 0.5393 | 1.0854 | 0.1389 | 2.2651 | 0.2589 |
| Netherlands | Taylor Rule Model | 1.0671 | -1.3596 | 0.0878 | 0.5340 | 0.9406 | 0.1734 | 2.3792 | 0.2634 |
| Netherlands | Taylor Rule Model+Investor Sentiment (8) | 1.1288 | -1.9319 | 0.0274 | 0.5550 | 1.5195 | 0.0643 | 1.6800 | 0.1892 |
| Netherlands | UIRP+Consumer Confidence | 1.0672 | -1.7203 | 0.0435 | 0.5864 | 2.3878 | 0.0085 | 4.8299 | 0.5829 |
| Norway | Real Oil Price | 1.0217 | -0.8050 | 0.2109 | 0.5602 | 1.6642 | 0.0480 | 3.8694 | 0.3573 |
| Norway | Real Oil Price+Investor Sentiment (17) | 1.5362 | -1.1407 | 0.1277 | 0.5654 | 1.8089 | 0.0352 | 3.5511 | 0.2794 |
| Norway | Oil Price+Investor Sentiment (17) | 1.5463 | -1.1396 | 0.1279 | 0.5707 | 1.9537 | 0.0254 | 4.2401 | 0.3418 |
| Portugal | PPP | 1.0379 | -1.4538 | 0.0738 | 0.5183 | 0.5065 | 0.3063 | 0.5800 | 0.0684 |
| Portugal | PPP+Investor Sentiment (17) | 1.1233 | -1.9789 | 0.0246 | 0.5812 | 2.2431 | 0.0124 | 4.1153 | 0.4123 |
| Portugal | PPP+Investor Sentiment (17) | 1.1233 | -1.9789 | 0.0246 | 0.5812 | 2.2431 | 0.0124 | 4.1153 | 0.4123 |
| Slovakia | Real Commodity Price | 1.0513 | -1.1214 | 0.1318 | 0.5497 | 1.3748 | 0.0846 | 1.1897 | 0.1037 |
| Slovakia | Real Commodity Price+Investor Sentiment (9) | 1.1337 | -1.8895 | 0.0302 | 0.5759 | 2.0984 | 0.0179 | 2.1189 | 0.1905 |
| Slovakia | PPP+Investor Sentiment (20) | 1.1415 | -1.5790 | 0.0580 | 0.5759 | 2.0984 | 0.0179 | 2.3655 | 0.2074 |
| Slovenia | Real Commodity Price | 1.0613 | -1.5953 | 0.0562 | 0.5340 | 0.9406 | 0.1734 | 0.5348 | 0.0616 |
| Slovenia | Real Commodity Price+Investor Sentiment (24) | 1.1449 | -1.5507 | 0.0613 | 0.5759 | 2.0984 | 0.0179 | 2.8587 | 0.3073 |
| Slovenia | Real Commodity Price+Investor Sentiment (24) | 1.1449 | -1.5507 | 0.0613 | 0.5759 | 2.0984 | 0.0179 | 2.8587 | 0.3073 |
| Spain | Real Commodity Price | 1.0630 | -1.6377 | 0.0516 | 0.5183 | 0.5065 | 0.3063 | 0.7880 | 0.0889 |
| Spain | Real Commodity Price+Investor Sentiment (24) | 1.1933 | -2.7193 | 0.0036 | 0.5497 | 1.3748 | 0.0846 | 0.8933 | 0.1071 |
| Spain | Terms of Trade+Investor Sentiment (24) | 1.1471 | -2.7082 | 0.0037 | 0.5602 | 1.6642 | 0.0480 | 0.9553 | 0.1036 |
| Sweden | Oil Price | 1.0176 | -0.6011 | 0.2742 | 0.5497 | 1.3748 | 0.0846 | 4.0603 | 0.4076 |
| Sweden | Oil Price+Investor Sentiment (19) | 1.0624 | -1.2256 | 0.1109 | 0.5812 | 2.2431 | 0.0124 | 4.8942 | 0.4245 |
| Sweden | Oil Price+Investor Sentiment (19) | 1.0624 | -1.2256 | 0.1109 | 0.5812 | 2.2431 | 0.0124 | 4.8942 | 0.4245 |
| Switzerland | Terms of Trade | 1.0454 | -1.7570 | 0.0403 | 0.5183 | 0.5065 | 0.3063 | 1.0406 | 0.1051 |
| Switzerland | Terms of Trade+Business Confidence | 1.0648 | -2.2814 | 0.0118 | 0.5236 | 0.6512 | 0.2575 | 1.1716 | 0.1348 |
| Switzerland | Commodity Price+VIX | 1.0852 | -1.7208 | 0.0435 | 0.5393 | 1.0854 | 0.1389 | 1.9231 | 0.2000 |
| United Kingdom | Oil Price | 0.9843 | 0.3346 | 0.3692 | 0.5333 | 0.8944 | 0.1855 | 1.6507 | 0.1870 |
| United Kingdom | Oil Price+Investor Sentiment (4) | 1.0202 | -0.2199 | 0.4131 | 0.5500 | 1.3416 | 0.0899 | 5.2122 | 0.4968 |
| United Kingdom | Oil Price+Investor Sentiment (4) | 1.0202 | -0.2199 | 0.4131 | 0.5500 | 1.3416 | 0.0899 | 5.2122 | 0.4968 |

Note: The column "Countries" shows the names of countries. The column "Models" defines the underlying model. The column "MSPE_ratio" indicates the $\text{MSPE}_{ratio}$ for the best macro models, the best behavioral models (i.e., the best macro model plus the behavioral factors), and the best models among all macro and behavioral models for a given country. The column "DM(t-stats)" shows the DM statistics for the best macro models, the best behavioral models, and the best models among all macro and behavioral models for each country. The column "DM(p-value)" shows the p-values associated with the DM statistics. The column "$\hat{p}$" indicates the proportions of correct predictions of direction for the best macro models, the best behavioral models, and the best models among all macro and behavioral models for each country. The column "PT(t-stats)" shows the Pesaran-Timmermann (PT) statistics for the best macro models, the best behavioral models, and the best models among all macro and behavioral models for each country and the column "PT(p-value)" shows the p-values associated with the PT statistics. All reported test statistics are one-sided. The significance level is 10%. The column "Avg. Return (%)" shows the annualized average returns in percentage, and the column " Sharp_ratio " shows the annualized Sharp-ratio.

The findings indicate the model that included the terms of trade index (in changes) is the best macroeconomic model for 27% of advanced countries in terms of the PT statistic. The results are consistent with the literature that changes in oil and commodity prices (both in real and nominal terms) are not successful in predicting point forecasts. However, the findings in this essay show that they correctly forecast the direction of changes in exchange rates for 31% of advanced countries. Except for Estonia and the United Kingdom, none of countries has a forecasting model with a $\text{MSPE}_{ratio}$ less than one.

This essay shows that adding investor sentiment variables (behavioral factors) improves the macroeconomic models' ability to forecast the direction of exchange rates changes correctly for 92% of advanced countries. Moreover, the findings highlight the advantage of using investor sentiment variable in forecasting models. Except for Luxembourg, the 'best' forecasting model based on PT statistics includes macroeconomic and behavioral factors for all advanced countries. Furthermore, the empirical evidence shows that investment strategies based on the behavioral models have positive and considerable annualized average returns and the annualized Sharpe-ratio values for most advanced countries, including Canada, Estonia, Iceland, Korea, and Norway.

Table 2.8 shows the results for emerging/developing countries at horizon=1. Similar to the results for advanced countries, adding behavioral factors to macroeconomic models improves the macroeconomic models' ability to correctly predict the direction of changes in exchange rates in emerging/developing countries. The findings are consistent with the literature that the PPP model has no predictive ability in point forecasts at the short horizon. However, this study shows that the PPP model successfully predicts the direction of changes in exchange rates for 45% of countries. Moreover, the results show the PPP model for India and Turkey has a $\text{MSPE}_{ratio}$ less than one. Except for Turkey, DM statistics are not statistically significant in any cases. The empirical evidence indicates that the 'best' forecasting models simultaneously include macroeconomic and investor sentiment variables, except for Columbia and India. In addition, the findings suggest that investment strategies based on behavioral models have positive and notable annualized average returns and annualized Sharpe-ratio values for all emerging/developing countries, except for Colombia.

Table 2.8: Empirical results for emerging/developing countries in horizon=1

| Countries | Models | MSPE_ratio | DM (t-stats) | DM (p-value) | $\hat{p}$ | PT(t-stats) | PT (p-value) | Avg. Return (%) | Sharpe_ratio |
|---|---|---|---|---|---|---|---|---|---|
| Brazil | Real Oil Price | 1.0015 | -0.0418 | 0.4833 | 0.5654 | 1.8089 | 0.0352 | 7.3046 | 0.5191 |
| Brazil | Real Oil Price+Business Confidence | 1.0438 | -1.0860 | 0.1394 | 0.5654 | 1.8089 | 0.0352 | 7.2122 | 0.5594 |
| Brazil | Real Commodity Price+Business Confidence | 1.0274 | -0.6595 | 0.2552 | 0.5707 | 1.9537 | 0.0254 | 9.4974 | 0.7047 |
| Chile | PPP | 1.0647 | -1.4951 | 0.0683 | 0.5079 | 0.2171 | 0.4141 | -1.5862 | -0.1243 |
| Chile | PPP+Investor Sentiment (5) | 1.0576 | -0.9876 | 0.1623 | 0.5550 | 1.5195 | 0.0643 | 2.4939 | 0.1813 |
| Chile | Commodity Price+Composite Leading Indicator | 1.0606 | -1.0591 | 0.1454 | 0.5654 | 1.8089 | 0.0352 | 3.8840 | 0.3188 |
| Colombia | PPP | 1.0441 | -1.6714 | 0.0481 | 0.5707 | 1.9537 | 0.0254 | -0.5341 | -0.0451 |
| Colombia | PPP+SKEW | 1.0618 | -2.0724 | 0.0198 | 0.5393 | 1.0854 | 0.1389 | -3.6343 | -0.3029 |
| Colombia | PPP | 1.0441 | -1.6714 | 0.0481 | 0.5707 | 1.9537 | 0.0254 | -0.5341 | -0.0451 |
| Hungary | Real Oil Price | 1.0253 | -0.8015 | 0.2119 | 0.4869 | -0.3618 | 0.6412 | 2.1977 | 0.1667 |
| Hungary | Real Oil Price+Investor Sentiment (4) | 1.1053 | -1.2709 | 0.1027 | 0.5340 | 0.9406 | 0.1734 | 2.6637 | 0.1961 |
| Hungary | Taylor Rule Model+Investor Sentiment (15) | 1.2173 | -1.9729 | 0.0250 | 0.5550 | 1.5195 | 0.0643 | 2.8531 | 0.2141 |
| India | PPP | 0.9737 | 0.7842 | 0.2170 | 0.5730 | 1.9851 | 0.0236 | 3.8507 | 0.4786 |
| India | PPP+SKEW | 0.9912 | 0.2500 | 0.4014 | 0.5730 | 1.9851 | 0.0236 | 4.0571 | 0.5539 |
| India | PPP | 0.9737 | 0.7842 | 0.2170 | 0.5730 | 1.9851 | 0.0236 | 3.8507 | 0.4786 |
| Mexico | Taylor Rule Model | 1.0114 | -0.3184 | 0.3753 | 0.5191 | 0.5175 | 0.3024 | 2.0572 | 0.1762 |
| Mexico | Taylor Rule Model+Composite Leading Indicator | 1.0362 | -1.0381 | 0.1503 | 0.5410 | 1.1088 | 0.1338 | 3.0948 | 0.2862 |
| Mexico | Oil Price+Investor Sentiment (2) | 1.0466 | -1.3133 | 0.0954 | 0.5410 | 1.1088 | 0.1338 | 6.5728 | 0.5428 |
| Peru | Terms of Trade | 1.0255 | -0.7578 | 0.2248 | 0.6178 | 3.2561 | 0.0006 | 1.5976 | 0.2820 |
| Peru | Terms of Trade+SKEW | 1.0358 | -1.0724 | 0.1425 | 0.6319 | 3.5580 | 0.0002 | 1.4482 | 0.2803 |
| Peru | Terms of Trade+SKEW | 1.0358 | -1.0724 | 0.1425 | 0.6319 | 3.5580 | 0.0002 | 1.4482 | 0.2803 |
| Poland | Terms of Trade | 1.0021 | -0.0969 | 0.4615 | 0.6073 | 2.9667 | 0.0015 | 5.3513 | 0.3184 |
| Poland | Terms of Trade+Investor Sentiment (5) | 1.0037 | -0.0400 | 0.4841 | 0.6178 | 3.2561 | 0.0006 | 10.4934 | 0.6912 |
| Poland | Terms of Trade+Investor Sentiment (5) | 1.0037 | -0.0400 | 0.4841 | 0.6178 | 3.2561 | 0.0006 | 10.4934 | 0.6912 |
| South Africa | PPP | 1.0204 | -0.9943 | 0.1607 | 0.4869 | -0.3618 | 0.6412 | 0.1124 | 0.0076 |
| South Africa | PPP+Investor Sentiment (17) | 1.1310 | -1.8260 | 0.0347 | 0.5497 | 1.3748 | 0.0846 | 6.8940 | 0.4432 |
| South Africa | PPP+Investor Sentiment (17) | 1.1310 | -1.8260 | 0.0347 | 0.5497 | 1.3748 | 0.0846 | 6.8940 | 0.4432 |
| Tunisia | Commodity Price | 0.9946 | 0.1224 | 0.4514 | 0.5879 | 2.3720 | 0.0088 | 4.1238 | 0.5098 |
| Tunisia | Commodity Price+SKEW | 1.0011 | -0.0247 | 0.4901 | 0.5934 | 2.5202 | 0.0059 | 5.4322 | 0.6884 |
| Tunisia | Commodity Price+SKEW | 1.0011 | -0.0247 | 0.4901 | 0.5934 | 2.5202 | 0.0059 | 5.4322 | 0.6884 |
| Turkey | PPP | 0.9538 | 1.3843 | 0.0839 | 0.5737 | 2.0313 | 0.0211 | 10.5444 | 0.6222 |
| Turkey | PPP+VIX | 0.9915 | 0.2107 | 0.4167 | 0.5947 | 2.6117 | 0.0045 | 13.1033 | 0.8699 |
| Turkey | Real Oil Price+VIX | 1.0337 | -0.8534 | 0.1972 | 0.6158 | 3.1921 | 0.0007 | 11.9540 | 0.8657 |

Note: The column "Countries" shows the names of countries. The column "Models" defines the underlying model. The column "MSPE_ratio" indicates the $\text{MSPE}_{ratio}$ for the best macro models, the best behavioral models (i.e., the best macro model plus the behavioral factors), and the best models among all macro and behavioral models for a given country. The column "DM(t-stats)" shows the DM statistics for the best macro models, the best behavioral models, and the best models among all macro and behavioral models for each country. The column "DM(p-value)" shows the p-values associated with the DM statistics. The column "$\hat{p}$" indicates the proportions of correct predictions of direction for the best macro models, the best behavioral models, and the best models among all macro and behavioral models for each country. The column "PT(t-stats)" shows the Pesaran-Timmermann (PT) statistics for the best macro models, the best behavioral models, and the best models among all macro and behavioral models for each country and the column "PT(p-value)" shows the p-values associated with the PT statistics. All reported test statistics are one-sided. The significance level is 10%. The column "Avg. Return (%)" shows the annualized average returns in percentage, and the column " Sharp_ratio " shows the annualized Sharp-ratio.

Table 2.9 reports the summary results at horizon=12 and the median of statistics for the best macroeconomic model, the behavioral model (i.e., the best macroeconomic model plus behavioral factors), and the best model among all available (macroeconomic and behavioral) models within each country group (advanced and emerging/developing).

Table 2.9: Summary results in horizon=12

| WEO Country Group | Definitions | MSPE_ratio | DM (t-stats) | DM (p-value) | $\hat{p}$ | PT(t-stats) | PT (p-value) | Avg. Return (%) | Sharpe_ratio |
|---|---|---|---|---|---|---|---|---|---|
| Advanced | Best Macro Model | 1.1563 | -1.0783 | 0.1413 | 0.5562 | 1.4615 | 0.0719 | 0.4121 | 0.0256 |
| | Best Behavioral Model | 1.2754 | -1.4154 | 0.0800 | 0.5828 | 2.1538 | 0.0159 | 1.1190 | 0.0602 |
| | Best Model | 1.3173 | -1.4830 | 0.0700 | 0.5828 | 2.1538 | 0.0159 | 1.2948 | 0.0802 |
| Emerging/Developing | Best Macro Model | 1.1838 | -0.8755 | 0.0693 | 0.6450 | 3.7692 | 0.0001 | 2.1263 | 0.1117 |
| | Best Behavioral Model | 1.1609 | -0.9556 | 0.1070 | 0.6509 | 3.9231 | 0.0000 | 3.1966 | 0.1436 |
| | Best Model | 1.1609 | -0.9556 | 0.1070 | 0.6509 | 3.9231 | 0.0000 | 3.1966 | 0.1436 |

Note: The column "WEO Country Group" shows the World Economic Outlook (WEO) countries' classifications. The column "Definitions" defines the underlying best forecasting model. The column "MSPE_ratio" indicates the median of the $MSPE_{ratio}$ for the best macro models, the best behavioral models (i.e., the best macro model plus behavioral factors), and the best models among all macro and behavioral models within each country group. The column "DM(t-stats)" shows the median of DM statistics for the best macro models, the best behavioral models, and the best models among all macro and behavioral models within each country group. The column "DM(p-value)" shows the p-values associated with the DM statistics. The column "$\hat{p}$" indicates the median of proportions of correct predictions of direction for the best macro models, the best behavioral models, and the best models among all macro and behavioral models within each country group. The column "PT(t-stats)" indicates the median of the Pesaran-Timmermann (PT) statistics for the best macro models, the best behavioral models, and the best models among all macro and behavioral models within each country group and the column "PT(p-value)" shows the p-values associated with the PT statistics. All reported test statistics are one-sided. The significance level is 10%. The column "Avg. Return (%)" shows the annualized average returns in percentage, and the column " Sharp_ratio " shows the annualized Sharp-ratio.

The results in Table 2.9 indicate, on average, the best macroeconomic model has a proportion of correct direction of changes greater than 0.5, and it is statistically significant at the 10% significance level in both groups. However, on average, none of the best macroeconomic models has a $MSPE_{ratio}$ statistic less than one and a positive and statistically significant DM statistic at horizon=12. The findings also indicate that adding behavioral factors, on average, improves the best macro models' ability to forecast the direction of changes correctly in both groups of countries.

Tables 2.10 and 2.11 show the detailed empirical results for advanced countries at horizon=12. The results show the best macroeconomic model, the best behavioral model, the best model among all available models for a given country. For example, for Norway, the best macroeconomic model is the model that includes changes in the oil price. The best behavioral model is the model that includes changes in the oil price, changes in VIX, and changes in SKEW variables for Norway at horizon=12.

