Title
Essays on Exchange Rates

Author
Zhou, Wenbo

Publication Date
2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
Los Angeles

Essays on Exchange Rates

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Economics

by

Wenbo Zhou

2017

ABSTRACT OF THE DISSERTATION

Essays on Exchange Rates

by

Wenbo Zhou
Doctor of Philosophy in Economics
University of California, Los Angeles, 2017
Professor Aaron Tornell, Chair

This dissertation studies the forward premium puzzle (FPP) and short-term exchange rate forecasting.

Chapter 1 studies the empirical behavior of the FPP over different subsamples instead of an average effect for the whole sample period as what is typically done in the literature. We find that the estimated slope coefficients from the Fama regression vary considerably from period to period. The signs of the slope estimate could be both significantly positive and negative. Our contribution is to show that the variation of the slope estimates is not random, rather it is driven by a common factor. We document a link between the variation and investors' long-run uncertainty about the economy. The long-run uncertainty index is specific to individual countries and defined as either a large fall in the real GDP growth rate or an inflation hike compared to past levels. We find that the long-run uncertainty index and its lags contribute to the positiveness of the slope estimate. The effect lasts longer for developed countries than emerging ones. The FPP exists if there is no long-run uncertainty about the economy but disappears with such uncertainty.

Chapter 2 provides a potential theoretical framework to understand the empirical facts described in Chapter 1 based on Li and Tornell (2015). They show that the robustness against model misspecification can generate both positive and negative Fama slope coefficients, depending on investors' beliefs about the relative importance of transitory and persistent interest rate shocks. But they miss one step linking the economic fundamentals to

the assumed interest rate differential model. We fill the gap using the long-run risk model with two variables: real consumption growth and inflation. We map the persistent interest rate shocks to long-run shocks to either consumption growth or inflation, which matches the long-run uncertainty defined in Chapter 1. We then qualitatively explain the empirical facts of time-varying slope estimates.

Chapter 3 implements an empirical forecasting strategy based on what the Federal Open Market Committee (FOMC) says after their regular meetings. We use several techniques from natural language processing including bag-of-words, latent semantic analysis and vector space model to construct nontraditional predictors from three types of text documents released by the FOMC. We apply a machine learning algorithm called support vector machine to forecast individual G10 currencies and also build a portfolio of all G10 currencies. For the portfolio, our out-of-sample forecasts have success ratios more than 50% for short-term prediction (less than 6 weeks) except for the 1-month horizon. Our best performance can be found for 1-week forecasting horizon. Eight out of nine currencies, as well as the portfolio, can beat the random walk model significantly using the weighted directional test.

The dissertation of Wenbo Zhou is approved.

Zhipeng Liao

Brett Trueman

Pierre-Olivier Weill

Aaron Tornell, Committee Chair

University of California, Los Angeles

2017

To my family

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor Professor Aaron Tornell for his inspiration, guidance, encouragement and support during my Ph.D. life. I am grateful to Professor Zhipeng Liao for comments on econometric techniques used in my dissertation. I would also like to thank the other members in my committee, Pierre-Olivier Weill and Brett Trueman for all their help. I thank everyone who participates the international proseminar to provide feedback on my research. Last, my deepest gratitude goes to my family for their unconditional love and support.

VITA

2009	B.Sc., Mathematics Renmin University of China Beijing, China
2012	M.A., Finance Renmin University of China Beijing, China
2014	C.Phil., Economics University of California, Los Angeles Los Angeles, California

1. The Forward Premium Puzzle and Robust Control: Empirics

1.1 Introduction

The uncovered interest parity (UIP) condition predicts that high interest rate currencies should depreciate to offset the interest rate differential between two countries. However, the empirical evidence noted in Fama (1984) and subsequent studies consistently reject the UIP. Currencies with relatively higher interest rates either appreciate or do not depreciate sufficiently to offset the interest rate differential. More specifically, the Fama regression of the exchange rate change on the interest rate differential should yield a slope coefficient of unity, but such regressions typically yield coefficients smaller than one or even negative. This stylized fact is known as the "forward premium puzzle" (FPP).

This chapter studies the empirical behavior of the FPP over different subsamples instead of an average effect for the whole sample period as what is typically done in the literature. We find that the estimated slope coefficients from the Fama regression vary considerably from period to period. The signs of the slope estimate could be both significantly positive and negative. Our contribution is to show that the variation of the slope estimates is not random, rather it is driven by a common factor. We document a link between the variation and investors' long-run uncertainty about the economy. The long-run uncertainty index is specific to individual countries and defined as either a large fall in the real GDP growth rate or an inflation hike compared to past levels. Investors feel uncertain about long-run performance of an economy if such events occur. We find that the long-run uncertainty index and its lags contribute to the positiveness of the slope estimates and the effect lasts longer for developed countries than emerging ones. The FPP exists if there is no long-run uncertainty about the economy but disappears with such uncertainty.