## Table 2.10: Empirical results for advanced countries at horizon=12

| Countries | Models | MSPE_ratio | DM (t-stats) | DM (p-value) | $\hat{p}$ | PT(t-stats) | PT (p-value) | Avg. Return (%) | Sharpe_ratio |
|---|---|---|---|---|---|---|---|---|---|
| Austria | Commodity Price | 1.1556 | -1.0564 | 0.1462 | 0.5562 | 1.4615 | 0.0719 | 0.4121 | 0.0256 |
| Austria | Commodity Price+Investor Sentiment (2) | 1.1233 | -0.7362 | 0.2313 | 0.5917 | 2.3846 | 0.0085 | 1.9132 | 0.1126 |
| Austria | Real Oil Price+Investor Sentiment (8) | 1.1437 | -0.8334 | 0.2029 | 0.5976 | 2.5385 | 0.0056 | 1.7364 | 0.1055 |
| Belgium | Commodity Price | 1.1556 | -1.0564 | 0.1462 | 0.5562 | 1.4615 | 0.0719 | 0.4121 | 0.0256 |
| Belgium | Commodity Price+Investor Sentiment (10) | 1.2180 | -1.3104 | 0.0959 | 0.5976 | 2.5385 | 0.0056 | 1.2413 | 0.0781 |
| Belgium | Oil Price+Investor Sentiment (10) | 1.2116 | -1.3246 | 0.0936 | 0.6095 | 2.8462 | 0.0022 | 1.3156 | 0.0814 |
| Canada | UIRP | 1.3162 | -1.5337 | 0.0635 | 0.6188 | 3.0042 | 0.0013 | 1.4261 | 0.0868 |
| Canada | UIRP+VIX | 1.3281 | -1.5860 | 0.0574 | 0.5938 | 2.3717 | 0.0089 | 1.2281 | 0.0768 |
| Canada | UIRP | 1.3162 | -1.5337 | 0.0635 | 0.6188 | 3.0042 | 0.0013 | 1.4261 | 0.0868 |
| Czech Republic | Real Oil Price | 1.1197 | -0.6085 | 0.2718 | 0.6036 | 2.6923 | 0.0035 | 1.9605 | 0.0836 |
| Czech Republic | Real Oil Price+VIX | 1.1322 | -0.6605 | 0.2549 | 0.5917 | 2.3846 | 0.0085 | 1.6725 | 0.0720 |
| Czech Republic | PPP+Investor Sentiment (6) | 1.4527 | -1.5574 | 0.0606 | 0.6036 | 2.6923 | 0.0035 | 2.3166 | 0.0982 |
| Denmark | Terms of Trade | 1.3187 | -1.4602 | 0.0730 | 0.5799 | 2.0769 | 0.0189 | 1.6337 | 0.0962 |
| Denmark | Terms of Trade+SKEW | 1.3313 | -1.4918 | 0.0688 | 0.5562 | 1.4615 | 0.0719 | 0.9437 | 0.0623 |
| Denmark | Terms of Trade | 1.3187 | -1.4602 | 0.0730 | 0.5799 | 2.0769 | 0.0189 | 1.6337 | 0.0962 |
| Estonia | Oil Price | 1.1525 | -1.0261 | 0.1532 | 0.5562 | 1.4615 | 0.0719 | 0.6378 | 0.0370 |
| Estonia | Oil Price+Consumer Confidence | 1.1676 | -1.1257 | 0.1309 | 0.5740 | 1.9231 | 0.0272 | 1.0099 | 0.0581 |
| Estonia | Terms of Trade+VIX | 1.2185 | -1.3741 | 0.0856 | 0.5740 | 1.9231 | 0.0272 | 1.0515 | 0.0635 |
| Finland | Terms of Trade | 1.2618 | -1.3932 | 0.0827 | 0.5799 | 2.0769 | 0.0189 | 1.6542 | 0.0910 |
| Finland | Terms of Trade+Investor Sentiment (14) | 1.3185 | -1.3390 | 0.0912 | 0.6982 | 5.1538 | 0.0000 | 3.3179 | 0.2066 |
| Finland | Terms of Trade+Investor Sentiment (14) | 1.3185 | -1.3390 | 0.0912 | 0.6982 | 5.1538 | 0.0000 | 3.3179 | 0.2066 |
| France | Commodity Price | 1.1556 | -1.0564 | 0.1462 | 0.5562 | 1.4615 | 0.0719 | 0.4121 | 0.0256 |
| France | Commodity Price+Consumer Confidence | 1.2406 | -1.5842 | 0.0575 | 0.5621 | 1.6154 | 0.0531 | 0.5965 | 0.0375 |
| France | Commodity Price+Consumer Confidence | 1.2406 | -1.5842 | 0.0575 | 0.5621 | 1.6154 | 0.0531 | 0.5965 | 0.0375 |
| Germany | Commodity Price | 1.1556 | -1.0564 | 0.1462 | 0.5562 | 1.4615 | 0.0719 | 0.4121 | 0.0256 |
| Germany | Commodity Price+Investor Sentiment (26) | 1.3461 | -1.4965 | 0.0682 | 0.6036 | 2.6923 | 0.0035 | 1.6092 | 0.1162 |
| Germany | Real Commodity Price+Investor Sentiment (26) | 1.3491 | -1.5058 | 0.0670 | 0.6095 | 2.8462 | 0.0022 | 1.7764 | 0.1220 |
| Greece | Commodity Price | 1.1560 | -1.0593 | 0.1455 | 0.5562 | 1.4615 | 0.0719 | 0.4121 | 0.0256 |
| Greece | Commodity Price+Consumer Confidence | 1.1935 | -1.2409 | 0.1082 | 0.5680 | 1.7692 | 0.0384 | 0.8055 | 0.0479 |
| Greece | Commodity Price+Consumer Confidence | 1.1935 | -1.2409 | 0.1082 | 0.5680 | 1.7692 | 0.0384 | 0.8055 | 0.0479 |
| Iceland | Extended UIRP | 2.0757 | -1.7776 | 0.0386 | 0.4615 | -1.0000 | 0.8413 | -0.9335 | -0.0273 |
| Iceland | Extended UIRP+Investor Sentiment (1) | 2.0871 | -1.7852 | 0.0380 | 0.4793 | -0.5385 | 0.7049 | -0.7447 | -0.0220 |
| Iceland | Extended UIRP+Investor Sentiment (1) | 2.0871 | -1.7852 | 0.0380 | 0.4793 | -0.5385 | 0.7049 | -0.7447 | -0.0220 |
| Ireland | Terms of Trade | 1.2030 | -1.0973 | 0.1371 | 0.5621 | 1.6154 | 0.0531 | 0.8319 | 0.0481 |
| Ireland | Terms of Trade+Investor Sentiment (26) | 1.4801 | -1.8956 | 0.0299 | 0.6036 | 2.6923 | 0.0035 | 1.2740 | 0.0830 |
| Ireland | Terms of Trade+Investor Sentiment (26) | 1.4801 | -1.8956 | 0.0299 | 0.6036 | 2.6923 | 0.0035 | 1.2740 | 0.0830 |
| Israel | Taylor Rule Model | 1.7568 | -2.7984 | 0.0029 | 0.4852 | -0.3846 | 0.6497 | -0.6297 | -0.0461 |
| Israel | Taylor Rule Model+Investor Sentiment (4) | 1.8759 | -3.0823 | 0.0012 | 0.4852 | -0.3846 | 0.6497 | -0.5803 | -0.0482 |
| Israel | UIRP+Investor Sentiment (5) | 1.5115 | -2.7690 | 0.0031 | 0.5150 | 0.3869 | 0.3494 | 0.1080 | 0.0069 |

Note: The column "Countries" shows the names of countries. The column "Models" defines the underlying model. The column "MSPE_ratio" indicates the $\text{MSPE}_{ratio}$ for the best macro models, the best behavioral models (i.e., the best macro model plus the behavioral factors), and the best models among all macro and behavioral models for a given country. The column "DM(t-stats)" shows the DM statistics for the best macro models, the best behavioral models, and the best models among all macro and behavioral models for each country. The column "DM(p-value)" shows the p-values associated with the DM statistics. The column "$\hat{p}$" indicates the proportions of correct predictions of direction for the best macro models, the best behavioral models, and the best models among all macro and behavioral models for each country. The column "PT(t-stats)" shows the Pesaran-Timmermann (PT) statistics for the best macro models, the best behavioral models, and the best models among all macro and behavioral models for each country and the column "PT(p-value)" shows the p-values associated with the PT statistics. All reported test statistics are one-sided. The significance level is 10%. The column "Avg. Return (%)" shows the annualized average returns in percentage, and the column " Sharp_ratio " shows the annualized Sharp-ratio.

Table 2.11: Empirical results for advanced countries at horizon=12, continued

| Countries | Models | MSPE_ratio | DM (t-stats) | DM (p-value) | $\hat{p}$ | PT(t-stats) | PT (p-value) | Avg. Return (%) | Sharpe_ratio |
|---|---|---|---|---|---|---|---|---|---|
| Italy | Commodity Price | 1.1556 | -1.0564 | 0.1462 | 0.5562 | 1.4615 | 0.0719 | 0.4121 | 0.0256 |
| Italy | Commodity Price+VIX | 1.1567 | -1.0524 | 0.1471 | 0.5621 | 1.6154 | 0.0531 | 0.6363 | 0.0391 |
| Italy | Taylor Rule Model+Investor Sentiment (14) | 1.6410 | -1.9990 | 0.0236 | 0.5799 | 2.0769 | 0.0189 | 1.5996 | 0.0968 |
| Japan | PPP | 1.6031 | -2.5478 | 0.0059 | 0.5562 | 1.4615 | 0.0719 | 0.1816 | 0.0093 |
| Japan | PPP+Investor Sentiment (25) | 1.7960 | -3.1515 | 0.0010 | 0.5680 | 1.7692 | 0.0384 | 0.7937 | 0.0399 |
| Japan | PPP+Investor Sentiment (25) | 1.7960 | -3.1515 | 0.0010 | 0.5680 | 1.7692 | 0.0384 | 0.7937 | 0.0399 |
| Korea | Terms of Trade | 1.4418 | -2.7138 | 0.0037 | 0.4024 | -2.5385 | 0.9944 | -2.7630 | -0.1318 |
| Korea | Terms of Trade+Investor Sentiment (23) | 1.5229 | -3.3279 | 0.0005 | 0.4911 | -0.2308 | 0.5913 | -1.8221 | -0.0865 |
| Korea | Terms of Trade+Investor Sentiment (23) | 1.5229 | -3.3279 | 0.0005 | 0.4911 | -0.2308 | 0.5913 | -1.8221 | -0.0865 |
| Luxembourg | Commodity Price | 1.1556 | -1.0564 | 0.1462 | 0.5562 | 1.4615 | 0.0719 | 0.4121 | 0.0256 |
| Luxembourg | Commodity Price+Investor Sentiment (13) | 1.1591 | -1.0810 | 0.1406 | 0.5680 | 1.7692 | 0.0384 | 0.7211 | 0.0438 |
| Luxembourg | Commodity Price+Investor Sentiment (13) | 1.1591 | -1.0810 | 0.1406 | 0.5680 | 1.7692 | 0.0384 | 0.7211 | 0.0438 |
| Netherlands | Terms of Trade | 1.2376 | -1.3177 | 0.0947 | 0.5740 | 1.9231 | 0.0272 | 0.7773 | 0.0466 |
| Netherlands | Terms of Trade+Investor Sentiment (15) | 1.3821 | -2.0314 | 0.0219 | 0.5740 | 1.9231 | 0.0272 | 0.8131 | 0.0515 |
| Netherlands | Terms of Trade+Investor Sentiment (15) | 1.3821 | -2.0314 | 0.0219 | 0.5740 | 1.9231 | 0.0272 | 0.8131 | 0.0515 |
| Norway | Oil Price | 1.0845 | -0.6990 | 0.2428 | 0.5976 | 2.5385 | 0.0056 | 0.9268 | 0.0437 |
| Norway | Oil Price+Investor Sentiment (1) | 1.0989 | -0.8162 | 0.2078 | 0.6095 | 2.8462 | 0.0022 | 1.5576 | 0.0790 |
| Norway | Oil Price+Investor Sentiment (1) | 1.0989 | -0.8162 | 0.2078 | 0.6095 | 2.8462 | 0.0022 | 1.5576 | 0.0790 |
| Portugal | PPP | 1.3090 | -1.7625 | 0.0399 | 0.5740 | 1.9231 | 0.0272 | 1.5508 | 0.0935 |
| Portugal | PPP+Business Confidence | 1.3463 | -1.8677 | 0.0318 | 0.5976 | 2.5385 | 0.0056 | 1.7941 | 0.1085 |
| Portugal | Real Commodity Price+Investor Sentiment (6) | 1.1542 | -1.0347 | 0.1512 | 0.6036 | 2.6923 | 0.0035 | 1.2215 | 0.0729 |
| Slovakia | Oil Price | 1.0717 | -0.3452 | 0.3652 | 0.5680 | 1.7692 | 0.0384 | 2.0569 | 0.1010 |
| Slovakia | Oil Price+Investor Sentiment (14) | 1.2489 | -0.8633 | 0.1946 | 0.5917 | 2.3846 | 0.0085 | 2.7730 | 0.1418 |
| Slovakia | Commodity Price+Investor Sentiment (2) | 1.2372 | -0.8963 | 0.1857 | 0.6036 | 2.6923 | 0.0035 | 2.2988 | 0.1148 |
| Slovenia | Oil Price | 1.1567 | -1.2302 | 0.1102 | 0.5503 | 1.3077 | 0.0955 | 0.5417 | 0.0311 |
| Slovenia | Oil Price+Investor Sentiment (3) | 1.1822 | -1.1543 | 0.1250 | 0.5799 | 2.0769 | 0.0189 | 1.4290 | 0.0831 |
| Slovenia | Oil Price+Investor Sentiment (3) | 1.1822 | -1.1543 | 0.1250 | 0.5799 | 2.0769 | 0.0189 | 1.4290 | 0.0831 |
| Spain | Commodity Price | 1.1556 | -1.0564 | 0.1462 | 0.5562 | 1.4615 | 0.0719 | 0.4121 | 0.0256 |
| Spain | Commodity Price+Investor Sentiment (7) | 1.2034 | -1.3139 | 0.0953 | 0.5858 | 2.2308 | 0.0128 | 0.2692 | 0.0170 |
| Spain | Commodity Price+Investor Sentiment (7) | 1.2034 | -1.3139 | 0.0953 | 0.5858 | 2.2308 | 0.0128 | 0.2692 | 0.0170 |
| Sweden | Commodity Price | 1.2115 | -1.5122 | 0.0662 | 0.6036 | 2.6923 | 0.0035 | -0.0744 | -0.0035 |
| Sweden | Commodity Price+Composite Leading Indicator | 1.3018 | -1.7149 | 0.0441 | 0.6095 | 2.8462 | 0.0022 | 1.2360 | 0.0566 |
| Sweden | Terms of Trade+Investor Sentiment (26) | 1.4788 | -2.1580 | 0.0162 | 0.6154 | 3.0000 | 0.0013 | 1.7977 | 0.0947 |
| Switzerland | Oil Price | 1.1025 | -0.5835 | 0.2802 | 0.5621 | 1.6154 | 0.0531 | 1.3091 | 0.0823 |
| Switzerland | Oil Price+Investor Sentiment (18) | 1.1064 | -0.6172 | 0.2690 | 0.6213 | 3.1538 | 0.0008 | 1.8663 | 0.1241 |
| Switzerland | Real Oil Price+Investor Sentiment (21) | 1.1194 | -0.6726 | 0.2511 | 0.6213 | 3.1538 | 0.0008 | 1.6348 | 0.1227 |
| United Kingdom | PPP | 1.4493 | -2.3177 | 0.0109 | 0.5633 | 1.5911 | 0.0558 | -0.3655 | -0.0179 |
| United Kingdom | PPP+Investor Sentiment (7) | 1.3928 | -2.2298 | 0.0136 | 0.5759 | 1.9093 | 0.0281 | -0.0201 | -0.0010 |
| United Kingdom | PPP+Investor Sentiment (7) | 1.3928 | -2.2298 | 0.0136 | 0.5759 | 1.9093 | 0.0281 | -0.0201 | -0.0010 |

Note: The column "Countries" shows the names of countries. The column "Models" defines the underlying model. The column "MSPE$_ratio$" indicates the MSPE$_{ratio}$ for the best macro models, the best behavioral models (i.e., the best macro model plus the behavioral factors), and the best models among all macro and behavioral models for a given country. The column "DM(t-stats)" shows the DM statistics for the best macro models, the best behavioral models, and the best models among all macro and behavioral models for each country. The column "DM(p-value)" shows the p-values associated with the DM statistics. The column "$\hat{p}$" indicates the proportions of correct predictions of direction for the best macro models, the best behavioral models, and the best models among all macro and behavioral models for each country. The column "PT(t-stats)" shows the Pesaran-Timmermann (PT) statistics for the best macro models, the best behavioral models, and the best models among all macro and behavioral models for each country and the column "PT(p-value)" shows the p-values associated with the PT statistics. All reported test statistics are one-sided. The significance level is 10%. The column "Avg. Return (%)" shows the annualized average returns in percentage, and the column " Sharp+ratio " shows the annualized Sharp-ratio.

The findings in this essay are consistent with the literature that the PPP and UIRP models are weak predictors of exchange rates in terms of the MSPE$_{ratio}$ and DM statistics at longer horizons. Yet, the empirical evidence in this essay indicates that these models correctly forecast the direction of changes in exchange rates for some countries, including

Canada, Japan, Portugal, and the United Kingdom. Similar to the results at horizon=1, changes in oil and commodity prices (in both nominal and real terms) are not successful in predicting point forecasts. However, these variables do well in predicting the direction of changes in exchange rates for some advanced countries, including the Czech Republic, Germany, Sweden, and Switzerland.

Tables 2.10 and 2.11 show that adding investor sentiment variables leads to improved out-of-sample performance of the best macroeconomic models in terms of the PT statistic for 73% of advanced countries at horizon=12. For the rest of countries, the process does not impact the predictive ability of macroeconomic models. Moreover, the findings show that for all countries, except Canada and Denmark, the best forecasting models are those that contain both macroeconomic and behavioral factors. The annualized average returns for the remaining advanced countries are either positive but small or negative.

Table 2.12 shows the results for emerging/developing countries at horizon=12. The results show that incorporating behavioral factors into macroeconomic models improves models' predictive ability in terms of the PT statistic for all emerging/developing counties at horizon=12. In addition, changes in commodity and oil prices outperform the benchmark model according to the PT statistic for some emerging/developing countries, including Brazil, Chile, and Hungary. Also, the empirical evidence shows that investment strategies using behavioral models have positive and notable annualized average returns for all emerging/developing countries, except for Columbia, Peru, and Poland.

Table 2.12: Empirical results for emerging/developing countries in horizon=12

| Countries | Models | MSPE_ratio | DM (t-stats) | DM (p-value) | $\hat{p}$ | PT(t-stats) | PT (p-value) | Avg. Return (%) | Sharpe_ratio |
|---|---|---|---|---|---|---|---|---|---|
| Brazil | Oil Price | 1.1102 | -0.7495 | 0.2273 | 0.6095 | 2.8462 | 0.0022 | 4.1699 | 0.1312 |
| Brazil | Oil Price+Investor Sentiment (11) | 1.0976 | -0.6995 | 0.2426 | 0.6331 | 3.4615 | 0.0003 | 4.0317 | 0.1394 |
| Brazil | Real Oil Price+Investor Sentiment (11) | 1.0969 | -0.6944 | 0.2442 | 0.6331 | 3.4615 | 0.0003 | 4.0317 | 0.1394 |
| Chile | Oil Price | 1.1154 | -0.8499 | 0.1983 | 0.5503 | 1.3077 | 0.0955 | 0.6534 | 0.0325 |
| Chile | Oil Price+Composite Leading Indicator | 0.9277 | 0.6271 | 0.2657 | 0.6036 | 2.6923 | 0.0035 | 2.0750 | 0.1025 |
| Chile | Oil Price+Composite Leading Indicator | 0.9277 | 0.6271 | 0.2657 | 0.6036 | 2.6923 | 0.0035 | 2.0750 | 0.1025 |
| Colombia | PPP | 1.7164 | -1.6597 | 0.0494 | 0.6036 | 2.6923 | 0.0035 | 1.1227 | 0.0396 |
| Colombia | PPP+SKEW | 1.7286 | -1.6834 | 0.0471 | 0.6154 | 3.0000 | 0.0013 | 1.4403 | 0.0532 |
| Colombia | PPP+SKEW | 1.7286 | -1.6834 | 0.0471 | 0.6154 | 3.0000 | 0.0013 | 1.4403 | 0.0532 |
| Hungary | Real Oil Price | 1.0231 | -0.1926 | 0.4237 | 0.6568 | 4.0769 | 0.0000 | 2.7960 | 0.1242 |
| Hungary | Real Oil Price+Business Confidence | 1.0633 | -0.5716 | 0.2842 | 0.6627 | 4.2308 | 0.0000 | 3.1966 | 0.1436 |
| Hungary | Real Oil Price+Business Confidence | 1.0633 | -0.5716 | 0.2842 | 0.6627 | 4.2308 | 0.0000 | 3.1966 | 0.1436 |
| India | PPP | 1.2683 | -1.8594 | 0.0324 | 0.6545 | 3.9703 | 0.0000 | 1.6444 | 0.0954 |
| India | PPP+Composite Leading Indicator | 1.2436 | -1.2477 | 0.1070 | 0.7030 | 5.2159 | 0.0000 | 3.0469 | 0.1881 |
| India | PPP+Composite Leading Indicator | 1.2436 | -1.2477 | 0.1070 | 0.7030 | 5.2159 | 0.0000 | 3.0469 | 0.1881 |
| Mexico | Extended UIRP | 1.2406 | -1.5208 | 0.0651 | 0.5776 | 1.9703 | 0.0244 | 2.1263 | 0.1117 |
| Mexico | Extended UIRP+Business Confidence | 1.1609 | -0.9556 | 0.1704 | 0.6460 | 3.7041 | 0.0001 | 4.5361 | 0.2357 |
| Mexico | Extended UIRP+Business Confidence | 1.1609 | -0.9556 | 0.1704 | 0.6460 | 3.7041 | 0.0001 | 4.5361 | 0.2357 |
| Peru | Real Oil Price | 1.1865 | -1.4883 | 0.0693 | 0.6188 | 3.0042 | 0.0013 | 0.5018 | 0.0421 |
| Peru | Real Oil Price+VIX | 1.1977 | -1.6022 | 0.0555 | 0.6125 | 2.8460 | 0.0022 | 0.4373 | 0.0372 |
| Peru | Real Oil Price | 1.1865 | -1.4883 | 0.0693 | 0.6188 | 3.0042 | 0.0013 | 0.5018 | 0.0421 |
| Poland | Taylor Rule Model | 1.2696 | -1.4621 | 0.0728 | 0.6450 | 3.7692 | 0.0001 | 1.9042 | 0.0686 |
| Poland | Taylor Rule Model+Investor Sentiment (17) | 1.7891 | -1.8882 | 0.0304 | 0.6509 | 3.9231 | 0.0000 | 1.6900 | 0.0639 |
| Poland | Taylor Rule Model+Investor Sentiment (17) | 1.7891 | -1.8882 | 0.0304 | 0.6509 | 3.9231 | 0.0000 | 1.6900 | 0.0639 |
| South Africa | UIRP | 1.1838 | -0.8755 | 0.1913 | 0.6982 | 5.1538 | 0.0000 | 5.8649 | 0.2106 |
| South Africa | UIRP+Investor Sentiment (7) | 1.2021 | -1.0026 | 0.1587 | 0.7278 | 5.9231 | 0.0000 | 6.7078 | 0.2646 |
| South Africa | UIRP+Investor Sentiment (7) | 1.2021 | -1.0026 | 0.1587 | 0.7278 | 5.9231 | 0.0000 | 6.7078 | 0.2646 |
| Tunisia | Oil Price | 0.6719 | 2.3782 | 0.0093 | 0.7938 | 7.4314 | 0.0000 | 5.0533 | 0.2991 |
| Tunisia | Oil Price+SKEW | 0.6747 | 2.3645 | 0.0096 | 0.7875 | 7.2732 | 0.0000 | 4.9716 | 0.2944 |
| Tunisia | Oil Price | 0.6719 | 2.3782 | 0.0093 | 0.7938 | 7.4314 | 0.0000 | 5.0533 | 0.2991 |
| Turkey | UIRP | 0.9340 | 1.8640 | 0.0320 | 0.7619 | 6.7893 | 0.0000 | 7.9859 | 0.2405 |
| Turkey | UIRP+Business Confidence | 0.8838 | 2.3385 | 0.0103 | 0.8214 | 8.3324 | 0.0000 | 9.6139 | 0.3070 |
| Turkey | UIRP+Business Confidence | 0.8838 | 2.3385 | 0.0103 | 0.8214 | 8.3324 | 0.0000 | 9.6139 | 0.3070 |

Note: The column "Countries" shows the names of countries. The column "Models" defines the underlying model. The column "MSPE_ratio" indicates the $\text{MSPE}_{ratio}$ for the best macro models, the best behavioral models (i.e., the best macro model plus the behavioral factors), and the best models among all macro and behavioral models for a given country. The column "DM(t-stats)" shows the DM statistics for the best macro models, the best behavioral models, and the best models among all macro and behavioral models for each country. The column "DM(p-value)" shows the p-values associated with the DM statistics. The column "$\hat{p}$" indicates the proportions of correct predictions of direction for the best macro models, the best behavioral models, and the best models among all macro and behavioral models for each country. The column "PT(t-stats)" shows the Pesaran-Timmermann (PT) statistics for the best macro models, the best behavioral models, and the best models among all macro and behavioral models for each country and the column "PT(p-value)" shows the p-values associated with the PT statistics. All reported test statistics are one-sided. The significance level is 10%. The column "Avg. Return (%)" shows the annualized average returns in percentage, and the column " Sharp_ratio " shows the annualized Sharp-ratio.

Overall findings across advanced and emerging/developing countries emphasize the advantage of incorporating investor sentiment variables (behavioral factors) in the forecasting process. It is important to note that investors analyze and use available information differently to make economic and investment decisions. The results confirm that investors' opinions and perceptions of financial market risk could play essential roles in forming better forecasts for exchange rates.

Moreover, the results highlight the importance of forecast evaluation methods. Few behavioral/macroeconomic models could improve forecasting results in terms of $\text{MSPE}_{ratio}$ and DM statistics. However, according to PT statistics, most behavioral/macroeconomic models outperform the WN model across different countries and horizons. These findings are critical for policymakers and practitioners in financial markets. The empirical evidence indicates, on average, that models which include investor sentiment indices have higher values of $\hat{p}$ and PT(t-stats) and are more successful in predicting changes in exchange rates in emerging/developing versus advanced countries. The different results could come from differences in economic policies, the maturity of financial markets, people's confidence about the stability of economic conditions.

## 2.6 Conclusion

This essay took a new direction and incorporated behavioral economics by adding investor sentiment indices to macroeconomic models. Using macroeconomic models, I evaluated monthly changes in exchange rates predictability for the US dollar against 37 currencies. I explored the predictive ability of the change of terms of trade index as a single predictor and an additional predictor in uncovered interest rate parity (UIRP). Moreover, I evaluated the out-of-sample performance of uncovered interest rate parity augmented by the inflation rate differential (Extended UIRP). The inflation rate differential means the difference between the inflation rate of a foreign country and the US.

Next, I added investor sentiment variables to the 'best' macroeconomic model in terms of PT statistics for each country and horizon to examine whether out-of-sample performance of the macro model could be improved. These indices include consumer confidence, business confidence, composited leading indicator, VIX, and SKEW indices (in changes). Finally, I included investor sentiment variable to all macroeconomic models for each country and horizon and examined whether this method could improve the forecasting results for exchange rates. I assessed the out-of-sample predictive ability of models in two ways: their ability to predict point forecasts and their ability to forecast the direction of change in exchange rates.

Therefore, I used various statistics, including $MSPE_{ratio}$, DM, and PT statistics.

The findings indicate forecasting models that include macroeconomic and behavioral factors outperform the WN model in terms of PT statistics for most countries at the 1-month-ahead and 12-months-ahead forecasts. The results confirm that people's perceptions about the financial market and its stability could affect their economic and investment decisions. Therefore, they use different information to forecast future movements of exchange rates.

The remarkable results that emerge from the empirical analysis are that traditional macroeconomic models may not be successful in predicting point forecasts. Still, they correctly forecast the direction of changes in exchange rates for most countries. Also, there are similar outcomes when traditional macroeconomic models are combined with behavioral variables. Therefore, these models and their extension by investor sentiment variables could be very informative for policymakers and financial markets participants.

# Essay 3

# Ensemble (Combinations) Methods and Exchange Rate Forecasting

## 3.1 Introduction

In practice, it is common to use a set of models to forecast exchange rates. Models are constructed based upon different approaches and assumptions. Forecasts could come from subjective judgments from experts or quantitative methods, i.e., linear or non-linear econometric models with constant or time-varying parameters. Furthermore, predictions could come from contemporaneous or historical data.

When there is a set of forecasting models, one important question is which model is the 'best.' Since the study of Bates and Granger (1969), forecast combinations have been used in central banks, among private-sector forecasters, and academic studies. The forecast combinations have been viewed as a simple and efficient method to improve the forecasting performance over the performance of individual forecasting models. Bates and Granger (1969) argue that if the objective is to have a good forecast, an individual forecasting model with poor forecasting results should not be discarded. Individuals forecasting models may have independent information that comes from either different variables or information in each model or different assumptions about the relationship between the variables.

According to the study by Timmermann (2006), forecast combinations have been used successfully in different areas, including management science, finance, economics, and psychology. The author explains that combining different forecasts from various models could decrease the impacts of misspecification biases and measurements errors from individual models.

In my opinion, forecast combination approaches in predicting changes in exchange rates can be helpful because, based on the empirical evidence in the literature and this dissertation, no single forecasting model is successful for all currencies. In addition, an individual forecasting model with better out-of-sample performance than other models is expected to have a higher weight than others in forecast combination procedures, except in an equal weighting method.

There are two main stages to constructing combined forecasts: (1) Model selection and (2) weight assignment. In the first stage, $N$ models are selected. The set of selected forecasting models could be fixed or change over time. The literature review shows that even if underlying forecasting models are fixed over time, their parameters (coefficients) are updated and re-estimated at each period using either a rolling window estimation approach or an expanding window estimation approach. Then, estimated parameters are used to form a forecast for the next period.

In the second stage, weights should be assigned to each forecast. Weights could be fixed or time-varying. Almost all of the cited studies in this essay develop weights that get updated over time. Several weighting approaches are reported in the literature. Empirical analyses in many studies highlight the advantage of these approaches and forecast combinations.

The first weighting approach is standard and uses a simple averaging implementing three methods. 1) Equal weight forecasts at each period; 2) the trimmed-mean that first discards the smallest and largest forecasts and then use equal weight forecasts at each period; 3) the median of panel of forecasts at each period. This approach has been used in Chan, Stock, and Watson (1999), Stock and Watson (2003), Stock and Watson (2004), Smith and Wallis (2009), Della Corte and Tsiakas (2012), Morales-Arias and Moura (2013), Hsiao and Wan (2014), and Kouwenberg et al. (2017) among other studies. These studies show that in general combined forecasts generated by the equal weight method are more precise compared to more complex weighting methods. Clearly, this can only be possible if the equal weight scheme is not a possible result of more complex approaches.

The second standard weighting approach is a linear model approach. Many studies in the forecasting literature use a linear forecast combination following the study by Granger and Ramanathan (1984). Elliott and Timmermann (2005), Genre, Kenny, Meyler, and Timmermann (2013), Morales-Arias and Moura (2013), and Kouwenberg et al. (2017) suggest using a linear regression with time-varying weights. These studies use either a rolling window approach or an expanding window approach[15] to estimate and update weights at each period. Their empirical results indicate that combining forecasts yields more accurate predictions compared to predictions from individual models. However, few studies, for instance, Granger and Ramanathan (1984) keep weights fixed overtimes; a reason could be that they only had 24 observations.

The third standard weighting approach assigns time-varying weights to forecasts from individual models based on the out-of-sample forecasting performance of underlying models. This approach uses the Mean Squared Prediction Error (MSPE) statistic. Bates and Granger (1969), Chan et al. (1999), Stock and Watson (2004), Della Corte and Tsiakas (2012), Morales-Arias and Moura (2013), Hsiao and Wan (2014), and Kouwenberg et al. (2017) use the approach and they argue that using the recent out-of-sample information can make forecasts most robust to structural breaks and model specifications.

Kouwenberg et al. (2017) use the *smooth* MSPE. Whereas, Stock and Watson (2004), Smith and Wallis (2009), Della Corte and Tsiakas (2012), and Genre et al. (2013) calculate *discounted* MSPE at each period for individual models and use them to assign weights. Morales-Arias and Moura (2013) propose assigning weights based on each individual model's rank. *Rank* is the rank of individual model based on its MSPE calculated at each period. They update MSPE, *Ranks*, and weights at each period using both rolling and expanding window approaches, respectively.

[15] A rolling approach makes use of a fixed-length window of data to re-estimate parameters, whereas an expanding (recursive) approach makes use of an increasing window to re-estimate parameters.

The fourth typical weighting approach is the smooth Bayesian information criterion (BIC), which has close linked to the Bayesian Model Averaging (BMA) approach use in the studies by Wright (2008) and Della Corte, Sarno, and Tsiakas (2009). In this approach, investors have a prior probability of which model is the best and compute posterior probabilities after receiving additional information. In this approach, weights calculated from BIC could be fixed or time-varying. This weighting approach focuses on the in-sample performance of individual models and assigns weights based on that.

Some studies including Stock and Watson (2004) and Kouwenberg et al. (2017) select models dynamically over time and assign time-varying weights. Therefore, both forecasting models and weights could change over time. Other studies such as Elliott and Timmermann (2005), Aiolfi, Capistrán, and Timmermann (2010), and Genre et al. (2013) use subjective survey forecasts in forecasts combination procedures. They only have survey forecasts, but they do not have access to individual forecasters' information sets.

Elliott and Timmermann (2005) consider combinations of subjective survey forecasts and time series regression forecasts. They obtain survey forecasts from the Survey of Professional Forecasters, and these forecasts are available from the Philadelphia Fed's Web site. Forecasts are based on the mean across survey participants. They also form forecasts from time-series regressions based on AR(p). The parameters of AR(p) are re-estimated at each period. They consider predictions of six macroeconomic variables, for example, the nominal Gross Domestic Product (GDP). Elliott and Timmermann (2005) assign weights to the forecasts (from the survey and AR(p)) using different approaches, for example, a simple averaging, a linear regression, and Markov switching weights.

There are two reasons to combine forecasts from various individual models. First, individual investors, corporations, and governments, among other stakeholders, face uncertainty about the foreign exchange market. Second, individual models may be poor predictors of changes in exchange rates. Still, they include some information.

This essay contributes to the literature by addressing model uncertainty through forecasts combination methods. I use standard weighting approaches, including equal weights, MSPE, linear regression, and new forecast combination approaches to predict monthly changes in exchange rates the US dollar versus 37 (advanced and emerging/developing) currencies using various models per currency pair.

In addition, this essay proposes a new weighting (combination) method to combine forecasts from individual models introduced in the second essay in this dissertation. In standard methods, statistics used to calculate weights measure models' ability to correctly forecast changes in exchange rates. Unlike these methods, this essay uses a statistic that measures a model's ability to predict the direction of changes in exchange rates and calls it Directional Prediction.

Furthermore, previous studies did not use linear regression (both linear combination and convex combination) methods with many individual models. In contrast to these studies, this essay uses linear combination and convex combination weighting methods to combine forecasts from various models using a regularization method, in particular, a Ridge regression. Unlike previous studies that considered data for only the United States or a few advanced and emerging/developing countries, this essay examines several countries.

The study by Elliott and Timmermann (2005) uses the Diebold-Mariano statistic, and the studies by Morales-Arias and Moura (2013) and Kouwenberg et al. (2017) use the Diebold-Mariano-McCracken statistic [McCracken (2007)] and the Clark-West statistic [Clark and West (2007)] for the out-of-sample evaluation of forecasts combinations. In addition, Kouwenberg et al. (2017) use Pesaran-Timmermann (PT) statistic [developed by Pesaran and Timmermann (1992)] for the forecast combinations evaluation.

For assessments, I compare out-of-sample forecasting performance of weighting approaches with a benchmark model, the White Noise (WN) model of $\Delta s_t$, assuming that the level of exchange rate, $S_t$, follows a random walk model. I use the following statistics, the ratio of Mean Squared Prediction Error (MSPE$_{ratio}$), Diebold-Mariano (DM) [developed by Diebold and Mariano (1995)], and PT statistics. I use the DM and PT statistics for forecast combinations but cautiously interpret p-values and significant cases because the asymptotic distribution of DM or PT may not remain the same in forecast combinations. In addition to statistical forecast evaluation measures, I use the Sharpe-ratio statistic to evaluate investment strategies, using the expected exchange rate changes to calculate investment returns.

The rest of the essay is organized as follows: Section 3.2 presents modeling approaches. Section 3.3 describes data sets used in individual models and Section 3.4 explains methodologies and implementations. Section 3.5 reports empirical results and Section 3.6 concludes.

## 3.2 Modeling Approaches

In this section, I first explain the general optimization framework used to determine weights and the common approaches that fall under the framework. Then, I present weighing approaches used by previous studies and introduce new weighting approaches[16].

### 3.2.1 Weight Assignment-General Framework

Suppose $y_t$ is a quantitative variable of interest and there are $N$ forecasts conditioned on information at $t-1$. Then, there is a vector of forecasts $\mathbf{\Delta\hat{s}} = (\hat{y}_{1t}, ..., \hat{y}_{Nt})'$, formed at time $t-1$, for $y_t$ where $t = T_0, ..., T_1$. The goal is to find a linear combination $\mathbf{w}_t = (\omega_{1t}, ..., \omega_{Nt})'$, an $N \times 1$ vector of weights to form a new forecast $\widehat{y}_{t+1} = \mathbf{w}_t'\mathbf{\hat{y}_{t+1}}$, $t = T_2, ..., T$ which is optimal in terms of a given objective function. Determining $\mathbf{w}_t$ requires a set of previous forecasts and realized values of $y_t$. A special case is to choose $\mathbf{w}_t = \mathbf{w}$ for $t >= T_2$ such that weights are fixed. In the literature, weights are often time-varying (which I discuss in the subsequent sections).

To be more specific, the objective is to find $\mathbf{w}_t$ to minimize a loss function denoted by $L$

$$\min_{\mathbf{w}_t} \; L(y_t, \widehat{y}_t). \tag{3.1.1}$$

To solve the above, a certain amount of data $\mathbf{D}_t$ consisting of forecasts and realized $y$'s are required to estimate the objective function and perform the optimization. For instance, for one period ahead out-of-sample forecasting at time $t$, I use only the forecasts up to time

[16] I am grateful to Professor Andrei Semenov for his comments and inputs, particularly development of the Directional Prediction combination method.

$t-1$ and their associated realized values. This implies that I use data available at time $t$ to estimate and optimize $L(y_t, \widehat{y}_t)$. Let $\hat{\mathbf{w}}_t'$ be the solution to equation (3.1.1), then the combined forecast for the next period will be $\hat{\mathbf{w}}_t'\hat{\mathbf{y}}_{t+1}$.

In this essay, I use three common weighting approaches and introduce a new approach as follows:

#### 3.2.1.1 Linear Combination

The linear model (linear combination and convex combination) has been used in several studies. Most studies use linear regression with time-varying weights. Each forecasting model can be viewed as an approximation of data-generating process for a particular time. Therefore, their ability to approximate may change over time because of structural breaks and changes in economic conditions. These changes in economic conditions motivate using time-varying weights.

For the linear combination case, $L_{Linear} = E(\Delta s_t - \Delta\hat{s}_{ct})^2$, with $\Delta s_t$ denotes a realized observation for changes in exchange rates and $\Delta\hat{s}_{ct}$ is a combined forecast given by $\Delta\hat{s}_{ct} = w_1\Delta\hat{s}_{1t} + w_2\Delta\hat{s}_{2t} + \cdots + w_N\Delta\hat{s}_{Nt}$, is the loss function to be minimized in order to provide weights. More specifically, the linear regression model is as follows:

$$\Delta s_t = \omega_0 + \mathbf{w}'\mathbf{\Delta\hat{s}_t} + u_t, \tag{3.1.2}$$

where $\Delta s_t$ denotes the realized observations of changes in exchange rates, $\omega_0$ is an intercept, $\mathbf{w}$ is a N-Vector of regression coefficients or weights ($\mathbf{w} = (\omega_1, ..., \omega_N)'$), $\mathbf{\Delta\hat{s}_t}$ is N-Vector of forecasts formed by individual models $i = 1, ..., N$ ($\mathbf{\Delta\hat{s}_t} = (\Delta\hat{s}_{1t}, ..., \Delta\hat{s}_{Nt})'$), and $u_t$ is the error term.

#### 3.2.1.2 Convex Combination

For the convex combination, the loss function to be minimized is $L_{Convex} = E(\Delta s_t - \Delta\hat{s}_{ct})^2$ subject to weights being positive and adding up to 1. Hence, weights are obtained from constrained regression as follows:

$$\Delta s_t = \omega_0 + \mathbf{w}'\mathbf{\Delta\hat{s}_t} + u_t \;\; subject\ to \sum_{i=1}^{N} \omega_i = 1 \;\; and \;\; 0 \leq \omega_i \leq 1, \tag{3.1.3}$$

where $\Delta s_t$ denotes the realized observations of changes in exchange rates, $\omega_0$ is an intercept, $\mathbf{w}$ is an N-Vector of regression coefficients or weights ($\mathbf{w} = (\omega_1, ..., \omega_N)'$). $\mathbf{\Delta\hat{s}_t}$ is N-Vector of forecasts formed by individual models $i = 1, ..., N$, and $u_t$ is the error term.

In equations (3.1.2) and (3.1.3), $\omega_0$ and $\mathbf{w}$ are estimated using a fixed-rolling-window approach and the information up to $t$ (the forecasts formed up to and including time $t-1$ and their associated realized values). I explain the detailed methodologies and implementations of these approaches in Section 3.4.

#### 3.2.1.3 Equal Weights

Earlier studies show that in general forecasts generated by equal weighted combinations are more precise compared to more complex combination methods. Weights obtained by minimizing $L_{Equal} = E(\Delta s_t - \Delta\hat{s}_{ct})^2$ as the loss function, with $\Delta s_t$ denotes a realized observation for changes in exchange rates and $\Delta\hat{s}_{ct}$ is a combined forecast given by $\Delta\hat{s}_{ct} = \frac{1}{N}\Delta\hat{s}_{1t} + \frac{1}{N}\Delta\hat{s}_{2t} + \cdots + \frac{1}{N}\Delta\hat{s}_{Nt}$, and $N$ is the number of individual forecasts ($\Delta\hat{s}_t$) for a realized value, $\Delta s_t$, and weights are:

$$\omega_{i,t} = \frac{1}{N}, \quad i = 1, ..., N. \tag{3.1.4}$$

Weights are fixed over time. For example, if the number of individual models is 5. The weight for each model is $\frac{1}{5}$ over time. More details are provided in Section 3.4.

#### 3.2.1.4 Mean Squared Prediction Error (MSPE)

Bates and Granger (1969) suggest calculating weights using MSPE$_t$ measures computed from the past prediction errors of $\upsilon$ forecasts of $\Delta s$ prior to $t$. Let $MSPE_t^i = \frac{1}{\upsilon}\sum_{\tau=\upsilon-t}^{t} \hat{e}_{i,\tau}^2$ be the $i$th forecasting model's MSPE at time $t$. Then weights are

$$\omega_{i,t} = \frac{(MSPE_t^i)^{-1}}{\sum_{j=1}^{N}(MSPE_t^j)^{-1}}. \tag{3.1.5}$$

I update MSPE and weights at time $t$ using a fixed-rolling window approach. I present the detailed methodology of this approach in Section 3.4.

#### 3.2.1.5 Directional Prediction

As a new approach, I use a metric that measures the models' ability to accurately predict the direction (sign) of changes in exchange rates, i.e., the sign of $\Delta s_t$. Intuitively, the objective is to minimize the loss function $L_{DP} = E[sign(\Delta s_t) - sign(\Delta \hat{s}_{ct})]^2$, with $\Delta s_t$ denotes a realized observation for changes in exchange rates and $\Delta \hat{s}_{ct}$ is a combined forecast given by $\Delta \hat{s}_{ct} = \omega_1 \Delta \hat{s}_{1t} + \omega_2 \Delta \hat{s}_{2t} + \cdots + \omega_N \Delta \hat{s}_{Nt}$ and $N$ is the number of individual forecasts ($\Delta \hat{s}_t$) for a realized value, $\Delta s_t$. In this case, favorable outcomes would be combined forecasts and their associated realized values have the same sign. I propose equation (3.1.6), which is a heuristic that could lead to favorable outcomes, but not necessarily optimal ones. I calculate the weight for the $i$th forecast using the information up to $t$ and $\mathrm{CD}_t^i$ as follows:

$$\omega_{i,t} = \frac{CD_t^i}{\sum_{j=1}^{N} CD_t^j}, \quad i = 1, ..., N, \tag{3.1.6}$$

where $\omega_{i,t}$ is the weight for forecast $i$, which was formed by model $i$ at time $t$ (using the information up to and including time $t-1$), $\mathrm{CD}_t^i$ is the number of forecasts with correct directions for the $i$ at time $t$ (using the information up to and including time $t-1$), and $N$ is the number of individual forecasts ($\Delta \hat{s}_t$) for a realized value, $\Delta s_t$. Correct directions means that both a forecast and its associated realized value have the same sign. I update $CD_t^i$ and weights using the fixed-rolling window approach. I present the detailed methodology of this approach in Section 3.4.

I provide an example to explain the method more precisely. For example, individual forecasting models A, B, C form 100 forecasts when predicting 100 past values. Suppose for model A, the 65 out of 100 forecasts have correct signs, for model B, the 73 out of 100 forecasts have correct signs, and for model C, the 82 out of 100 forecasts have correct signs. Therefore, the $CD_t^A = 65$, $CD_t^B = 73$, and $CD_t^C = 82$.

To ensure $\omega_{A,t} + \omega_{B,t} + \omega_{C,t} = 1$, I calculate and normalize weights as follows: $\omega_{A,t} = \frac{65}{220}$,

$\omega_{B,t} = \frac{73}{220}$, and $\omega_{C,t} = \frac{82}{220}$. Therefore, weights are $\omega_{A,t} = 0.3$, $\omega_{B,t} = 0.33$, and $\omega_{C,t} = 0.37$. Clearly, $CD_t = 0$ and therefore $\omega_t = 0$ for an individual model with no correct sign forecast.