Specifically, we estimate the Fama regression with rolling subsamples from 1975 to 2014 for different window sizes. First, it is motivated by the fact that the slope estimates from Fama (1984) regression in the literature are different in different sample periods. If we consider the U.S. and Germany currency pair as example, the slope estimate is -1.32 in Fama (1984) for sample period between 1973 and 1982, while the number becomes 0.43 in Verdelhan (2013) for sample period between 1983 and 2010. Second, the FPP does not exist for less developed and emerging countries and Bansal and Dahlquist (2000) show the Fama slope coefficients are positive. Theoretical works (Bekaert (1996), Alvarez, Atkeson and Kehoe (2009), Verdelhan (2010) among others) tend to focus on explaining the FPP in developed countries but ignore the empirical fact in emerging ones. Our paper includes 15 developed ones and 15 less developed and emerging countries which are simply called emerging countries later. We find that these two phenomena are not separate cases. The signs of the slope estimates could be both positive and negative for both groups of countries during different sample periods. Consider developed countries as example, the signs become positive especially after the recent 2008 global financial crisis. Hence, we could understand both facts within an unified framework by checking the variation of the slope estimates.

The other side of the FPP is the profitability of carry trade, which is a simple trading strategy of borrowing from low interest rate country and investing in high interest rate country. If the FPP does not exist, carry trade is not profitable. Empirical evidence also shows positive excess returns for developed currency pairs. Figure 1.1 shows both cumulative total returns and excess returns of implementing carry trades with G10 currencies[1]. Even though this strategy is profitable over the whole sample period, we do observe a huge loss during the recent global financial crisis. The vertical line means the starting date of the Deutsche Bank carry trade strategy and we can see the excess return is near zero between 2007 and 2015. In terms of Fama regression, the slope estimates would be positive if we run the Fama regression with that sample period. This piece of evidence confirms that slope

[1] The data is downloaded from Deutsche Bank and called the Deutsche Bank G10 Currency Future Harvest Index. Investors can buy ETF which tracks this index. The strategy is implemented in a way of longing three lowest interest rate currencies and shorting three highest interest rate currencies. Weights on each currency are equal and this portfolio is balanced quarterly.

Figure 1.1: Cumulative Return of Deutsche Bank Carry Trade Index

This figure plots daily cumulative gross return and excess returns of G10 currency Future Harvest Index. Data is downloaded from DB Currency Future Harvest Index. We normalize the index to be 100 at the starting date. The vertical green line indicates the starting date of this DB ETF. t-test shows that excess return is significantly different from 0 at 95% significance level.

coefficients vary in different periods.

To match the periods when signs of the slope estimates become positive, we define a long-run uncertainty index motivated by the observation above. We follow the idea of Gourinchas and Tornell (2004) and Li and Tornell (2015) that investors' distorted belief about the importance of transitory and persistent interest rate differential shocks can determine the exchange rate in equilibrium. To be more specific, the FPP disappears if investors overestimate the importance of persistent interest rate shocks. This is what happens during crisis when investors believe the negative shocks about an economy are persistent and worry about the long-run performance of the economy. But existing crisis measure is usually constructed with ex post facts, which contains future information at the moment when investors make decision. To correct and quantify this idea, we define the country specific long-run uncertainty index as either a large fall in the real GDP growth rate or an inflation hike compared to previous 10 year moving average levels. In this sense, persistent shocks are equivalent to the long-run uncertainty from investors' perspective.

Based on the slope estimates from rolling windows regressions, we implement a second step regression of these rolling slope estimates on the long-run uncertainty index and its lags, which is in the same way as Fama and MacBeth (1973). We use both pooled and panel regressions to control the country effects and their results are similar. We find that the long-run uncertainty and its lags contribute to the positiveness of the slope estimates. Without any long-run uncertainty in an economy, the Fama regression generates negative slope estimates as the literature. For different groups of countries, the effects of the uncertainty are different. It lasts longer and affects stronger in developed countries than emerging ones which includes both less developed and emerging countries. It is a more general result which discuss both groups of countries.