## 3.3 Data Description

I combine forecasts formed by individual models, which have been presented and introduced in the second essay of this dissertation [see details in Section 2.2]. Data sets used in individual models and methodologies to prepare variables for empirical analyses are explained in the second essay of this dissertation [see details in Section 2.3].

## 3.4 Methodology and Implementation

This section provides a brief review of individual models and explains estimation/calculation methods for weights and forecasts combinations. In particular, this section describes the following:

- Individual Forecasting Models
- Estimate/Calculate Weights and Form Combined Forecasts
- Performance Measures and Statistics

### 3.4.1 Individual Forecasting Models

In this section, I provide a brief review of individual forecasting models. As I extensively explained in the second essay, I used and evaluated the out-of-sample performance of macroeconomic models (11 models), where data were available for a given country and horizon. For individual forecasting models, criteria were MSPE, DM, and PT statistics. Here, a criterion is $\hat{L}$ at minimum, $\hat{L}_{min} \approx MSPE$ for an entire set of forecasting models. Explanations and motivations for using/introducing each model are provided in Section 2.2.

I added behavioral factors (investor sentiment indices) to macroeconomic models and created behavioral models. Based on data availability, there are up to 5 behavioral factors, and thus there are up to $11 \times 31 = 341$ behavioral models for a given country and horizon. There are up to 352 (341 behavioral models + 11 non-behavioral models) individual forecasting models for a given country and horizon from the second essay. I consider short (horizon=1) and long (horizon=12) horizons. I implement forecast combination of all models per country.

Tables 2.3 and 2.4 present macroeconomic models and combinations of behavioral factors (investor sentiment indices), respectively.

### 3.4.2 Estimate/Calculate Weights and Form Combined Forecasts

All individual models for a given country have the same number of forecasts for the same out-of-sample period. In other words, each model for a given country provides forecasts for every future value in the out-of-sample period. In horizon=1, all countries' forecasts start from Jan 2005 to Dec 2020. I use a fixed-length of 60 observations rolling window method to estimate/calculate weights for each of models in the below weighting approaches. A reason to use a fixed-length rolling window method is that there are changes over time, such as regime switching or financial crises. Also, this approach is consistent with ones used in the first and second essays. Because the number of forecasting models is high, I adjust the procedures of weight selection referred to as linear and convex combination approaches and explained in Sections 3.2.1.1 and 3.2.1.2 for collinearity by using a Ridge regression approach.

I use the following weighting approaches to calculate/estimate weights:

1. Linear Regression (Linear Combination with a Ridge Regression)

2. Linear Regression (Convex Combination with a Ridge Regression)

3. Equal Weights

4. Mean Squared Prediction Error (MSPE)

5. Directional Prediction

#### 3.4.2.1 Linear Regression (Linear Combination with a Ridge Regression)

Here, I explain the regression model used in this essay:

$$\Delta s_t = \omega_0 + \mathbf{w}' \, \mathbf{\Delta \hat{s}_t} + u_t, \tag{3.1.1}$$

where $\Delta s_t$ denotes realized observations for changes in exchange rates, $\omega_0$ is an intercept, $\mathbf{w}$ is an N-Vector of regression coefficients or weights ($\mathbf{w} = (\omega_1, ..., \omega_N)'$), $\mathbf{\Delta \hat{s}_t}$ is N-Vector of forecasts formed by selected models $i = 1, ..., N$ ($\mathbf{\Delta \hat{s}_t} = (\Delta \hat{s}_{1,t}, ..., \Delta \hat{s}_{N,t})'$), $u_t$ is the error term.

As I explained in Section 3.4.1, there are up to 352 individual forecasting models for a given country and horizon. When there are up to 352 models used in the models' combination process, in fact, there are up to 352 regressors in equation (3.1.1) for a given country. Because the number of regressors becomes large, there may be over-fitting problems. According to Kuhn and Johnson (2013), when there are over-fitting problems, or when there are issues with collinearity, the linear regression parameter estimates may become inflated[17]. The magnitude of these estimates could be regularized to reduce the sum of the squared errors (SSE). Regularizing parameter estimates can be achieved by adding a penalty to the SSE. Hoerl and Kennard (1970) introduce a Ridge regression, which adds a penalty on the sum of the squared regression parameters as follows:

$$SSE_{L_2} = \sum_{i=1}^{M} (\Delta s_i - \sum_{j=1}^{N} \omega_j \Delta \hat{s}_{ij})^2 + \lambda \sum_{j=1}^{N} \omega_j^2, \tag{3.1.2}$$

where SSE is the sum of the squared error, $\Delta s_i$ is a realized observation for changes in exchange rates, $\omega_j \Delta \hat{s}_{ij}$ is a combined forecast, $\lambda$ is a hyperparameter, and $\omega_j$ is a weight for a forecast. The $L_2$ emphasizes that a second-order penalty is being used on parameter estimates. In the Ridge method, estimates shrink toward 0 as the value of $\lambda$ increases. By adding the penalty, there is a trade-off between the model variance and bias. By introducing some bias, the variance could be reduced enough to make the overall MSE lower than unbiased models [Kuhn and Johnson (2013)].

[17] Putting restrictions on coefficients in the convex combination estimation mitigates the problem of inflated parameters to some degree.

I examine different values of $\lambda$ to find optimal values used for all countries in each horizon, $\lambda = (1e^{-6}, 1e^{-5}, 1e^{-3}, 1e^{-2}, 1e^{-1}, 0.1, 0.5, 1, 2.5, 5, 10, 15, 20, 25, 30, 35, 40, 50)$. Optimal values are $\lambda$=2.5 at horizon=1 and $\lambda$=30 at horizon=12, which lead to the highest average values of proportions of correct directions (i.e., the highest average value of $\hat{p}$) for exchange rates across all countries.

I should point out a Lasso regression (an alternative regularization technique) allows some parameters to be estimated as 0. Therefore, there is a selection of models, and some models would be eliminated in prediction procures. I leave this technique for future research.

After finding the optimal value for $\lambda$, I estimate weights in equation (3.1.1) using the penalty term. I should emphasize that the results are not very sensitive to the choice of $\lambda$. Specifically, I take the following steps to estimate weights in equation (3.1.1) and calculate combined forecasts:

1. First, I use forecasts and their associated realized values from Jan 2005 to Dec 2009 (60 forecasts) to estimate weights in equation above. I apply estimated weights to forecasts at date Jan 2010 to form a combined forecast for the date Jan 2010 using the following equation:

$$\Delta \hat{s}_{combine,2010M1} = \hat{\omega}^L_{0,2009M12} + \sum_{i=1}^{N} \hat{\omega}^L_{i,2009M12}\ \Delta \hat{s}_{i,2010M1}, \tag{3.1.3}$$

   where $\hat{\omega}^L_{0,2009M12}$ is the estimated intercept (using the information up to and including time Dec 2009 and the linear estimation), $\omega^L_{i,2009M12}$ is the estimated weight for the forecast $i$ at Jan 2010 (using the information up to and including time Dec 2009 and the linear estimation), $\Delta \hat{s}_{i,2010M1}$ is the forecast $i$ at Jan 2010, and $N$ is the number of individual forecasts ($\Delta \hat{s}_t$) for a realized value, $\Delta s_t$.

2. Then, I roll forecasts forward by one and use forecasts and their associated realized values from Feb 2005 to Jan 2010 to estimate weights in equation (3.1.1). Therefore, I use 60 forecasts as in the previous step. I apply estimated weights to forecasts at Feb 2010 to form a combined forecast for Feb 2010 using the following equation:

$$\Delta\hat{s}_{combine,2010M2} = \hat{\omega}^L_{0,2010M1} + \sum_{i=1}^{N} \hat{\omega}^L_{i,2010M1}\ \Delta\hat{s}_{i,2010M2}. \tag{3.1.4}$$

3. I repeat the above steps until I reach the end of the data (forecasts). I use the rolling window of length 60 forecasts to estimate weights and calculate combined forecasts.

#### 3.4.2.2 Linear Regression (Convex Combination with a Ridge Regression)

Here, I explain the constrained regression model used in this essay:

$$\Delta s_t = \omega_0 + \mathbf{w}'\ \mathbf{\Delta\hat{s}_t} + u_t\ \ \textit{subject to} \sum_{i=1}^{N} \omega_i = 1\ \ \textit{and}\ \ 0 \leq \omega_i \leq 1, \tag{3.1.5}$$

where $\Delta s_t$ denotes realized observations for changes in exchange rates, $\omega_0$ is an intercept, $\mathbf{w}$ is an N-Vector of regression coefficients or weights ($\mathbf{w} = (\omega_1, ..., \omega_N)'$), $\mathbf{\Delta\hat{s}_t}$ is N-Vector of forecasts formed by the selected models $i = 1, ..., N$ ($\mathbf{\Delta\hat{s}_t} = (\Delta\hat{s}_{1,t}, ..., \Delta\hat{s}_{N,t})'$). The number of individual forecasting models is the same as the previous section. I use equation (3.1.2) and examine different values of $\lambda$ to find the optimal values used for all countries in each horizon. The values of $\lambda$ to examine are the same as in the previous section.

The optimal values are $\lambda$=0.5 and $\lambda$=10 at horizon=1 and horizon=12, respectively, leading to the highest average values of proportions of correct directions (i.e., the highest average value of $\hat{p}$) for exchange rates across all countries. After finding the optimal value for $\lambda$, I estimate weights in the equation above using the penalty term. I should emphasize that the results are not very sensitive to the choice of $\lambda$.

Taking the below steps, I estimate weights in equation (3.1.5) and calculate combined forecasts:

1. Using forecasts and their associated realized values from Jan 2005 to Dec 2009 (60 forecasts), I estimate weights in the equation above. I apply estimated weights to forecasts at Jan 2010 to form a combined forecast for Jan 2010 using the following equation:

$$\Delta\hat{s}_{combine,2010M1} = \hat{\omega}^{C}_{0,2009M12} + \sum_{i=1}^{N} \hat{\omega}^{C}_{i,2009M12}\ \Delta\hat{s}_{i,2010M1}, \tag{3.1.6}$$

where $\hat{\omega}^{C}_{0,2009M12}$ is the estimated intercept (using information up to and including time Dec 2009 and the convex estimation method), $\omega^{C}_{i,2009M12}$ is the estimated weight for the forecast $i$ at Jan 2010 (using the information up to and including time Dec 2009 and the convex estimation method), $\Delta\hat{s}_{i,2010M1}$ is the forecast $i$ at Jan 2010, and $N$ is the number of individual forecasts ($\Delta\hat{s}_t$) for a realized value, $\Delta s_t$.

2. Rolling forecasts forward by one, I use forecasts and their associated realized values from Feb 2005 to Jan 2010 to estimate weights in equation (3.1.5). Therefore, I use 60 forecasts as in the previous step. I apply estimated weights to forecasts at Feb 2010 to form a combined forecast for Feb 2010 using the following equation:

$$\Delta\hat{s}_{combine,2010M2} = \hat{\omega}^{C}_{0,2010M1} + \sum_{i=1}^{N} \hat{\omega}^{C}_{i,2010M1}\ \Delta\hat{s}_{i,2010M2}. \tag{3.1.7}$$

3. Repeating the above steps, I reach the end of the data (forecasts). I use the rolling window of length 60 forecasts to estimate weights and calculate combined forecasts.

#### 3.4.2.3 Equal Weights

Here, I explain the equal weighting approach in this essay. Weights are calculated as follows:

$$\omega_{i,t} = \frac{1}{N}, \quad i = 1, ..., N, \tag{3.1.8}$$

where $N$ is the number of individual forecasts ($\Delta\hat{s}_t$) for a realized value, $\Delta s_t$. Weights are fixed over time. For example, if the number of individual forecasts is 32. The weight for each forecast is $\frac{1}{32}$ over time. In this approach, I do not need to use any sub-sample of forecasts to calculate weights. This means I can calculate combined forecasts from Jan 2005. However, because I would like all weighting approaches to have the same number of combined forecasts for evaluation (same number of forecasts used in the out-of-sample statistics), I calculate combined forecasts from Jan 2010 using the following equation:

$$\Delta\hat{s}_{combine,t} = \sum_{i=1}^{N} \omega_i^E \ \Delta\hat{s}_{i,t}, \tag{3.1.9}$$

where $\omega_i^E$ is the weight for the forecast $i$ using the equal-weight weighting method and $\Delta\hat{s}_{i,t}$ is the forecast $i$ at time $t$.

#### 3.4.2.4 Mean Squared Prediction Error (MSPE)

Here, I present the MSPE weighting approach and explain the methodology. Weights are given by:

$$\omega_{i,t} = \frac{(MSPE_t^i)^{-1}}{\sum_{j=1}^{N}(MSPE_{j,t})^{-1}}, \quad i = 1, ..., N, \tag{3.1.10}$$

where $\omega_{i,t}$ is the weight for forecast $i$, which was formed by model $i$ at time $t$ (using the information up to and including time $t-1$), $\text{MSPE}_t^i$ is the mean squared prediction error for model $i$ at time $t$, and $N$ is the number of individual forecasts ($\Delta\hat{s}_t$) for a realized value, $\Delta s_t$. I update MSPE and weights at time $t$ using the fixed-length rolling window approach. I use the below procedures to calculate weights using equation above and calculate combined forecasts:

1. I initially use forecasts and their associated realized values from Jan 2005 to Dec 2009 (60 forecasts) to calculate $\text{MSPE}^i$. Then, I use $\text{MSPE}^i$ in equation above to calculate $\hat{\omega}_i^M$. I apply calculated weights to forecasts at Jan 2010 to form a combined forecast at Jan 2010 using the following equation:

$$\Delta\hat{s}_{combine,2010M1} = \sum_{i=1}^{N} \hat{\omega}_{i,2009M12}^M \ \Delta\hat{s}_{i,2010M1}, \tag{3.1.11}$$

where $\hat{\omega}_{i,2009M12}^M$ is the calculated weight for forecast $i$ at time Jan 2010 (using the information up to and including time Dec 2009 and using the MSPE weighting approach), $\Delta\hat{s}_{i,2010M1}$ is the forecast $i$ at time Jan 2010, and $N$ is the number of individual forecasts ($\Delta\hat{s}_t$) for a realized value, $\Delta s_t$.

2. Next, I calculate MSPE$^i$ for each model $i$ by rolling the forecasts forward by one and use forecasts and their associated realized values from Feb 2005 to Jan 2010. I use the MSPE$^i$ for each model $i$ in equation (3.1.10) to calculate $\hat{\omega}_i^M$. I apply calculated weights to forecasts at date Feb 2010 to form a combined forecast for the date Feb 2010 using the following equation:

$$\Delta\hat{s}_{combine,2010M2} = \sum_{i=1}^{N} \hat{\omega}_{i,2010M1}^{M} \ \Delta\hat{s}_{i,2010M2}. \tag{3.1.12}$$

3. I reach the end of the data (forecasts) by repeating the above procedures. I use the rolling window of length 60 forecasts to calculate weights and combined forecasts.

#### 3.4.2.5 Directional Prediction

Here, I present the new approach, thus I use the proportion of correct direction of change of exchange rates to calculate a weight for the $i$th forecast using the information up to $t$ and CD$_t^i$ as follows:

$$\omega_{i,t} = \frac{CD_t^i}{\sum_{j=1}^{N} CD_{j,t}}, \quad i = 1, ..., N, \tag{3.1.13}$$

where $\omega_{i,t}$ is the weight for forecast $i$, which was created by model $i$ at time $t$ (using the information up to and including time $t-1$), CD$_t^i$ is the number of forecasts with correct directions for the model $i$, and $N$ is the number of individual forecasts ($\Delta\hat{s}_t$) for a realized value, $\Delta s_t$. As a reminder, the correct directions (sign) mean both a forecast and its associated realized value have the same sign (i.e., both are negative or positive). I provided an example in Section 3.2.1.5 to make this weighting approach more clear. I update CD$^i$ and weights at time $t$ using the fixed-length rolling window approach. Following the below steps, I compute weights in equation above and calculate combined forecasts:

1. Utilizing the forecasts from Jan 2005 to Dec 2009 (60 forecasts), I calculate the value of CD$^i$, and use the value in equation above to calculate $\hat{\omega}_i$. I apply the calculated

weights to the forecasts at Jan 2010 to form the combined forecast for Jan 2010 using the following equation:

$$\Delta\hat{s}_{combine,2010M1} = \sum_{i=1}^{N} \hat{\omega}_{i,2009M12}^{DP}\ \Delta\hat{s}_{i,2010M1}, \tag{3.1.14}$$

where $\hat{\omega}_{i,2009M12}^{DP}$ is the calculated weight for the forecast $i$, which was formed by model $i$ at time Jan 2010 (using the information up to and including time Dec 2009 and the Directional Prediction weighting method), $\Delta\hat{s}_{i,2010M1}$ is the forecast for the model $i$ at time Jan 2010, and $N$ is the number of individual forecasts ($\Delta\hat{s}_t$) for a realized value, $\Delta s_t$.

2. Moving forecasts forward by one, I use forecasts from Feb 2005 to Jan 2010 to compute $\text{CD}^i$ and use it in equation (3.1.13) to calculate $\hat{\omega}_i$. I apply calculated weights to forecasts at date Feb 2010 to form a combined forecast for the date Feb 2010 using the following equation:

$$\Delta\hat{s}_{combine,2010M2} = \sum_{i=1}^{N} \hat{\omega}_{i,2010M1}^{DP}\ \Delta\hat{s}_{i,2010M2}. \tag{3.1.15}$$

3. Iterating the above processes and rolling the fixed window of length 60 forecasts, I reach the end of forecasts sets.

In the next section, I explain statistics used to evaluate weighting approaches.

### 3.4.3 Performance Measures and Statistics

As explained in Section 3.1, Elliott and Timmermann (2005) use the Diebold-Mariano statistic for the out-of-sample evaluation of combined forecasts. However, they emphasize that the DM statistic should be interpreted with caution and results are best viewed as a diagnostic because the asymptotic distribution of the test statistic relies on a sampling experiment that a researcher considers. As mentioned in Section 3.1, Kouwenberg et al. (2017) use PT statistics for out-of-sample evaluation of combined forecasts. However, I believe the PT statistic should be cautiously interpreted as well since the asymptotic distribution of the test statistic is unknown.

To evaluate the out-of-sample performance of forecasting models with the White Noise (WN) (the benchmark) model, I use statistics, including the ratio of Mean Squared Prediction Error (MSPE$_{ratio}$)$\approx \hat{L}_{min}$ of the combined forecasts, Diebold-Mariano (DM), and the Pesaran-Timmermann (PT). In addition to these statistics, I use the Sharpe-ratio statistic to evaluate investment strategies, using the expected exchange rate changes to calculate investment returns. Explanations and motivations for using these statistics are provided in Section 1.4.3.

I use equations (1.4.19), (1.4.20), and (1.4.21) to calculate the MSPE$_{WN}$ (the MSPE statistic for the WN model), the MSPE$_{M}$ (the MSPE statistic for a weighting approach), and the MSPE$_{ratio}$ (the ratio of MSPE for a weighting approach over MSPE of the benchmark). MSPE$_{ratio} < 1$ implies that a forecasting model provides better out-of-sample performance than the benchmark model.

I use DM statistics for forecast combinations but cautiously interpret the p-values and significant cases. I use equation (1.4.22) to calculate $\hat{d}_{t+h}$. The DM statistic can be calculated by regression of the loss differential ($\hat{d}_{t+h}$) on a constant using HAC approaches to correct standard errors for autocorrelation and heteroskedasticity. I calculate the p-values associated with the DM statistics and set the 10% level of significance. I calculate the PT statistics using equation (1.4.23) and the p-values associated with the PT statistics. The proportion of correct direction larger than 0.5 (above 50%) indicates a better forecasting performance than the WN model. I interpret the p-values and significant cases with caution.

I calculate returns, $r_t$, for investment strategies using equation (1.4.24) and I regress $r_t$ on a constant and use the HAC approach to correct the standard errors for autocorrelation and heteroskedasticity. Using the estimated constant and an adjusted standard error, I calculate an annualized Sharpe-ratio for each weighting method. I also report the annualized average return in percentage for each weighting method for a given country.

I should emphasize to be able to compare the out-of-sample performance of individual forecasting models with weighted combined forecasts, I use the same out-of-sample periods (thus, the same number of forecasts) in the statistics for both sets of models.

## 3.5 Empirical Results

This section describes the out-of-sample forecasting performance of weighting approaches from the statistical and economic perspectives. I report the empirical results at horizon=1 and horizon=12. Recall that p-values need to be interpreted with caution if asymptotic validity is not known. Table 3.1 reports the median of statistics for a given weighting approach, the median of individual models[18], and the median of statistics for the best models in each country category at horizon=1.

The results show, on average, using the linear combination (Ridge regression) weighting method leads to the proportion of correct direction of changes in exchange rates greater than 0.5 for advanced countries. Still, on average, the results are not statistically significant at the 10% significance level. On average, none of weighting approaches provides better forecasting results than the best individual models, and no weighting method has a $\text{MSPE}_{ratio}$ less than one and a positive DM statistic at horizon=1.

In contrast to advanced countries' results, on average, incorporating all weighting methods improves the forecasting results for exchange rates changes in emerging/developing countries at horizon=1. However, on average, none of weighting methods provides $\text{MSPE}_{ratio} < 1$ and a positive DM statistic.

[18] First, I calculate the median of each test statistic (e.g. the $\text{MSPE}_{ratio}$) across the individual models for a given country. There are 26 advanced countries; therefore, there are 26 $\text{MSPE}_{ratio}$, DM, $\hat{p}$, PT(t-stats), average returns, and Sharpe-ratio statistics. Next, I calculate the median of these 26 statistics, such as 26 DM statistics. I report the numbers in Table 3.1. I repeat similar steps for emerging/developing countries.

Table 3.1: Summary results in horizon=1

| WEO country group | Weighting Approaches | MSPE_ratio | DM (t-stats) | DM (p-value) | $\hat{p}$ | PT (t-stats) | PT (p-value) | Average Return (%) | Sharpe_ratio |
|---|---|---|---|---|---|---|---|---|---|
| | Linear Combination (Ridge regression) | 1.0091 | -0.6610 | 0.2561 | 0.5344 | 0.7863 | 0.2158 | 0.7816 | 0.0978 |
| | Convex Combination (Ridge regression) | 1.0849 | -1.7734 | 0.0393 | 0.4962 | -0.0874 | 0.5348 | -0.1943 | -0.0230 |
| | MSPE | 1.0708 | -1.7058 | 0.0453 | 0.4885 | -0.2621 | 0.6034 | -1.1593 | -0.1199 |
| **Advanced** | Directional Prediction | 1.0728 | -1.7541 | 0.0409 | 0.4885 | -0.2621 | 0.6034 | -1.2194 | -0.1301 |
| | Equal Weight | 1.0713 | -1.7148 | 0.0444 | 0.4809 | -0.4369 | 0.6689 | -1.1982 | -0.1401 |
| | Median of Individual Models | 1.1425 | -2.1880 | 0.0152 | 0.4885 | -0.2621 | 0.6034 | -0.4332 | -0.0614 |
| | Best Models | 1.1011 | -1.4521 | 0.0744 | 0.5802 | 1.8348 | 0.0333 | 3.2060 | 0.3764 |
| | Linear Combination (Ridge regression) | 1.0090 | -0.3594 | 0.3412 | 0.5581 | 1.3207 | 0.0933 | 4.4002 | 0.3627 |
| | Equal Weight | 1.0509 | -1.0227 | 0.1542 | 0.5344 | 0.7863 | 0.2158 | 2.9800 | 0.3505 |
| | MSPE | 1.0517 | -1.0135 | 0.1564 | 0.5267 | 0.6116 | 0.2704 | 2.7522 | 0.3561 |
| **Emerging/Developing** | Directional Prediction | 1.0517 | -1.0205 | 0.1547 | 0.5267 | 0.6116 | 0.2704 | 3.2535 | 0.3561 |
| | Convex Combination (Ridge regression) | 1.0864 | -1.2184 | 0.1127 | 0.5115 | 0.2621 | 0.3966 | 2.5475 | 0.3610 |
| | Median of Individual Models | 1.1070 | -1.4040 | 0.0814 | 0.5076 | 0.1747 | 0.4309 | 0.1886 | 1.6990 |
| | Best Models | 1.0137 | -0.2125 | 0.2018 | 0.5772 | 1.7132 | 0.0433 | 6.1423 | 0.5723 |

Note: The column "WEO Country Group" shows the World Economic Outlook (WEO) countries' classifications. The column "Weighting Approaches" defines the underlying model or weighting approach. The column "MSPE_ratio" indicates the median of the $\text{MSPE}_{ratio}$ for each weighting approach, the median of individual models, the best models within each country group. The column "DM(t-stats)" shows the median of DM statistics for each weighting approach, the median of individual models, the best models within each country group. The column "DM(p-value)" shows the median of p-values associated with the DM statistics. The column "$\hat{p}$" indicates the median of proportions of correct predictions of direction for each weighting approach, the median of individual models, the best models within each country group. The column "PT(t-stats)" shows the median of Pesaran-Timmermann (PT) statistics for each weighting approach, the median of individual models, the best models within each country group and the column "PT(p-value)" shows the median of p-values associated with the PT statistics. All reported test statistics are one-sided. The significance level is 10%. The column "Avg. Return (%)" shows the median of annualized average returns in percentage, and the column "Sharp_ratio" shows the median of annualized Sharp-ratio values. To be clear, the median of individual models means I first calculate the median of underlying statistics using all statistics from all individual models for a given country, then calculate the median of the median of underlying statistics across countries in each country group.

Based on the evidence in Table 3.1, the linear combination (Ridge regression) is the most successful method in both advanced and emerging/developing countries. Therefore, to save space, I only report the detailed results for this weighting approach at horizon=1.

Tables 3.2 and 3.3 report the out-of-sample statistics for the linear combination (Ridge regression) weighting approach, the median of individual models, and the best individual model for a given advanced country at horizon=1. The findings show that the linear combination weighting approach for 77% of countries provides the correct direction of changes greater than 0.5 ($\hat{p} > 0.5$) and for the 23% of countries statistically significant at the 10% significance level. Recalling that there are up to 352 individual models for a given country. Therefore, there are up to 352 statistics, for example, the 352 $\text{MSPE}_{ratio}$ statistics. I calculate the median of these 352 statistics and report it in the row of "Median of Individual Models."

## Table 3.2: Empirical results for advanced countries in horizon=1

| Countries | Weighting Approaches | MSPE_ratio | DM (t-stats) | DM (p-value) | $\hat{p}$ | PT (t-stats) | PT (p-value) | Average Return (%) | Sharpe_ratio |
|---|---|---|---|---|---|---|---|---|---|
| Austria | Linear Combination (Ridge regression) | 1.0063 | -0.5006 | 0.3088 | 0.5420 | 0.9611 | 0.1683 | -0.4001 | -0.0462 |
| Austria | Median of Individual Models | 1.1419 | -2.2746 | 0.0123 | 0.5115 | 0.2621 | 0.3966 | 0.1457 | 0.0174 |
| Austria | Best Individual Model | 1.0661 | -1.0008 | 0.1594 | 0.5954 | 2.1843 | 0.0145 | 5.8135 | 0.6536 |
| Belgium | Linear Combination (Ridge regression) | 1.0049 | -0.4173 | 0.3386 | 0.5191 | 0.4369 | 0.3311 | -0.2083 | -0.0250 |
| Belgium | Median of Individual Models | 1.1733 | -2.3264 | 0.0108 | 0.4809 | -0.4369 | 0.6689 | -1.2922 | -0.1467 |
| Belgium | Best Individual Model | 1.0133 | -0.3635 | 0.3584 | 0.5878 | 2.0095 | 0.0222 | 5.7746 | 0.5811 |
| Canada | Linear Combination (Ridge regression) | 1.0173 | -1.0220 | 0.1544 | 0.5820 | 1.8107 | 0.0351 | 0.8367 | 0.1064 |
| Canada | Median of Individual Models | 1.1538 | -2.2936 | 0.0118 | 0.5328 | 0.7243 | 0.2344 | 0.6795 | 0.0912 |
| Canada | Best Individual Model | 1.1755 | -2.0417 | 0.0217 | 0.6475 | 3.2593 | 0.0006 | 5.8545 | 0.7274 |
| Czech Republic | Linear Combination (Ridge regression) | 1.0256 | -1.4982 | 0.0682 | 0.4656 | -0.7863 | 0.7842 | 0.4123 | 0.0368 |
| Czech Republic | Median of Individual Models | 1.1520 | -2.6157 | 0.0050 | 0.4733 | -0.6116 | 0.7296 | -1.7846 | -0.1651 |
| Czech Republic | Best Individual Model | 1.1425 | -2.7453 | 0.0035 | 0.5725 | 1.6600 | 0.0485 | 3.1750 | 0.3022 |
| Denmark | Linear Combination (Ridge regression) | 1.0043 | -0.3275 | 0.3719 | 0.5649 | 1.4853 | 0.0687 | 2.7916 | 0.3192 |
| Denmark | Median of Individual Models | 1.1331 | -2.1962 | 0.0149 | 0.4809 | -0.4369 | 0.6689 | -0.3703 | -0.0450 |
| Denmark | Best Individual Model | 1.1052 | -1.4536 | 0.0742 | 0.5573 | 1.3106 | 0.0950 | 2.4672 | 0.2875 |
| Estonia | Linear Combination (Ridge regression) | 1.0011 | -0.0746 | 0.4703 | 0.5496 | 1.1358 | 0.1280 | 1.0025 | 0.1199 |
| Estonia | Median of Individual Models | 1.0873 | -1.3735 | 0.0860 | 0.5038 | 0.0874 | 0.4652 | 0.6369 | 0.0732 |
| Estonia | Best Individual Model | 1.0229 | -0.4300 | 0.3340 | 0.5802 | 1.8348 | 0.0333 | 3.2369 | 0.3810 |
| Finland | Linear Combination (Ridge regression) | 1.0104 | -0.8175 | 0.2076 | 0.5115 | 0.2621 | 0.3966 | -0.6806 | -0.0767 |
| Finland | Median of Individual Models | 1.1208 | -2.0792 | 0.0198 | 0.4885 | -0.2621 | 0.6034 | 0.0072 | 0.0008 |
| Finland | Best Individual Model | 1.0971 | -1.8828 | 0.0310 | 0.5802 | 1.8348 | 0.0333 | 2.3075 | 0.3420 |
| France | Linear Combination (Ridge regression) | 1.0079 | -0.5525 | 0.2908 | 0.5231 | 0.5262 | 0.2994 | -0.3249 | -0.0371 |
| France | Median of Individual Models | 1.1376 | -2.5543 | 0.0059 | 0.4580 | -0.9611 | 0.8317 | -1.4248 | -0.1712 |
| France | Best Individual Model | 1.0670 | -1.0585 | 0.1459 | 0.5725 | 1.6600 | 0.0485 | 5.4618 | 0.6302 |
| Germany | Linear Combination (Ridge regression) | 1.0120 | -0.9155 | 0.1808 | 0.5191 | 0.4369 | 0.3311 | 0.0219 | 0.0024 |
| Germany | Median of Individual Models | 1.1045 | -1.7651 | 0.0399 | 0.4885 | -0.2621 | 0.6034 | -0.1811 | -0.0243 |
| Germany | Best Individual Model | 1.0515 | -1.5225 | 0.0652 | 0.6031 | 2.3590 | 0.0092 | 3.5665 | 0.4738 |
| Greece | Linear Combination (Ridge regression) | 1.0067 | -0.4494 | 0.3270 | 0.5344 | 0.7863 | 0.2158 | 1.1533 | 0.1305 |
| Greece | Median of Individual Models | 1.1303 | -1.9271 | 0.0281 | 0.5038 | 0.0874 | 0.4652 | 0.2669 | 0.0335 |
| Greece | Best Individual Model | 1.1501 | -1.8792 | 0.0312 | 0.5802 | 1.8348 | 0.0333 | 3.7044 | 0.4511 |
| Iceland | Linear Combination (Ridge regression) | 1.0610 | -2.6417 | 0.0046 | 0.4351 | -1.4853 | 0.9313 | -5.0485 | -0.5210 |
| Iceland | Median of Individual Models | 1.1851 | -2.6290 | 0.0048 | 0.5153 | 0.3495 | 0.3639 | 1.4966 | 0.1460 |
| Iceland | Best Individual Model | 1.1943 | -2.1337 | 0.0174 | 0.5802 | 1.8348 | 0.0333 | 3.6532 | 0.3551 |
| Ireland | Linear Combination (Ridge regression) | 1.0048 | -0.3790 | 0.3527 | 0.5573 | 1.3106 | 0.0950 | 2.8567 | 0.3240 |
| Ireland | Median of Individual Models | 1.1327 | -2.2456 | 0.0132 | 0.4733 | -0.6116 | 0.7296 | -1.5490 | -0.1850 |
| Ireland | Best Individual Model | 1.1148 | -1.8717 | 0.0317 | 0.5496 | 1.1358 | 0.1280 | 1.0743 | 0.1377 |
| Israel | Linear Combination (Ridge regression) | 1.0210 | -1.1001 | 0.1367 | 0.4651 | -0.7924 | 0.7859 | -2.8881 | -0.4661 |
| Israel | Median of Individual Models | 1.3332 | -1.7396 | 0.0422 | 0.5115 | 0.2621 | 0.3966 | -0.4506 | -0.0715 |
| Israel | Best Individual Model | 1.7026 | -1.1128 | 0.1339 | 0.6031 | 2.3590 | 0.0092 | 2.8140 | 0.4822 |

Note: The column "Countries" shows the names of countries. The column "Weighting Approaches" defines the underlying models. The column "MSPE_ratio" indicates the $\text{MSPE}_{ratio}$ statistics for the Linear Combination (Ridge regression) weighting approach, the $\text{MSPE}_{ratio}$ statistics for the median of individual models, and the $\text{MSPE}_{ratio}$ statistics for the best individual model for a given country. The column "DM(t-stats)" shows the DM statistics for the Linear Combination (Ridge regression) weighting approach, the DM statistic for the median of individual models, and the DM statistics for the best individual model for a given country. The column "DM(p-value)" shows the p-values associated with the DM statistics. The column "$\hat{p}$" indicates the proportions of correct predictions of direction for the Linear Combination (Ridge regression) weighting approach, the proportions of correct predictions of direction for the median of individual models, and the proportions of correct predictions of direction for the best individual model. The column "PT(t-stats)" indicates the Pesaran-Timmermann (PT) statistics for the Linear Combination (Ridge regression) weighting approach, the PT statistics for the median of individual models, and the PT statistics for the best individual model and the column "PT(p-value)" shows the p-values associated with the PT statistics. All reported test statistics are one-sided. The significance level is 10%. The column "Avg. Return (%)" shows the annualized average returns in percentage, and the column " Sharp_ratio " shows the annualized Sharp-ratio values for each underlying model for a given country.

Table 3.3: Empirical results for advanced countries in horizon=1, continued

| Countries | Weighting Approaches | MSPE_ratio | DM (t-stats) | DM (p-value) | $\hat{p}$ | PT (t-stats) | PT (p-value) | Average Return (%) | Sharpe_ratio |
|---|---|---|---|---|---|---|---|---|---|
| Italy | Linear Combination (Ridge regression) | 1.0071 | -0.5424 | 0.2942 | 0.5267 | 0.6116 | 0.2704 | 0.1336 | 0.0169 |
| Italy | Median of Individual Models | 1.1401 | -2.0666 | 0.0204 | 0.4962 | -0.0874 | 0.5348 | -0.2492 | -0.0298 |
| Italy | Best Individual Model | 1.0930 | -1.4466 | 0.0752 | 0.5954 | 2.1843 | 0.0145 | 5.1972 | 0.6851 |
| Japan | Linear Combination (Ridge regression) | 1.0519 | -1.1314 | 0.1300 | 0.5344 | 0.7863 | 0.2158 | 0.4271 | 0.0462 |
| Japan | Median of Individual Models | 1.1739 | -2.1129 | 0.0183 | 0.5038 | 0.0874 | 0.4652 | -0.6563 | -0.0732 |
| Japan | Best Individual Model | 1.1823 | -1.8381 | 0.0342 | 0.5802 | 1.8348 | 0.0333 | 3.6699 | 0.5096 |
| Korea | Linear Combination (Ridge regression) | 1.0221 | -1.7046 | 0.0453 | 0.4351 | -1.4853 | 0.9313 | -2.5571 | -0.3557 |
| Korea | Median of Individual Models | 1.1730 | -2.5955 | 0.0053 | 0.5191 | 0.4369 | 0.3311 | -1.0285 | -0.1248 |
| Korea | Best Individual Model | 1.0664 | -1.4507 | 0.0746 | 0.5878 | 2.0095 | 0.0222 | 2.8176 | 0.3719 |
| Luxembourg | Linear Combination (Ridge regression) | 1.0036 | -0.2761 | 0.3915 | 0.5420 | 0.9611 | 0.1683 | 1.9937 | 0.2336 |
| Luxembourg | Median of Individual Models | 1.0983 | -1.8466 | 0.0335 | 0.4656 | -0.7863 | 0.7842 | -0.7515 | -0.0889 |
| Luxembourg | Best Individual Model | 1.0499 | -1.0801 | 0.1410 | 0.5344 | 0.7863 | 0.2158 | 2.1641 | 0.2543 |
| Netherlands | Linear Combination (Ridge regression) | 1.0051 | -0.3599 | 0.3598 | 0.5573 | 1.3106 | 0.0950 | 1.4787 | 0.1611 |
| Netherlands | Median of Individual Models | 1.1456 | -2.2974 | 0.0116 | 0.4809 | -0.4369 | 0.6689 | -1.0516 | -0.1372 |
| Netherlands | Best Individual Model | 1.0466 | -1.0398 | 0.1502 | 0.5878 | 2.0095 | 0.0222 | 5.4029 | 0.6431 |
| Norway | Linear Combination (Ridge regression) | 1.0222 | -1.2935 | 0.0991 | 0.5420 | 0.9611 | 0.1683 | 2.0039 | 0.2030 |
| Norway | Median of Individual Models | 1.1855 | -1.4741 | 0.0714 | 0.5231 | 0.5262 | 0.2994 | -1.0194 | -0.0836 |
| Norway | Best Individual Model | 1.7560 | -1.1453 | 0.1271 | 0.5725 | 1.6600 | 0.0485 | 2.4731 | 0.1885 |
| Portugal | Linear Combination (Ridge regression) | 1.0064 | -0.4840 | 0.3146 | 0.5692 | 1.5787 | 0.0572 | 3.4375 | 0.4425 |
| Portugal | Median of Individual Models | 1.1539 | -2.2433 | 0.0133 | 0.4806 | -0.4402 | 0.6701 | -0.4158 | -0.0514 |
| Portugal | Best Individual Model | 1.0870 | -1.3135 | 0.0957 | 0.5649 | 1.4853 | 0.0687 | 3.4088 | 0.4179 |
| Slovakia | Linear Combination (Ridge regression) | 1.0187 | -0.7695 | 0.2215 | 0.5649 | 1.4853 | 0.0687 | 1.8415 | 0.2074 |
| Slovakia | Median of Individual Models | 1.1469 | -1.9795 | 0.0249 | 0.4885 | -0.2621 | 0.6034 | 0.1022 | 0.0127 |
| Slovakia | Best Individual Model | 1.1619 | -1.7181 | 0.0441 | 0.5649 | 1.4853 | 0.0687 | 3.1261 | 0.3970 |
| Slovenia | Linear Combination (Ridge regression) | 1.0041 | -0.3088 | 0.3790 | 0.5496 | 1.1358 | 0.1280 | 1.8163 | 0.2165 |
| Slovenia | Median of Individual Models | 1.1037 | -1.8745 | 0.0316 | 0.5115 | 0.2621 | 0.3966 | -0.0159 | -0.0018 |
| Slovenia | Best Individual Model | 1.1369 | -1.9573 | 0.0262 | 0.5954 | 2.1843 | 0.0145 | 2.5465 | 0.2821 |
| Spain | Linear Combination (Ridge regression) | 1.0105 | -0.7837 | 0.2173 | 0.5496 | 1.1358 | 0.1280 | 2.3581 | 0.2881 |
| Spain | Median of Individual Models | 1.1171 | -2.1797 | 0.0155 | 0.4962 | -0.0874 | 0.5348 | -0.1690 | -0.0215 |
| Spain | Best Individual Model | 1.1823 | -2.5615 | 0.0058 | 0.5573 | 1.3106 | 0.0950 | 1.7313 | 0.1958 |
| Sweden | Linear Combination (Ridge regression) | 1.0228 | -1.2185 | 0.1126 | 0.4962 | -0.0874 | 0.5348 | -0.2973 | -0.0300 |
| Sweden | Median of Individual Models | 1.1431 | -2.3613 | 0.0098 | 0.4885 | -0.2621 | 0.6034 | -1.3398 | -0.1265 |
| Sweden | Best Individual Model | 1.0913 | -1.2012 | 0.1159 | 0.5649 | 1.4853 | 0.0687 | 3.5412 | 0.3078 |
| Switzerland | Linear Combination (Ridge regression) | 1.0034 | -0.1300 | 0.4484 | 0.5191 | 0.4369 | 0.3311 | 0.7266 | 0.0893 |
| Switzerland | Median of Individual Models | 1.1266 | -1.6681 | 0.0488 | 0.4733 | -0.6116 | 0.7296 | -0.8312 | -0.0996 |
| Switzerland | Best Individual Model | 1.0511 | -1.4187 | 0.0792 | 0.5420 | 0.9611 | 0.1683 | 1.5230 | 0.1865 |
| United Kingdom | Linear Combination (Ridge regression) | 1.0368 | -1.1451 | 0.1272 | 0.4667 | -0.7303 | 0.7674 | 0.8756 | 0.1078 |
| United Kingdom | Median of Individual Models | 1.1944 | -2.5690 | 0.0057 | 0.4750 | -0.5477 | 0.7081 | -2.1980 | -0.2646 |
| United Kingdom | Best Individual Model | 1.1863 | -2.5601 | 0.0059 | 0.5417 | 0.9129 | 0.1807 | 1.2704 | 0.1552 |

Note: The column "Countries" shows the names of countries. The column "Weighting Approaches"" defines the underlying models. The column "MSPE_ratio" indicates the $\text{MSPE}_{ratio}$ statistics for the Linear Combination (Ridge regression) weighting approach, the $\text{MSPE}_{ratio}$ statistics for the median of individual models, and the $\text{MSPE}_{ratio}$ statistics for the best individual model for a given country. The column "DM(t-stats)" shows the DM statistics for the Linear Combination (Ridge regression) weighting approach, the DM statistic for the median of individual models, and the DM statistics for the best individual model for a given country. The column "DM(p-value)" shows the p-values associated with the DM statistics. The column "$\hat{p}$" indicates the proportions of correct predictions of direction for the Linear Combination (Ridge regression) weighting approach, the proportions of correct predictions of direction for the median of individual models, and the proportions of correct predictions of direction for the best individual model. The column "PT(t-stats)" indicates the Pesaran-Timmermann (PT) statistics for the Linear Combination (Ridge regression) weighting approach, the PT statistics for the median of individual models, and the PT statistics for the best individual model and the column "PT(p-value)" shows the p-values associated with the PT statistics. All reported test statistics are one-sided. The significance level is 10%. The column "Avg. Return (%)" shows the annualized average returns in percentage, and the column " Sharp_ratio " shows the annualized Sharp-ratio values for each underlying model for a given country.

The empirical evidence in Tables 3.2 and 3.3 shows that the linear combination weighting approach is not successful in terms of the $\text{MSPE}_{raio}$ and DM statistics. Of course, this result is not surprising because even the best individual forecasting model does not outperform the WN model for a given country based on the $\text{MSPE}_{raio}$ and DM statistics. The results show that the investment strategies based on the linear combination weighing approach have positive average returns for some countries: Denmark, Ireland, Norway, Portugal, and Spain. For the remaining countries, either the values are positive but small or negative at horizon=1.