The rest of paper is organized as follows. In Section 2, we discuss the related literature in empirical analysis. Section 3 reviews the forward premium puzzle in both Fama regression form and carry trade form. We describe our data in Section 4. In Section 5, we construct the long-run uncertainty index for every country and run a two-step regression to show the relationship between signs of slope coefficients and uncertainties. Section 6 concludes.

1.2 Literature

The forward premium puzzle is a well-established empirical fact in the literature. Early research finds evidence rejecting the UIP resulting in the FPP. Bilson (1981) and Fama (1984) provide early empirical work documenting the failure of UIP. They use data from early 1970s to early 1980s and conclude that UIP does not hold in many developed countries. Froot (1990) finds that the average slope coefficient across 75 published estimates is -0.88. Lewis (1995) and Engel (1996) survey the early work and establish two key results. First, the empirical tests routinely reject the null hypothesis that the forward rate is a conditionally unbiased predictor of future spot rates. Second, models of the risk premium have been unsuccessful at explaining the magnitude of this failure. Recent empirical and theoretical survey can be found in Engel (2014). Eichenbaum and Evans (1995) estimate a five variable

VAR model. They show that the currency continues to appreciate before depreciation after a U.S. monetary contraction which leads to higher domestic interest rate. It is called "delayed overshooting" and the maximum delayed period can last for two to three years depending on different currencies.

Not only the UIP condition fails with in-sample data evaluation, the ability of out-of-sample forecasting is also weak. Meese and Rogoff (1983) conclude that prediction from several forms of structured monetary models cannot significantly beat the ones generated by driftless random walk model. Mark (1995) estimates a simple regression of exchange rate changes on some fundamental values and finds the root mean square error of such models are lower compared to random walk model, especially at longer horizon (three to four years). But Faust and Rogers (2003) find that Mark (1995)'s conclusion depends on the sample period used and does not perform well in most other periods. Molodtsova, Nikolsko-Rzhevskyy and Papell (2008) and Molodtsova and Papel (2009) find evidence of forecasting power with Taylor rule models.

This result is a little different when we consider different sample periods, especially the one including the recent 2008 global financial crisis. Although we have noted some important exceptions, consensus seems to agree with the rejection of UIP condition. Burnside et al. (2006) estimate nine currency pairs of developed countries against British pound using monthly data from January 1976 to December 2005. In all the cases, the estimators are significantly less than one. For most countries, the estimated slope coefficient is significantly negative. Engel (2016) finds that 4 out of 6 countries reject UIP using data from 1979 to 2009. Verdelhan (2013) rejects UIP for less than half of the developed countries with data from November 1983 to December 2010. In summary, the behaviors of UIP seem to change across different sample periods.

However, several papers find support for the UIP condition. Huisman et al. (1998) find UIP holds in periods where the forward premiums are large. Bansal and Dahlquist (2000) show that the FPP does not seem to be present in emerging economies, and also when the U.S. interest rate exceeds the foreign rates, implying UIP is state-dependent. They use weekly data of 28 developed and emerging economies from January 1976 to May 1998.

Chaboud and Wright (2005) find that UIP holds well over a very short horizon but is rejected above six hours. Chinn and Meredith (2005) cannot reject UIP using interest rates on 5-year bonds for the U.S., Germany, Japan and Canada. Similar result can be found in Chinn (2006). Lothian and Wu (2011) find positive UIP slope coefficient for USD/GBP and USD/FRF by constructing ultra-long time series that span two centuries from 1800.

1.3 Background: the UIP and the FPP

We often use the UIP puzzle and the FPP interchangeably. The forward unbiasedness hypothesis (FUH) in the currency market simply states that the forward exchange rate is an unbiased predictor of the future spot exchange rate,

$$E_t(S_{t+1}) = F_t, \tag{1.1}$$

where S_{t+1} is the nominal spot exchange rate at time $t+1$ and F_t is the 1-period nominal forward exchange rate at time t. Throughout the paper we follow the convention that the exchange rate of a country is the domestic price of the foreign currency. We use USD (i.e., U.S. dollars) as the home currency. So an increase in the exchange rate is a depreciation of USD.