Table 3.4 reports out-of-sample statistics for the linear combination (Ridge regression) weighting approach, the median of individual models, and the best model for a given emerging/developing country at horizon=1. The results show the linear combination (Ridge regression) approach for the 73% of countries provides the proportions of correct direction of change greater than 0.5 ($\hat{p} > 0.5$), and for the 54% of results are statistically significant at the 10% significance level. For most countries, the linear combination (Ridge regression) method provides better forecasting results than the median of individual models.

The findings show the linear combination (Ridge regression) approach outperforms the WN model in terms of the $\text{MSPE}_{raio}$ for the 36% of countries, including India, Mexico, Tunisia, and Turkey.

The results show that investment strategies based on the linear combination (Ridge regression) weighting approach have positive average returns for most countries such as Brazil, Colombia, Mexico, Tunisia, and Turkey. The Sharpe-ratio values are noticeable for these countries. For the rest of the countries, values are not significant.

Table 3.4: Empirical results for emerging/developing countries in horizon=1

| Countries | Weighting Approaches | MSPE_ratio | DM (t-stats) | DM (p-value) | $\hat{p}$ | PT (t-stats) | PT (p-value) | Average Return (%) | Sharpe_ratio |
|---|---|---|---|---|---|---|---|---|---|
| Brazil | Linear Combination (Ridge regression) | 1.0218 | -0.5783 | 0.2821 | 0.5581 | 1.3207 | 0.0933 | 7.2794 | 0.4641 |
| Brazil | Median of Individual Models | 1.1412 | -1.4517 | 0.0745 | 0.4885 | -0.2621 | 0.6034 | 3.8666 | 0.2373 |
| Brazil | Best Models | 1.0068 | -0.1619 | 0.4358 | 0.5649 | 1.4853 | 0.0687 | 8.1263 | 0.6018 |
| Chile | Linear Combination (Ridge regression) | 1.0141 | -0.8614 | 0.1953 | 0.5649 | 1.4853 | 0.0687 | 3.1279 | 0.3317 |
| Chile | Median of Individual Models | 1.0369 | -0.9569 | 0.1702 | 0.5076 | 0.1747 | 0.4309 | 1.9565 | 0.1933 |
| Chile | Best Models | 1.0137 | -0.2125 | 0.4160 | 0.5573 | 1.3106 | 0.0950 | 5.5193 | 0.5060 |
| Colombia | Linear Combination (Ridge regression) | 1.0090 | -0.4787 | 0.3165 | 0.5802 | 1.8348 | 0.0333 | 5.4879 | 0.4947 |
| Colombia | Median of Individual Models | 1.0918 | -1.6562 | 0.0500 | 0.5191 | 0.4369 | 0.3311 | 2.3983 | 0.1886 |
| Colombia | Best Models | 1.0548 | -1.4832 | 0.0702 | 0.5725 | 1.6600 | 0.0485 | 6.3183 | 0.4998 |
| Hungary | Linear Combination (Ridge regression) | 1.0057 | -0.3364 | 0.3685 | 0.4427 | -1.3106 | 0.9050 | -3.4823 | -0.2423 |
| Hungary | Median of Individual Models | 1.1409 | -2.4787 | 0.0072 | 0.4580 | -0.9611 | 0.8317 | -1.9705 | -0.1731 |
| Hungary | Best Models | 1.2708 | -2.9348 | 0.0020 | 0.5344 | 0.7863 | 0.2158 | 0.9551 | 0.0886 |
| India | Linear Combination (Ridge regression) | 0.9958 | 0.1773 | 0.4298 | 0.5159 | 0.3563 | 0.3608 | 1.4445 | 0.1751 |
| India | Median of Individual Models | 1.2063 | -1.1046 | 0.1347 | 0.5433 | 0.9761 | 0.1645 | 1.6990 | 0.2116 |
| India | Best Models | 0.9522 | 1.1076 | 0.1351 | 0.6000 | 2.2361 | 0.0127 | 5.9028 | 0.7761 |
| Mexico | Linear Combination (Ridge regression) | 0.9956 | 0.2208 | 0.4128 | 0.4959 | -0.0902 | 0.5359 | 5.4825 | 0.4708 |
| Mexico | Median of Individual Models | 1.0593 | -1.4977 | 0.0684 | 0.5000 | 0.0000 | 0.5000 | 0.6869 | 0.0569 |
| Mexico | Best Models | 1.0342 | -1.1388 | 0.1285 | 0.5772 | 1.7132 | 0.0433 | 4.2871 | 0.3985 |
| Peru | Linear Combination (Ridge regression) | 1.0130 | -0.4101 | 0.3412 | 0.5984 | 2.1729 | 0.0149 | 1.5886 | 0.3454 |
| Peru | Median of Individual Models | 1.0399 | -0.8478 | 0.1992 | 0.5492 | 1.0864 | 0.1386 | 0.5006 | 0.1122 |
| Peru | Best Models | 1.0169 | -0.4054 | 0.3429 | 0.6475 | 3.2593 | 0.0006 | 2.0302 | 0.6269 |
| Poland | Linear Combination (Ridge regression) | 1.0114 | -0.3594 | 0.3600 | 0.5496 | 1.1358 | 0.1280 | 4.4002 | 0.3627 |
| Poland | Median of Individual Models | 1.1799 | -1.4040 | 0.0814 | 0.4962 | -0.0874 | 0.5348 | 0.0371 | 0.0027 |
| Poland | Best Models | 0.9964 | 0.1518 | 0.4398 | 0.6260 | 2.8832 | 0.0020 | 7.1586 | 0.5106 |
| South Africa | Linear Combination (Ridge regression) | 1.0184 | -0.9067 | 0.1831 | 0.4351 | -1.4853 | 0.9313 | 2.0614 | 0.1363 |
| South Africa | Median of Individual Models | 1.1070 | -1.6749 | 0.0482 | 0.4885 | -0.2621 | 0.6034 | -0.0216 | -0.0012 |
| South Africa | Best Models | 1.0603 | -1.5532 | 0.0614 | 0.5649 | 1.4853 | 0.0687 | 7.3226 | 0.5723 |
| Tunisia | Linear Combination (Ridge regression) | 0.9266 | 1.2974 | 0.0985 | 0.5984 | 2.1729 | 0.0149 | 7.4719 | 0.9563 |
| Tunisia | Median of Individual Models | 0.9651 | 0.5419 | 0.2944 | 0.5656 | 1.4486 | 0.0770 | 5.6596 | 0.7131 |
| Tunisia | Best Models | 0.9486 | 0.8381 | 0.2018 | 0.6148 | 2.5350 | 0.0056 | 6.1423 | 0.7131 |
| Turkey | Linear Combination (Ridge regression) | 0.9800 | 0.3823 | 0.3514 | 0.6000 | 2.2804 | 0.0113 | 14.2366 | 0.9721 |
| Turkey | Median of Individual Models | 1.4586 | -1.0193 | 0.1523 | 0.5692 | 1.5787 | 0.0572 | 8.8254 | 0.5954 |
| Turkey | Best Models | 0.9945 | 0.1199 | 0.4524 | 0.6308 | 2.9820 | 0.0014 | 14.2864 | 0.9696 |

Note: The column "Countries" shows the names of countries. The column "Weighting Approaches" defines the underlying models. The column "MSPE_ratio" indicates the MSPE$_{ratio}$ statistics for the Linear Combination (Ridge regression) weighting approach, the MSPE$_{ratio}$ statistics for the median of individual models, and the MSPE$_{ratio}$ statistics for the best individual model for a given country. The column "DM(t-stats)" shows the DM statistics for the Linear Combination (Ridge regression) weighting approach, the DM statistic for the median of individual models, and the DM statistics for the best individual model for a given country. The column "DM(p-value)" shows the p-values associated with the DM statistics. The column "$\hat{p}$" indicates the proportions of correct predictions of direction for the Linear Combination (Ridge regression) weighting approach, the proportions of correct predictions of direction for the median of individual models, and the proportions of correct predictions of direction for the best individual model. The column "PT(t-stats)" indicates the Pesaran-Timmermann (PT) statistics for the Linear Combination (Ridge regression) weighting approach, the PT statistics for the median of individual models, and the PT statistics for the best individual model and the column "PT(p-value)" shows the p-values associated with the PT statistics. All reported test statistics are one-sided. The significance level is 10%. The column "Avg. Return (%)" shows the annualized average returns in percentage, and the column " Sharp_ratio " shows the annualized Sharp-ratio values for each underlying model for a given country.

Table 3.5 reports the median of out-of-sample statistics for a given weighting methodology, the median of individual models, and the median of best models in each group at horizon=12. The findings indicate, on average, none of weighting approaches outperforms the WN model in terms of the MSPE$_{ratio}$ and DM statistics. The results show, on average, using weighting approaches does not provide promising results for advanced countries.

However, on average, all weighting methods have $\hat{p} > 0.5$, so they successfully predict the direction of changes in exchange rates for emerging/developing countries. Still, the weighting approaches, on average, do not provide better results than the WN model in terms of the MSPE$_{ratio}$ and DM statistics.

Table 3.5: Summary results in horizon=12

| WEO country group | Weighting Approaches | MSPE_ratio | DM (t-stats) | DM (p-value) | $\hat{p}$ | PT (t-stats) | PT (p-value) | Average Return (%) | Sharpe_ratio |
|---|---|---|---|---|---|---|---|---|---|
| Advanced | Linear Combination (Ridge regression) | 1.1609 | -2.0707 | 0.0206 | 0.4643 | -0.7071 | 0.7591 | -1.3009 | -0.1079 |
| | Convex Combination (Ridge regression) | 1.8241 | -2.6119 | 0.0053 | 0.4278 | -1.4215 | 0.9224 | -0.2805 | -0.0186 |
| | MSPE | 1.4357 | -2.3605 | 0.0101 | 0.4031 | -1.9193 | 0.9719 | -1.1091 | -0.0843 |
| | Directional Prediction | 1.4145 | -2.3293 | 0.0110 | 0.3929 | -2.1213 | 0.9826 | -1.2716 | -0.0933 |
| | Equal Weights | 1.3556 | -2.3598 | 0.0102 | 0.3918 | -2.1318 | 0.9831 | -1.1032 | -0.0910 |
| | Median of Individual Models | 1.4968 | -2.2548 | 0.0133 | 0.4898 | -0.2020 | 0.5801 | -0.1147 | -0.0089 |
| | Best Models | 1.5484 | -1.7321 | 0.0432 | 0.6122 | 2.2223 | 0.0131 | 1.6248 | 0.1102 |
| Emerging/Developing | Directional Prediction | 1.2199 | -0.9785 | 0.0518 | 0.6222 | 2.3190 | 0.0102 | 3.2481 | 0.1955 |
| | MSPE | 1.2242 | -1.0141 | 0.0520 | 0.6122 | 2.0190 | 0.0107 | 3.1471 | 0.1888 |
| | Equal Weights | 1.2688 | -1.0834 | 0.0640 | 0.6020 | 2.0203 | 0.0217 | 3.3433 | 0.1617 |
| | Linear Combination (Ridge regression) | 1.0573 | -0.4786 | 0.1556 | 0.5918 | 1.8183 | 0.0345 | 2.0302 | 0.1333 |
| | Convex Combination (Ridge regression) | 1.6021 | -1.2848 | 0.0168 | 0.5778 | 1.4757 | 0.0700 | 2.0955 | 0.1576 |
| | Median of Individual Models | 1.1510 | -1.0479 | 0.0602 | 0.5816 | 1.6162 | 0.0530 | 2.1007 | 0.1306 |
| | Best Models | 1.0436 | -0.3169 | 0.2247 | 0.6939 | 3.8386 | 0.0001 | 5.3710 | 0.2783 |

Note: The column "WEO Country Group" shows the World Economic Outlook (WEO) countries' classifications. The column "Weighting Approaches" defines the underlying model or weighting approach. The column "MSPE_ratio" indicates the median of the MSPE$_{ratio}$ for each weighting approach, the median of individual models, the best models within each country group. The column "DM(t-stats)" shows the median of DM statistics for each weighting approach, the median of individual models, the best models within each country group. The column "DM(p-value)" shows the median of p-values associated with the DM statistics. The column "$\hat{p}$" indicates the median of proportions of correct predictions of direction for each weighting approach, the median of individual models, the best models within each country group. The column "PT(t-stats)" shows the median of Pesaran-Timmermann (PT) statistics for each weighting approach, the median of individual models, the best models within each country group and the column "PT(p-value)" shows the median of p-values associated with the PT statistics. All reported test statistics are one-sided. The significance level is 10%. The column "Avg. Return (%)" shows the median of annualized average returns in percentage, and the column "Sharp_ratio" shows the median of annualized Sharp-ratio values. To be clear, the median of individual models means I first calculate the median of underlying statistics using all statistics from all individual models for a given country, then calculate the median of the median of underlying statistics across countries in each country group.

Based on the results in Table 3.5, the linear combination (Ridge regression) weighting approach has the highest value of $\hat{p}$ among other weighting approaches for advanced countries. In addition, the Directional Prediction weighting approach is a successful weighing approach among different weighting approaches in terms of $\hat{p}$ for emerging/developing countries. Therefore, I only present the detailed results for these weighting approaches at horizon=12.

Tables 3.6 and 3.7 report the out-of-sample statistics for the linear combination (Ridge regression) weighting approach, the median of individual models, and the best model for a given advanced country at horizon=12. The results show that the linear combination (Ridge regression) method for the 35% of countries provides the proportions of correct direction of change greater than 0.5 ($\hat{p} > 0.5$).

Table 3.6: Empirical results for advanced countries in horizon=12

| Countries | Models | MSPE_ratio | DM (t-stats) | DM (p-value) | $\hat{p}$ | PT (t-stats) | PT (p-value) | Average Return (%) | Sharpe_ratio |
|---|---|---|---|---|---|---|---|---|---|
| Austria | Linear Combination (Ridge regression) | 1.1518 | -1.3750 | 0.0861 | 0.5510 | 1.0102 | 0.1562 | 1.5243 | 0.0979 |
| Austria | Median of Individual Models | 1.3648 | -2.1650 | 0.0164 | 0.5000 | 0.0000 | 0.5000 | 0.1544 | 0.0131 |
| Austria | Best Model | 2.0116 | -1.9822 | 0.0251 | 0.6122 | 2.2223 | 0.0131 | 1.2371 | 0.0965 |
| Belgium | Linear Combination (Ridge regression) | 1.1645 | -2.0954 | 0.0194 | 0.4388 | -1.2122 | 0.8873 | 0.1090 | 0.0073 |
| Belgium | Median of Individual Models | 1.4207 | -2.1409 | 0.0174 | 0.5102 | 0.2020 | 0.4199 | 0.1997 | 0.0161 |
| Belgium | Best Model | 1.1063 | -0.3715 | 0.3555 | 0.6122 | 2.2223 | 0.0131 | 1.7959 | 0.1285 |
| Canada | Linear Combination (Ridge regression) | 1.1216 | -1.7254 | 0.0440 | 0.4944 | -0.1060 | 0.5422 | -1.3916 | -0.1153 |
| Canada | Median of Individual Models | 1.7566 | -3.3408 | 0.0007 | 0.4607 | -0.7420 | 0.7710 | -1.4982 | -0.1368 |
| Canada | Best Model | 1.5312 | -1.9625 | 0.0264 | 0.5618 | 1.1660 | 0.1218 | 1.1638 | 0.1072 |
| Czech Republic | Linear Combination (Ridge regression) | 1.2015 | -3.4195 | 0.0005 | 0.4082 | -1.8183 | 0.9655 | -3.8237 | -0.2829 |
| Czech Republic | Median of Individual Models | 1.5631 | -2.3907 | 0.0094 | 0.4796 | -0.4041 | 0.6569 | -0.0738 | -0.0052 |
| Czech Republic | Best Model | 1.2672 | -1.4450 | 0.0758 | 0.6224 | 2.4244 | 0.0077 | 1.4940 | 0.0921 |
| Denmark | Linear Combination (Ridge regression) | 1.0834 | -1.3005 | 0.0983 | 0.5000 | 0.0000 | 0.5000 | -1.3939 | -0.1004 |
| Denmark | Median of Individual Models | 1.5346 | -2.2962 | 0.0119 | 0.4592 | -0.8081 | 0.7905 | -0.8449 | -0.0663 |
| Denmark | Best Model | 1.6366 | -1.5146 | 0.0666 | 0.5714 | 1.4142 | 0.0786 | 1.7340 | 0.1117 |
| Estonia | Linear Combination (Ridge regression) | 1.1554 | -2.3447 | 0.0105 | 0.4694 | -0.6061 | 0.7278 | -2.3734 | -0.1801 |
| Estonia | Median of Individual Models | 1.3704 | -2.2134 | 0.0146 | 0.4694 | -0.6061 | 0.7278 | -0.1555 | -0.0126 |
| Estonia | Best Model | 1.3802 | -2.1962 | 0.0152 | 0.5918 | 1.8183 | 0.0345 | 1.6566 | 0.1073 |
| Finland | Linear Combination (Ridge regression) | 1.1615 | -1.6765 | 0.0484 | 0.5612 | 1.2122 | 0.1127 | 1.0877 | 0.0738 |
| Finland | Median of Individual Models | 1.5065 | -1.9887 | 0.0248 | 0.5204 | 0.4041 | 0.3431 | 0.2378 | 0.0197 |
| Finland | Best Model | 1.1807 | -0.7850 | 0.2172 | 0.7245 | 4.4447 | 0.0000 | 3.6828 | 0.3034 |
| France | Linear Combination (Ridge regression) | 1.1699 | -2.4553 | 0.0079 | 0.4184 | -1.6162 | 0.9470 | -1.7689 | -0.1302 |
| France | Median of Individual Models | 1.6919 | -2.6465 | 0.0047 | 0.4490 | -1.0102 | 0.8438 | -0.4356 | -0.0345 |
| France | Best Model | 1.8296 | -2.2729 | 0.0126 | 0.6224 | 2.4244 | 0.0077 | 1.5604 | 0.1168 |
| Germany | Linear Combination (Ridge regression) | 1.2168 | -2.9480 | 0.0020 | 0.3571 | -2.8284 | 0.9977 | -2.4590 | -0.1838 |
| Germany | Median of Individual Models | 1.4830 | -2.1169 | 0.0184 | 0.5102 | 0.2020 | 0.4199 | 0.4307 | 0.0377 |
| Germany | Best Model | 1.6025 | -1.9222 | 0.0288 | 0.6327 | 2.6264 | 0.0043 | 1.1595 | 0.0873 |
| Greece | Linear Combination (Ridge regression) | 1.1362 | -2.3900 | 0.0094 | 0.4286 | -1.4142 | 0.9214 | -1.0347 | -0.0923 |
| Greece | Median of Individual Models | 1.4787 | -2.2976 | 0.0119 | 0.4898 | -0.2020 | 0.5801 | -0.4255 | -0.0325 |
| Greece | Best Model | 1.8694 | -2.1335 | 0.0177 | 0.5714 | 1.4142 | 0.0786 | 2.0324 | 0.1430 |
| Iceland | Linear Combination (Ridge regression) | 1.1612 | -2.5454 | 0.0062 | 0.2449 | -5.0508 | 1.0000 | -5.1158 | -0.2949 |
| Iceland | Median of Individual Models | 1.6597 | -3.0960 | 0.0013 | 0.3367 | -3.2325 | 0.9994 | -3.0750 | -0.1809 |
| Iceland | Best Model | 2.0343 | -3.1837 | 0.0010 | 0.4898 | -0.2020 | 0.5801 | 0.1041 | 0.0056 |
| Ireland | Linear Combination (Ridge regression) | 1.1177 | -2.1908 | 0.0154 | 0.5000 | 0.0000 | 0.5000 | -0.2715 | -0.0189 |
| Ireland | Median of Individual Models | 1.4933 | -2.4327 | 0.0084 | 0.5000 | 0.0000 | 0.5000 | -0.7458 | -0.0600 |
| Ireland | Best Model | 1.8758 | -1.7603 | 0.0408 | 0.5918 | 1.8183 | 0.0345 | 1.5929 | 0.1087 |
| Israel | Linear Combination (Ridge regression) | 1.1535 | -1.5470 | 0.0626 | 0.3750 | -2.4495 | 0.9928 | -1.6369 | -0.1489 |
| Israel | Median of Individual Models | 1.3612 | -2.1928 | 0.0154 | 0.4330 | -1.3195 | 0.9055 | -0.7845 | -0.0888 |
| Israel | Best Model | 1.2027 | -1.5237 | 0.0654 | 0.5816 | 1.6162 | 0.0530 | 0.7664 | 0.0891 |

Note: The column "Countries" shows the names of countries. The column "Weighting Approaches" defines the underlying models. The column "MSPE_ratio" indicates the MSPE$_{ratio}$ statistics for the Linear Combination (Ridge regression) weighting approach, the MSPE$_{ratio}$ statistics for the median of individual models, and the MSPE$_{ratio}$ statistics for the best individual model for a given country. The column "DM(t-stats)" shows the DM statistics for the Linear Combination (Ridge regression) weighting approach, the DM statistic for the median of individual models, and the DM statistics for the best individual model for a given country. The column "DM(p-value)" shows the p-values associated with the DM statistics. The column "$\hat{p}$" indicates the proportions of correct predictions of direction for the Linear Combination (Ridge regression) weighting approach, the proportions of correct predictions of direction for the median of individual models, and the proportions of correct predictions of direction for the best individual model. The column "PT(t-stats)" indicates the Pesaran-Timmermann (PT) statistics for the Linear Combination (Ridge regression) weighting approach, the PT statistics for the median of individual models, and the PT statistics for the best individual model and the column "PT(p-value)" shows the p-values associated with the PT statistics. All reported test statistics are one-sided. The significance level is 10%. The column "Avg. Return (%)" shows the annualized average returns in percentage, and the column " Sharp_ratio " shows the annualized Sharp-ratio values for each underlying model for a given country.

Table 3.7: Empirical results for advanced countries in horizon=12, continued

| Countries | Weighting Approaches | MSPE_ratio | DM (t-stats) | DM (p-value) | $\hat{p}$ | PT (t-stats) | PT (p-value) | Average Return (%) | Sharpe_ratio |
|---|---|---|---|---|---|---|---|---|---|
| Italy | Linear Combination (Ridge regression) | 1.3555 | -2.6181 | 0.0051 | 0.3776 | -2.4244 | 0.9923 | -1.2103 | -0.0837 |
| Italy | Median of Individual Models | 1.6674 | -2.3745 | 0.0098 | 0.4898 | -0.2020 | 0.5801 | -0.0538 | -0.0045 |
| Italy | Best Model | 1.7159 | -1.4107 | 0.0808 | 0.6429 | 2.8284 | 0.0023 | 2.3739 | 0.1621 |
| Japan | Linear Combination (Ridge regression) | 1.4999 | -3.5194 | 0.0003 | 0.3061 | -3.8386 | 0.9999 | -4.1112 | -0.2214 |
| Japan | Median of Individual Models | 2.1746 | -3.8080 | 0.0001 | 0.3776 | -2.4244 | 0.9923 | -2.9978 | -0.1597 |
| Japan | Best Model | 2.0159 | -4.1534 | 0.0000 | 0.4694 | -0.6061 | 0.7278 | -2.0623 | -0.1051 |
| Korea | Linear Combination (Ridge regression) | 1.3394 | -1.9261 | 0.0285 | 0.3776 | -2.4244 | 0.9923 | -1.5083 | -0.1661 |
| Korea | Median of Individual Models | 2.1708 | -3.6378 | 0.0002 | 0.3061 | -3.8386 | 0.9999 | -2.3693 | -0.2976 |
| Korea | Best Model | 2.1402 | -3.2062 | 0.0009 | 0.4694 | -0.6061 | 0.7278 | -0.4234 | -0.0463 |
| Luxembourg | Linear Combination (Ridge regression) | 1.1238 | -1.5890 | 0.0577 | 0.5408 | 0.8081 | 0.2095 | 0.5626 | 0.0404 |
| Luxembourg | Median of Individual Models | 1.2736 | -2.0781 | 0.0202 | 0.5204 | 0.4041 | 0.3431 | 0.4077 | 0.0296 |
| Luxembourg | Best Model | 1.5656 | -1.2997 | 0.0984 | 0.6122 | 2.2223 | 0.0131 | 2.1069 | 0.1384 |
| Netherlands | Linear Combination (Ridge regression) | 1.1606 | -1.3788 | 0.0856 | 0.5306 | 0.6061 | 0.2722 | -0.5214 | -0.0389 |
| Netherlands | Median of Individual Models | 1.9640 | -2.5217 | 0.0067 | 0.4898 | -0.2020 | 0.5801 | -0.6884 | -0.0526 |
| Netherlands | Best Model | 1.6782 | -1.7150 | 0.0448 | 0.6327 | 2.6264 | 0.0043 | 2.2868 | 0.1528 |
| Norway | Linear Combination (Ridge regression) | 1.0138 | -0.1816 | 0.4281 | 0.6327 | 2.6264 | 0.0043 | 0.9391 | 0.0516 |
| Norway | Median of Individual Models | 1.2632 | -1.9932 | 0.0245 | 0.5459 | 0.9091 | 0.1829 | 1.0248 | 0.0583 |
| Norway | Best Model | 1.0628 | -0.5312 | 0.2982 | 0.6429 | 2.8284 | 0.0023 | 3.9593 | 0.2667 |
| Portugal | Linear Combination (Ridge regression) | 1.1939 | -3.2271 | 0.0009 | 0.3673 | -2.6264 | 0.9957 | -2.7251 | -0.2486 |
| Portugal | Median of Individual Models | 1.5003 | -2.1805 | 0.0158 | 0.5102 | 0.2020 | 0.4199 | 0.5859 | 0.0456 |
| Portugal | Best Model | 1.4941 | -1.7492 | 0.0417 | 0.6429 | 2.8284 | 0.0023 | 2.2756 | 0.1496 |
| Slovakia | Linear Combination (Ridge regression) | 1.0850 | -1.4275 | 0.0783 | 0.5714 | 1.4142 | 0.0786 | 0.6007 | 0.0438 |
| Slovakia | Median of Individual Models | 1.2876 | -2.1724 | 0.0161 | 0.5306 | 0.6061 | 0.2722 | -0.0426 | -0.0033 |
| Slovakia | Best Model | 1.2087 | -1.9885 | 0.0248 | 0.6122 | 2.2223 | 0.0131 | 0.5300 | 0.0390 |
| Slovenia | Linear Combination (Ridge regression) | 1.1754 | -2.4497 | 0.0080 | 0.4592 | -0.8081 | 0.7905 | -2.6429 | -0.1979 |
| Slovenia | Median of Individual Models | 1.3386 | -1.8201 | 0.0359 | 0.5102 | 0.2020 | 0.4199 | 0.1257 | 0.0089 |
| Slovenia | Best Model | 1.4967 | -1.4077 | 0.0812 | 0.6735 | 3.4345 | 0.0003 | 2.8626 | 0.2012 |
| Spain | Linear Combination (Ridge regression) | 1.3285 | -4.2438 | 0.0000 | 0.4184 | -1.6162 | 0.9470 | -1.8884 | -0.1572 |
| Spain | Median of Individual Models | 1.3484 | -2.0165 | 0.0233 | 0.5204 | 0.4041 | 0.3431 | 0.1185 | 0.0092 |
| Spain | Best Model | 1.2955 | -0.9247 | 0.1787 | 0.6531 | 3.0305 | 0.0012 | 2.4260 | 0.2008 |
| Sweden | Linear Combination (Ridge regression) | 1.1292 | -1.2056 | 0.1155 | 0.5306 | 0.6061 | 0.2722 | -1.1917 | -0.0665 |
| Sweden | Median of Individual Models | 1.6900 | -3.0448 | 0.0015 | 0.4337 | -1.3132 | 0.9043 | -2.0198 | -0.1267 |
| Sweden | Best Model | 1.2240 | -1.5798 | 0.0587 | 0.6020 | 2.0203 | 0.0217 | -0.4539 | -0.0287 |
| Switzerland | Linear Combination (Ridge regression) | 1.2737 | -2.0459 | 0.0217 | 0.5102 | 0.2020 | 0.4199 | -0.8505 | -0.1156 |
| Switzerland | Median of Individual Models | 2.3059 | -3.1863 | 0.0010 | 0.4592 | -0.8081 | 0.7905 | -0.9616 | -0.1408 |
| Switzerland | Best Model | 2.3407 | -3.1437 | 0.0011 | 0.5612 | 1.2122 | 0.1127 | -0.0535 | -0.0079 |
| United Kingdom | Linear Combination (Ridge regression) | 1.0069 | -0.1192 | 0.4527 | 0.5747 | 1.3937 | 0.0817 | 1.2584 | 0.0843 |
| United Kingdom | Median of Individual Models | 1.3440 | -2.1699 | 0.0164 | 0.5057 | 0.1072 | 0.4573 | 0.4373 | 0.0333 |
| United Kingdom | Best Model | 1.3545 | -1.1816 | 0.1203 | 0.6437 | 2.6803 | 0.0037 | 1.8879 | 0.1419 |

Note: The column "Countries" shows the names of countries. The column "Weighting Approaches" defines the underlying models. The column "MSPE_ratio" indicates the $\text{MSPE}_{ratio}$ statistics for the Linear Combination (Ridge regression) weighting approach, the $\text{MSPE}_{ratio}$ statistics for the median of individual models, and the $\text{MSPE}_{ratio}$ statistics for the best individual model for a given country. The column "DM(t-stats)" shows the DM statistics for the Linear Combination (Ridge regression) weighting approach, the DM statistic for the median of individual models, and the DM statistics for the best individual model for a given country. The column "DM(p-value)" shows the p-values associated with the DM statistics. The column "$\hat{p}$" indicates the proportions of correct predictions of direction for the Linear Combination (Ridge regression) weighting approach, the proportions of correct predictions of direction for the median of individual models, and the proportions of correct predictions of direction for the best individual model. The column "PT(t-stats)" indicates the Pesaran-Timmermann (PT) statistics for the Linear Combination (Ridge regression) weighting approach, the PT statistics for the median of individual models, and the PT statistics for the best individual model and the column "PT(p-value)" shows the p-values associated with the PT statistics. All reported test statistics are one-sided. The significance level is 10%. The column "Avg. Return (%)" shows the annualized average returns in percentage, and the column " Sharp_ratio " shows the annualized Sharp-ratio values for each underlying model for a given country.

The findings in Tables 3.6 and 3.7 show that the linear combination (Ridge regression) weighting approach is not successful in terms of the $\text{MSPE}_{raio}$ and DM statistics. Moreover, investment strategies using the linear weighting approach will not be profitable given the empirical evidence for advanced countries at horizon=12.

Table 3.8 reports out-of-sample statistics for the Directional Prediction weighting approach, the median of individual models, and the best model for a given emerging/developing country at horizon=12. The results show the Directional Prediction method for 64% of countries provides the proportion of correct direction of changes greater than 0.5 ($\hat{p} > 0.5$). For some countries such as Brazil, Mexico, and Tunisia, the Directional Prediction method provides better forecasting results than the median of individual models. The findings show that the weighting approach is not successful in terms of the $\text{MSPE}_{raio}$ and DM statistics for all countries, except for Mexico, South Africa, Tunisia, and Turkey. Investment strategies using the Directional Prediction weighting approaches have noteworthy average returns for some countries: Brazil, India, South Africa, and Tunisia. The values of Sharpe-ratio are notable for these countries. For the remaining countries, values are not considerable.

Table 3.8: Empirical results for emerging/developing countries in horizon=12

| Countries | Weighting Approaches | MSPE_ratio | DM (t-stats) | DM (p-value) | $\hat{p}$ | PT (t-stats) | PT (p-value) | Average Return (%) | Sharpe_ratio |
|---|---|---|---|---|---|---|---|---|---|
| Brazil | Directional Prediction | 1.3041 | -0.9785 | 0.1651 | 0.6429 | 2.8284 | 0.0023 | 5.9934 | 0.1955 |
| Brazil | Median of Individual Models | 1.3342 | -1.2817 | 0.1015 | 0.5816 | 1.6162 | 0.0530 | 3.4858 | 0.1097 |
| Brazil | Best Model | 0.9952 | 0.0312 | 0.4876 | 0.7041 | 4.0406 | 0.0000 | 7.5187 | 0.2756 |
| Chile | Directional Prediction | 1.2199 | -1.7108 | 0.0452 | 0.3980 | -2.0203 | 0.9783 | -1.6486 | -0.0879 |
| Chile | Median of Individual Models | 1.1465 | -1.0479 | 0.1486 | 0.4949 | -0.1010 | 0.5400 | 0.6177 | 0.0363 |
| Chile | Best Model | 1.0436 | -0.3169 | 0.3760 | 0.5510 | 1.0102 | 0.1562 | 2.1838 | 0.1355 |
| Colombia | Directional Prediction | 1.3377 | -1.6428 | 0.0518 | 0.5000 | 0.0000 | 0.5000 | -0.4104 | -0.0159 |
| Colombia | Median of Individual Models | 1.2281 | -1.8393 | 0.0345 | 0.5255 | 0.5051 | 0.3077 | -0.3003 | -0.0117 |
| Colombia | Best Model | 1.4672 | -1.8285 | 0.0353 | 0.6327 | 2.6264 | 0.0043 | 5.5360 | 0.2167 |
| Hungary | Directional Prediction | 1.2763 | -1.1768 | 0.1211 | 0.5612 | 1.2122 | 0.1127 | 1.7996 | 0.1105 |
| Hungary | Median of Individual Models | 1.5306 | -1.5671 | 0.0602 | 0.5918 | 1.8183 | 0.0345 | 2.1007 | 0.1306 |
| Hungary | Best Model | 1.1292 | -0.7122 | 0.2390 | 0.6939 | 3.8386 | 0.0001 | 4.1681 | 0.2783 |
| India | Directional Prediction | 1.0856 | -0.2583 | 0.3984 | 0.7021 | 3.9194 | 0.0000 | 3.2481 | 0.3029 |
| India | Median of Individual Models | 1.1146 | -0.3525 | 0.3629 | 0.7128 | 4.1257 | 0.0000 | 3.3628 | 0.3341 |
| India | Best Model | 0.9581 | 0.1537 | 0.4391 | 0.7447 | 4.7445 | 0.0000 | 3.8453 | 0.3809 |
| Mexico | Directional Prediction | 0.9706 | 0.1742 | 0.4311 | 0.6222 | 2.3190 | 0.0102 | 3.8016 | 0.2511 |
| Mexico | Median of Individual Models | 1.1510 | -0.8540 | 0.1977 | 0.5556 | 1.0541 | 0.1459 | 2.0809 | 0.1351 |
| Mexico | Best Model | 1.1911 | -0.6432 | 0.2609 | 0.6556 | 2.9515 | 0.0016 | 5.3710 | 0.3282 |
| Peru | Directional Prediction | 1.5289 | -3.6525 | 0.0002 | 0.4157 | -1.5900 | 0.9441 | -1.9352 | -0.2043 |
| Peru | Median of Individual Models | 1.5161 | -3.6818 | 0.0002 | 0.4270 | -1.3780 | 0.9159 | -1.9253 | -0.2047 |
| Peru | Best Model | 1.5251 | -3.1966 | 0.0010 | 0.4607 | -0.7420 | 0.7710 | -1.5921 | -0.1527 |
| Poland | Directional Prediction | 1.4561 | -2.0833 | 0.0199 | 0.4388 | -1.2122 | 0.8873 | -0.9919 | -0.0684 |
| Poland | Median of Individual Models | 1.4852 | -1.7084 | 0.0454 | 0.5102 | 0.2020 | 0.4199 | 0.1663 | 0.0110 |
| Poland | Best Model | 1.2411 | -1.3013 | 0.0981 | 0.6531 | 3.0305 | 0.0012 | 2.1500 | 0.1519 |
| South Africa | Directional Prediction | 0.8952 | 0.4944 | 0.3111 | 0.7755 | 5.4548 | 0.0000 | 7.5782 | 0.3486 |
| South Africa | Median of Individual Models | 1.0526 | -0.2249 | 0.3724 | 0.7398 | 4.7477 | 0.0000 | 6.5049 | 0.3031 |
| South Africa | Best Model | 0.8345 | 0.7596 | 0.2247 | 0.8571 | 7.0711 | 0.0000 | 9.6980 | 0.5738 |
| Tunisia | Directional Prediction | 0.8815 | 1.0260 | 0.0016 | 0.7876 | 7.3140 | 0.0000 | 8.3556 | 0.6612 |
| Tunisia | Median of Individual Models | 0.8878 | 1.8402 | 0.0028 | 0.6764 | 7.1020 | 0.0000 | 8.2418 | 0.6424 |
| Tunisia | Best Model | 0.8981 | 1.7527 | 0.0002 | 0.8089 | 7.5260 | 0.0000 | 8.3616 | 0.6624 |
| Turkey | Directional Prediction | 0.9115 | 1.5824 | 0.0003 | 0.8278 | 6.4274 | 0.0000 | 16.6577 | 0.7755 |
| Turkey | Median of Individual Models | 0.9357 | 1.9962 | 0.0017 | 0.8969 | 7.8182 | 0.0000 | 15.5825 | 0.6853 |
| Turkey | Best Model | 0.8795 | 2.4309 | 0.0000 | 0.8278 | 6.4274 | 0.0000 | 16.6577 | 0.7755 |

Note: The column "Countries" shows the names of countries. The column "Weighting Approaches" defines the underlying models. The column "MSPE_ratio" indicates the $\text{MSPE}_{ratio}$ statistics for the Linear Combination (Ridge regression) weighting approach, the $\text{MSPE}_{ratio}$ statistics for the median of individual models, and the $\text{MSPE}_{ratio}$ statistics for the best individual model for a given country. The column "DM(t-stats)" shows the DM statistics for the Linear Combination (Ridge regression) weighting approach, the DM statistic for the median of individual models, and the DM statistics for the best individual model for a given country. The column "DM(p-value)" shows the p-values associated with the DM statistics. The column "$\hat{p}$" indicates the proportions of correct predictions of direction for the Linear Combination (Ridge regression) weighting approach, the proportions of correct predictions of direction for the median of individual models, and the proportions of correct predictions of direction for the best individual model. The column "PT(t-stats)" indicates the Pesaran-Timmermann (PT) statistics for the Linear Combination (Ridge regression) weighting approach, the PT statistics for the median of individual models, and the PT statistics for the best individual model and the column "PT(p-value)" shows the p-values associated with the PT statistics. All reported test statistics are one-sided. The significance level is 10%. The column "Avg. Return (%)" shows the annualized average returns in percentage, and the column " Sharp_ratio " shows the annualized Sharp-ratio values for each underlying model for a given country.

Overall results show that the linear combination (Ridge regression) and Directional Prediction weighting approaches presented in this essay for the first time successfully predict the direction of changes in exchange rates correctly for most countries in both advanced and emerging/developing countries across different horizons. The empirical evidence indicates that combining forecasts from different individual models yields improved forecasting results for changes in exchange rates.

Moreover, the results highlight the importance of forecast evaluation methods. In a few cases, weighting approaches could improve forecasting results in terms of the $\text{MSPE}_{ratio}$ and DM statistics. However, according to the $\hat{p}$ measure, these approaches have a higher proportion of correct directions than the WN model. These findings are critical for participants and policymakers in financial markets. The empirical evidence indicates, on average, weighting methods have higher values of $\hat{p}$ in the emerging/developing group than the advanced group. A reason could be that the number of individual models used in the forecast combinations is not the same across advanced and emerging/developing countries because of the lack of data availability. Also, the different results could come from differences in the economic condition stability, insufficient financial market maturity, and economic strategies.

## 3.6 Conclusion

This essay addressed model uncertainty by forecast combination methods. Individual investors, companies, and governments, among other stakeholders, face uncertainty about financial markets, including the foreign exchange market. In addition, many individual forecasting models may have weak out-of-sample performance, but they include some information. Combining forecasts from individual models could solve model uncertainty.

I proposed the Directional Prediction method as a new weighting (combination) method. I used the linear and convex combinations estimations using the regularization technique, specifically the Ridge regression. This essay is the first to combine the forecasts formed by a large number of individual models that included macroeconomic and investor sentiment (behavioral) variables for advanced and emerging/developing countries, to the best of my knowledge. Note that the total number of individual models because of the limited data is not the same in both countries' groups.

I used the $\text{MSPE}_{ratio}$, DM, and PT statistics, to evaluate the out-of-sample performance of weighting methods. The results show even though, in general, weighting approaches do not outperform the WN model in terms of the $\mathrm{MSPE}_{ratio}$ and DM statistics, these methods yield the higher proportions of correct direction (sign) according to the $\hat{p}$ values.

The findings show weighting methods have higher values of $\hat{p}$ in emerging/developing countries compared to advanced countries. There could be some reasons. The number of available individual models in advanced countries is not similar to some emerging/developing countries because of limited data in those emerging/developing countries. In addition, economic strategies and policies are not necessarily the same across advanced and emerging/developing countries. Moreover, the different results could come from insufficient development of financial markets in emerging/developing countries.

Policymakers and participants in the financial market could use the introduced weighting approaches in this essay and not be worried about eliminating any model and information. The empirical evidence highlights the advantage of using the new weighting approaches, including the Directional Prediction and the linear combination (Ridge regression). Using these two weighting approaches as investment strategies could lead to profits, especially in emerging/developing countries at horizon=1, because the values of Sharpe-ratio are considerable.

# Conclusion

This dissertation aimed to examine and perhaps improve upon out-of-sample forecasting results for changes in exchange rates using behavioral economics and ensemble (combination) methods to account for model uncertainty. Previous studies indicated that predicting exchange rates is a challenging task. As mentioned before, the performance of existing models depends on different elements such as the selection of predictors, estimation methods, evaluation tests, and horizons. Therefore, no single forecasting model in the literature provides consistent performance across all currencies, horizons, and evaluation tests.

Behavioral economics has increasingly been used as an alternative approach to explain some anomalies and unexpected behaviors in financial markets. Given influence of people's judgments and insights on economic and financial decisions, this dissertation used behavioral economics in the first and second essays to develop out-of-sample forecasting models for changes in exchange rates for both advanced and emerging/developing countries. This dissertation took a different approach in the third essay. Individual investors, corporations, and governments, among other stakeholders, face uncertainty about the foreign exchange market. I used ensemble (combination) methods to address model uncertainty and form combined forecasts for changes in exchange rates.

For assessments in all three essays, I compared out-of-sample forecasting performance of models with the benchmark, White Noise (WN) model of $\Delta s_t$, assuming the level of exchange rates, $S_t$ follows a random walk. I used different statistics, including the ratio of Mean Squared Prediction Error ($\text{MSPE}_{ratio}$), Diebold-Mariano (DM), and Pesaran-Timmermann (PT). In addition to statistical forecast evaluation measures, I used Sharpe-ratio statistics to evaluate the investment strategies, using the expected exchange rate changes to calculate investment returns.

The first essay applied behavioral economics using behavioral heuristics, including availability heuristic, conservatism, pessimism, optimism, overconfidence, and underconfidence. I also used the median of exchange rate historical data as an anchor and presented two heuristics, anchoring-toward and anchoring-away. Using these heuristics either individually or in combinations, I altered the probability distribution of data and used subjective probabilities instead of objective probabilities. In other words, I assigned different weights to the exchange rate historical data. I used these weights in the autoregressive models to estimate coefficients and form forecasts for exchange rates.

The empirical results using $\text{MSPE}_{ratio}$ or DM statistics showed that none of models outperform the WN model. However, PT statistics showed that forecasting models under the anchoring-toward and optimism assumptions beat the WN model in correctly predicting the direction of changes in exchange rates at 1-month-ahead and 12-month-ahead forecasts in some emerging/developing countries. While forecasting models did not provide improved point forecasts, they gave better results than the WN model in predicting the direction of changes in exchange rates.

The second essay took a different approach and incorporated behavioral economics using investor sentiment variables. This essay investigated whether including these variables in macroeconomic models could improve forecasting results for changes in exchange rates. In doing so,

1. I evaluated the exchange rates' predictability using macroeconomic models. Macroeconomic models included purchasing power parity (PPP), uncovered interest rate parity (UIRP), and Taylor Rule models. In addition, I explored the predictive ability of the terms of trade index (in changes), changes in the commodity price (in both real and nominal terms), and changes in the oil price (in both real and nominal terms). I also inspected out-of-sample performance of the Extended UIRP model, where US and foreign country inflation rates are added to the standard uncovered interest rate parity

model.

2. I included investor sentiment variables, including the business confidence, composite leading indicator, consumer confidence, VIX, and SKEW, to macroeconomic models to examine whether this approach could improve the predictive ability of the underlying macroeconomic models.

The findings showed that forecasting models did not outperform the WN model in terms of $\text{MSPE}_{ratio}$ or DM statistics. However, according to PT statistics, these models did well in predicting the direction of changes in exchange rates. The empirical results indicated that underlying models included investor sentiment variables outperformed the WN model for several countries at 1-month-ahead and 12-months-ahead forecasts. The empirical evidence showed using changes in oil price and commodity price (both in nominal and real terms) provided promising forecasting results for some countries.

The third essay took a different point of view and applied ensemble (combination) methods to forecast changes in exchange rates. Combining forecasts has a practical value in addressing model uncertainty that individual investors, corporations, and governments face in the foreign exchange market. Moreover, individual models may have weak out-of-sample performance in predicting the exchange rates. A critical advantage of combining forecasts is that there is no need to identify the best individual model and prevents a look-ahead bias.

I proposed a new combination (weighting) approach that used the Directional Prediction ability to combine forecasts. I combined predictions from individual models that included macroeconomic and investor sentiment variables simultaneously. To the best of my knowledge, no previous studies considered this approach. The individual models were introduced and explained in the second essay. I also applied the regularization technique (Ridge regression) in the linear and convex combination estimations to form combined forecasts for changes in exchange rates.

The findings indicated while weighting (combination) approaches were not successful in terms of $\text{MSPE}_{ratio}$ or DM statistics, these approaches provided promising results in predicting the direction of change in exchange rates using PT statistics.

In emerging/developing countries, the empirical evidence showed that the linear combination (Ridge regression) and Directional Prediction methods improved the $\hat{p}$ value versus the median of individual models at 1-month-ahead and 12-months-ahead forecasts. In advanced countries, the linear combination (Ridge regression) method had improved the $\hat{p}$ value versus the median of individual models at 1-month-ahead forecasts. Still, this method did not statistically significantly outperform the WN model.

Taken together, the findings in three essays suggest incorporating and adding behavioral heuristics and investor sentiment indices, and also combining forecasts from individual models could lead to better forecasting results for changes in exchange rates, especially in emerging/developing countries. There are reasons for different results across advanced and emerging/developing countries, including the lack of deep foreign exchange markets in emerging/developing countries, various economic policies, trades volume, market speculations, etc.

Even though results are not necessarily the same across countries, they indicate it is essential to note that people analyze and use available information differently to make economic and investment decisions. The empirical evidence confirms that people's opinions and perceptions of financial market risk could notably impact forming better forecasts for exchange rates.

The exchange rate is a crucial financial variable that significantly impacts financial decisions made by foreign exchange investors, international companies, bankers, importers, exporters, financial institutions, businesses, policymakers, and tourists. Therefore, forecasting exchange rates is essential for economic and managerial decision-making. The proposed and applied approaches in this dissertation should help investors and policymakers better position themselves for both short and long terms. Having some ideas about the direction of exchange rates' movements could put policymakers and investors in active positions, not reactive. In addition, outcomes of this dissertation could aid governments, corporations, and individual investors in identifying the best places to invest.

Moreover, this dissertation highlights the importance of using different statistical metrics to evaluate out-of-sample performance of models. The results show that forecasting models are mostly not successful in correctly forecasting point forecasts. Yet, these models perform well to predict the direction of changes in exchange rates accurately.

While any modeling approach faces certain limitations, this dissertation shows that (a) behavioral economics and ensemble (combination) methods lead to promising forecasting results, (b) using ensemble (combination) methods mitigates model uncertainty.

# Bibliography


Abel, A. B. (2002). An exploration of the effects of pessimism and doubt on asset returns. *Journal of Economic dynamics and control*, *26*(7), 1075–1092.

Aiolfi, M., Capistrán, C., & Timmermann, A. (2010). Forecast combinations. In *The oxford handbook of economic forecasting.*

Akerlof, G. A., & Shiller, R. J. (2010). *Animal spirits: How human psychology drives the economy, and why it matters for global capitalism.* Princeton university press.

Alba, J. D., Park, D., & Xie, T. (2015). Predictability of exchange rates with taylor rule fundamentals: Evidence from inflation-targeting emerging countries. *Emerging Markets Finance and Trade*, *51*(4), 714–728.

Baicker, K., Congdon, W. J., & Mullainathan, S. (2012). Health insurance coverage and take-up: Lessons from behavioral economics. *Milbank Quarterly*, *90*(1), 107–134.

Bates, J. M., & Granger, C. W. (1969). The combination of forecasts. *Journal of the Operational Research Society*, *20*(4), 451–468.

Beckmann, J., & Schüssler, R. A. (2014). Forecasting exchange rates under parameter and model uncertainty. *Available at SSRN 2502382*.

Boudoukh, J., Richardson, M., & Whitelaw, R. (1998). The best of both worlds. *Risk*, *11*(5), 64–67.

Broda, C. (2004). Terms of trade and exchange rate regimes in developing countries. *Journal of International economics*, *63*(1), 31–58.

Campbell, J. Y., W.Lo, A., & MacKinlay, A. C. (1997). *The econometrics of financial markets.* Princeton university press.

Cboe. (2020a). http://www.cboe.com/skew.

Cboe. (2020b). http://www.cboe.com/vix.

Cerra, V., & Saxena, S. C. (2010). The monetary model strikes back: Evidence from the world. *Journal of International Economics*, *81*(2), 184–196.

Chan, Y. L., Stock, J. H., & Watson, M. W. (1999). A dynamic factor model framework for forecast combination. *Spanish Economic Review*, *1*(2), 91–121.

Chen, & Rogoff, K. (2003). Commodity currencies. *Journal of international Economics*, *60*(1), 133–160.

Chen, Rogoff, K., & Rossi, B. (2010). Can exchange rates forecast commodity prices? *Quarterly Journal of Economics*, *125*(3), 1145–1194.

Chen, S., & Hsu, C.-C. (2019). Do stock markets have predictive content for exchange rate movements? *Journal of Forecasting*, *38*(7), 699–713.

Cheung, Y.-W., Chinn, M. D., Pascual, A. G., & Zhang, Y. (2019). Exchange rate prediction redux: new models, new data, new currencies. *Journal of International Money and Finance*, *95*, 332–362.

Chow, K. V., & Denning, K. C. (1993). A simple multiple variance ratio test. *Journal of Econometrics*, *58*(3), 385–401.

Clark, T. E., & West, K. D. (2006). Using out-of-sample mean squared prediction errors to test the martingale difference hypothesis. *Journal of Econometrics*, *135*(1), 155–186.

Clark, T. E., & West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of econometrics*, *138*(1), 291–311.

Della Corte, P., Sarno, L., & Tsiakas, I. (2009). An economic evaluation of empirical exchange rate models. *Review of Financial Studies*, *22*(9), 3491–3530.

Della Corte, P., & Tsiakas, I. (2012). Statistical and economic methods for evaluating exchange rate predictability. *Handbook of exchange rates*, 221–263.

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, *13*(3).

Edwards, W. (1968). Conservatism in human information processing. *Formal representation* *of human judgment.*

Elliott, G., & Timmermann, A. (2005). Optimal forecast combination under regime switching. *International Economic Review*, *46*(4), 1081–1102.

Engel, C., Lee, D., Liu, C., Liu, C., & Wu, S. P. Y. (2019). The uncovered interest parity puzzle, exchange rate forecasting, and taylor rules. *Journal of International Money and Finance*, *95*, 317–331.

Engel, C., Mark, N. C., & West, K. D. (2015). Factor model forecasts of exchange rates. *Econometric Reviews*, *34*(1-2), 32–55.

Farmer, L., Schmidt, L., & Timmermann, A. (2019). Pockets of predictability. *Available at SSRN 3152386.*

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

Fisher, I. (1896). Appreciation and interest. *Publications of the American Economic Association*, *11*(4), 1–98.

Fong, W. M., Koh, S. K., & Ouliaris, S. (1997). Joint variance-ratio tests of the martingale hypothesis for exchange rates. *Journal of Business & Economic Statistics*, *15*(1), 51–59.

Genre, V., Kenny, G., Meyler, A., & Timmermann, A. (2013). Combining expert forecasts: Can anything beat the simple average? *International Journal of Forecasting*, *29*(1), 108–121.

Gourieroux, C., Jasiak, J., & Tong, M. (2021). Convolution-based filtering and forecasting: An application to wti crude oil prices. *Journal of Forecasting.*

Granger, C. W., & Ramanathan, R. (1984, Apr). Improved methods of combining forecasts. *Journal of Forecasting (pre-1986)*, *3*(2), 197.

Hauzenberger, N., & Huber, F. (2020). Model instability in predictive exchange rate regressions. *Journal of Forecasting*, *39*(2), 168–186.

Hochberg, Y. (1974). Some generalizations of the t-method in simultaneous inference.

*Journal of Multivariate Analysis*, *4*(2), 224–234.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, *12*(1), 55–67.

Hsiao, C., & Wan, S. K. (2014). Is there an optimal forecast combination? *Journal of Econometrics*, *178*, 294–309.

Ince, O. (2014). Forecasting exchange rates out-of-sample with panel methods and real-time data. *Journal of International Money and Finance*, *43*, 1–18.

Ito, T., Masuda, M., Naito, A., & Takeda, F. (2021). Application of google trends-based sentiment index in exchange rate prediction. *Journal of Forecasting*.

Jaworski, K. (2021). Forecasting exchange rates for central and eastern european currencies using country-specific factors. *Journal of Forecasting*.

Kouwenberg, R., Markiewicz, A., Verhoeks, R., & Zwinkels, R. C. (2017). Model uncertainty and exchange rate forecasting. *Journal of Financial and Quantitative Analysis*, *52*(1), 341–363.

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling* (Vols. 26, p.13). Springer.

Lo, A. W., & MacKinlay, A. C. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of financial studies*, *1*(1), 41–66.

Markiewicz, A. (2012). Model uncertainty and exchange rate volatility*. *International Economic Review*, *53*(3), 815–844.

McCracken, M. W. (2007). Asymptotics for out of sample tests of granger causality. *Journal of Econometrics*, *140*(2), 719–752.

Meese, R., & Rogoff, K. (1983a). Empirical exchange rate models of the seventies: Do they fit out of sample? *Journal of international economics*, *14*(1), 3–24.

Meese, R., & Rogoff, K. (1983b). The out-of-sample failure of empirical exchange rate models: sampling error or misspecification? In *Exchange rates and international macroeconomics* (pp. 67–112). University of Chicago Press.

Meese, R., & Rogoff, K. (1988). Was it real? the exchange rate-interest differential relation over the modern floating-rate period. *the Journal of Finance*, *43*(4), 933–948.

Molodtsova, T., & Papell, D. H. (2009). Out-of-sample exchange rate predictability with taylor rule fundamentals. *Journal of international economics*, *77*(2), 167–180.

Morales-Arias, L., & Moura, G. V. (2013). Adaptive forecasting of exchange rates with panel data. *International Journal of Forecasting*, *29*(3), 493–509.

OECD. (2020a). Business confidence index (BCI) (indicator). doi: 10.1787/3092dc4f-en.

OECD. (2020b). Composite leading indicator (CLI) (indicator). doi: 10.1787/4a174487-en.

OECD. (2020c). Consumer confidence index (CCI) (indicator). doi: 10.1787/46434d78-en.

Pesaran, M. H., & Timmermann, A. (1992). A simple nonparametric test of predictive performance. *Journal of Business & Economic Statistics*, *10*(4), 461–465.

Rossi, B. (2013). Exchange rate predictability. *Journal of Economic Literature*, *51*(4), 1063–1119.

Semenov, A. (2008). Historical simulation approach to the estimation of stochastic discount factor models. *Quantitative Finance*, *8*(4), 391–404.

Semenov, A. (2009). Departures from rational expectations and asset pricing anomalies. *The Journal of Behavioral Finance*, *10*(4), 234–241.

Semenov, A. (2019). Behavioral heuristics and the subjective probability of asset returns. *Working paper, York university*.

Sharpe, W. F. (1966). Mutual fund performance. *The Journal of business*, *39*(1), 119–138.

Šidák, Z. (1967). Rectangular confidence regions for the means of multivariate normal distributions. *Journal of the American Statistical Association*, *62*(318), 626–633.

Smith, J., & Wallis, K. F. (2009). A simple explanation of the forecast combination puzzle. *Oxford Bulletin of Economics and Statistics*, *71*(3), 331–355.

Stock, J. H., & Watson, M. W. (2003). How did leading indicator forecasts perform during the 2001 recession? *FRB Richmond Economic Quarterly*, *89*(3), 71–90.

Stock, J. H., & Watson, M. W. (2004). Combination forecasts of output growth in a seven-country data set. *Journal of forecasting*, *23*(6), 405–430.

Stracca, L. (2004). Behavioral finance and asset prices: Where do we stand? *Journal of economic psychology*, *25*(3), 373–405.

Timmermann, A. (2006). Forecast combinations. *Handbook of economic forecasting*, *1*, 135–196.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive psychology*, *5*(2), 207–232.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science*, *185*(4157), 1124–1131.

Tversky, A., & Kahneman, D. (1975). Judgment under uncertainty: Heuristics and biases. In *Utility, probability, and human decision making* (pp. 141–162). Springer.

Washington, DC: Central Intelligence Agency. (2019). *The world factbook.* (https://www.cia.gov/library/publications/the-world-factbook/index.html). (Accessed: 2019-15-10)

Wright, J. H. (2008). Bayesian model averaging and exchange rate forecasts. *Journal of Econometrics*, *146*(2), 329–341.

Wu, C.-H., Wu, C.-S., & Liu, V. W. (2009). The conservatism bias in an emerging stock market: Evidence from taiwan. *Pacific-Basin Finance Journal*, *17*(4), 494–505.

# Appendices

## Appendix A: Joint Variance Ratio

In this Appendix, I explain the joint variance ratio test in details. Because the variance ratio for all $q$ selected should be equal to one, the common method to examine the null is the multiple comparison. To conduct the multiple comparison, I proceed as in the Chow and Denning (1993) and Fong et al. (1997) studies.

Let $VR_m = [VR(q_1), \ldots, VR(q_m)]'$ be an $m \times 1$ vector of variances ratios, $\{q_i | i = 1, \ldots, m\}$ be a set of preselected lags such that $q_1 (= 2) < q_2 < \ldots < q_m \leq \frac{T_1}{2}$, and $T_1$ be the sample size. I consider $l_m$ as an $m \times 1$ unit vector, and define $\Psi = [\Psi(q_1), \ldots, \Psi(q_m)]' = \hat{W}^{((\frac{-1}{2}))}(\hat{V}R_m - l_m)$ as a vector of Lo-MacKinlay standardized variance ratio test statistics, where is the estimator of the covariance matrix of $VR_m$ Fong et al. (1997). From the Chow and Denning (1993) study, the distribution of $\hat{V}R_m$ converges asymptotically to an m-variate distribution $N(0, W)$, where $W$ is a diagonal matrix with the elements equal to $\theta(q_i)$. The distribution of $\Psi$ converges asymptotically to an m-variate standard normal distribution $N(0, l_m)$under the null hypothesis.

Because the null hypothesis (the random walk hypothesis) can be rejected if any of the $\hat{V}R(q_i)'$s significantly differ from one. I should only focus on the maximum absolute value of the vector of test statistics, that is $\max[|\Psi(q_1)|, \ldots, |\Psi(q_m)|]$. The simplest approach to control for the joint significance of the test is to make a Bonferroni adjustment, using the following probability inequality:

$\max[|\Psi(q_1)|, \ldots, |\Psi(q_m)|] \leq z_{\frac{\alpha^*}{2}}] \geq 1 - \alpha$ , where $z_{\frac{\alpha^*}{2}}$ is the upper $\frac{\alpha^*}{2}$ point of the standard normal distribution and $\alpha^* = \alpha/m$. Assuming that $\Psi$ are independent, Šidák (1967) obtains a slightly sharp confidence interval than the Bonferroni inequality. In the Hochberg (1974) study, using Sidak results that even under the condition that $\Psi$ are correlated with an arbitrary correlation matrix $\Omega$, the following holds: $\Pr[\max[|\Psi(q), \ldots, |\Psi(q)|] \leq SMM(\alpha; m; T_1)] \geq (1 - \alpha)$, where $SMM(\alpha; m; T_1)$ is the upper $\alpha$ point of the Studentized

Maximum Modulus (SMM) distribution with parameter $m$ and $T_1$ (sample size) degrees of freedom. The Hochberg inequality is asymptotically ($T_1 = \infty$) equivalent to Sidak inequality (Chow and Denning (1993)).

Chow and Denning (1993) control the size of multiple variance ratio test by comparing the calculated values of standardized test statistic $\Psi$ with the SMM critical values. If the maximum absolute value of $\Psi$ is greater than the SMM critical value at a predetermined significance level then the random walk hypothesis is rejected.

## Appendix B: WEO Countries' Categories

According to WEO, the group of emerging market and developing economies (154) includes all those that are not classified as advanced economies. Emerging market and developing economies are also classified according to analytical criteria. The analytical criteria reflect the composition of export earnings and a distinction between net creditor and net debtor economies. The analytical criterion source of export earnings distinguishes between the categories fuel (Standard International Trade Classification) and nonfuel and then focuses on nonfuel primary product. Economies are categorized into one of these groups when their main source of export earnings exceeded 50 percent of total exports on average between 2012 and 2016. According to the World Economic Outlook, the financial criteria focus on net creditor economies, net debtor economies, heavily indebted poor countries, and lo-income developing countries. Economies are categorized as net debtors when their latest net international investment position, where available, was less than zero or their current account balance accumulations from 1972 (or earliest available data) to 2016 were negative. Net debtor economies are further differentiated on the basis of experience with debt servicing.

# Appendix C: Data Descriptions and Sources

Table C.1: Data Descriptions and Sources

| Countries | Time-span In Estimation | Indicator | Source & Direct Link |
|---|---|---|---|
| All | Jan 2000 - Dec 2020 | Exchange Rates, US Dollar per Domestic Currency, End of Period | IMF-International Financial Statistics https://data.imf.org/?sk=322eb2eb-4d17-405a-8713-f6132af6e0ed&hide_uv=1 |
| All | Jan 2000 - Dec 2020 | Prices, Consumer Price Index, All items | IMF-International Financial Statistics https://data.imf.org/?sk=e4afee0f-039b-47f6-9508-81e57659c564&hide_uv=1 |
| All[1] | Jan 2000 - Oct 2020 | Economic Activity, Industrial Production, Seasonally adjusted | IMF-International Financial Statistics https://data.imf.org/?sk=0828146b-47d2-49cf-ba7a-aebed8b5434d&hide_uv=1 |
| All | Jan 2000 - Dec 2020 | Crude Oil (petroleum), West Texas Intermediate 40 API, Midland Texas, US$ per barrel | IMF- Primary Commodity Price System https://data.imf.org/?sk=471DDDF8-D8A7-499A-81BA-5B332C01F8B9 |
| All | Jan 2000 - Dec 2020 | All Commodity Price Index, includes both Fuel and Non-Feul Price Indices | IMF- Primary Commodity Price System https://data.imf.org/?sk=471DDDF8-D8A7-499A-81BA-5B332C01F8B9 |
| All[2] | Jan 2000 - Dec 2020 | Immediate Rates (< 24 Hours), Call Money (Interbank Rate) | Refinitiv [3] https://eikon.thomsonreuters.com/index.html |
| All[4] | Jan 2000 - Dec 2020 | Money Supply M1, Standardized, Seasonally adjusted, USD | Refinitiv https://eikon.thomsonreuters.com/index.html |

[1] Because of limited data, Chile, Colombia, Iceland, and India are excluded.

[2] Because of limited data, Estonia and Slovenia are excluded.

[3] Refinitiv (formerly Thomson Reuters Financial & Risk Business) is not open-source.

[4] Because of limited data, Belgium, Brazil, Mexico, Slovakia, Slovenia, Spain, and Turkey are excluded.

## Table C.2: Data Descriptions and Sources, Continued

| Countries | Time-span In Estimation | Indicator | Source & Direct Link |
|---|---|---|---|
| All[1] | Jan 2000 - Dec 2020 | Terms of Trade, Index | Refinitive[2] https://eikon.thomsonreuters.com/index.html |
| All | Jan 2000 - Dec 2020 | Cboe SKEW Index (SKEW) | Cboe Global Markets https://www.cboe.com/us/indices/dashboard/skew/ |
| All | Jan 2000 - Dec 2020 | Cboe Volatility Index (VIX) | Cboe Global Markets https://www.cboe.com/tradable_products/vix/vix_historical_data/ |
| All[3] | Jan 2000 - Dec 2020 | Business confidence index (BCI) | The Organisation for Economic Co-operation and Development (OECD) https://data.oecd.org/leadind/business-confidence-index-bci.htm#indicator-chart |
| All[4] | Jan 2000 - Dec 2020 | Consumer confidence index (CCI) | The Organisation for Economic Co-operation and Development (OECD) https://data.oecd.org/leadind/consumer-confidence-index-cci.htm#indicator-chart |
| All | Jan 2000 - Dec 2020 | Composite leading indicator (CLI) | The Organisation for Economic Co-operation and Development (OECD) https://data.oecd.org/leadind/composite-leading-indicator-cli.htm#indicator-chart |

[1] Because of limited data, Chile, Hungary, India, and Slovakia are excluded. Since monthly data are not available, Austria, Israel, Luxembourg, Norway, South Africa, and United Kingdom are excluded.

[2] Refinitiv (formerly Thomson Reuters Financial & Risk Business) is not open-source.

[3] Because of limited data, Chile is excluded.

[4] Because of limited data, Chile, Colombia, Israel, Luxembourg, Mexico, Poland, and Turkey are excluded.