The UIP condition holds by assuming risk neutrality and rational expectations,

$$\frac{E_t\left(S_{t+1}\right)}{S_t} = \frac{1+i_t}{1+i_t^*}. \tag{1.2}$$

On the other hand, the covered interest parity (CIP) connects UIP and the FPP. Let i_t be the domestic risk-free rate and i_t^* be the analogous foreign rate. One investor uses one USD to buy $\frac{1}{S_t}$ units of foreign currency at time t, which grows at the foreign risk-free rate to $\frac{1}{S_t}\left(1+i_t^*\right)$ at time $t+1$. At the same time, she has a long position in the one-month forward contract to buy back USD at rate F_t. The payoff of investing in foreign currency is $\frac{F_t}{S_t}\left(1+i_t^*\right)$, which should equal the return in domestic investment $1+i_t$ by a no arbitrage argument, leading to CIP. Then we have

$$\frac{E_t\left(S_{t+1}\right)}{S_t} = \frac{1+i_t}{1+i_t^*} = \frac{F_t}{S_t}, \tag{1.3}$$

where the second equality comes from CIP. As a result, UIP provides the economic foundation of the FUH. If the UIP does not hold, then $E_t\left(S_{t+1}\right) \neq F_t$.

Usually we write the UIP condition in log form,

$$E_t s_{t+1} - s_t \approx i_t - i_t^* \tag{1.4}$$

$$\approx f_t - s_t \tag{1.5}$$

where s_{t+1} and s_t are the log of the exchange rates at time $t+1$ and t, f_t is the log of the forward exchange rate at time t, and the second equality still comes from CIP. The equations can also be exactly equal if we consider continuous compounding with $\exp(i_t) = \exp(i_t^*)\frac{F_t}{S_t}$.

Therefore empirical test of the UIP involves either (1.4) or (1.5) by estimating the following regression,

$$\triangle s_{t+1} = \alpha + \beta\left(i_t - i_t^*\right) + \epsilon_{t+1} \tag{1.6}$$

$$\triangle s_{t+1} = \alpha + \beta\left(f_t - s_t\right) + \epsilon_{t+1} \tag{1.7}$$

where $\triangle s_{t+1}$ is the percentage change in the exchange rate. This is commonly referred to as the "Fama regression".

Under the null hypothesis, the regression coefficients are $\alpha = 0$ and $\beta = 1$. In words, the realized depreciation of the spot exchange rates is equal to the interest difference, on average. Instead, a long history of empirical work has found that the estimated value of β to be significantly less than one, and usually less than zero. This is also referred to as the forward premium puzzle, which implies that high-interest currencies tend to appreciate rather than depreciate and forms the basis of the widely-used carry trade strategies in active currency management. In this paper, we test (1.6) directly.

Another empirical fact about the FPP is the profitability of carry trade, which exploits the failure of the UIP condition by borrowing a low interest rate currency and lending a high interest rate currency. This type of simple trade is used by practitioners. The expected excess return rx_{t+1} is

$$E_t(rx_{t+1}) = \begin{cases} E_t s_{t+1} - s_t + i_t^* - i_t & \text{if } i_t \leq i_t^* \\ i_t - \left(E_t s_{t+1} - s_t + i_t^*\right) & \text{if } i_t > i_t^* \end{cases} \tag{1.8}$$

which would be zero if the FPP does not exist.

1.4 Data

Our full sample consists of the following 16 developed countries: Australia, Belgium, Canada, Denmark, France, Germany, Italy, Japan, the Netherlands, New Zealand, Norway, Spain, Sweden, Switzerland, the United Kingdom and the United States. Because we consider the United States as the home country, there are 15 currency pairs of exchange rates among them. Since Belgium, France, Germany, Italy, the Netherlands and Spain are part of the euro area, we actually include the G10 currencies and one more Danish krone after 1999. In its recent Triennial Central Bank Survey of foreign exchange market[2], the Bank for International Settlements (BIS) reported that as of end-April 2013, the global daily turnover is $5.3 trillion per day. Foreign exchange transactions with the G10 currencies on one side of the transaction represented 178.8% of all deals[3]. There is no doubt that the G10 currencies play the most important role in the currency market.

We also include 15 less developed and emerging countries: Argentina, Brazil, Chile, Colombia, Czech Republic, Indonesia, India, Korea, Malaysia, Mexico, Philippines, Poland, South Africa, Thailand and Turkey. We balance the emerging countries across all continents. In the literature, this group of countries behave differently from developed ones and there exists no forward premium puzzle for most countries in this group. We simply call this group as emerging countries thereafter.

The exchange rates data are downloaded from Global Financial Data (GFD). For each country, the exchange rate is the daily rate on the last trading day in a month. All the exchange rates have been converted to the dollar price of one unit of foreign currency. The sample period covered is 40 years long from January 1975 to June 2014. Belgium, France,

[2]Triennial Central Bank Survey Foreign exchange turnover in April 2013: preliminary global results: `http://www.bis.org/publ/rpfx13fx.pdf`, and Triennial Central Bank Survey Global foreign exchange market turnover in 2013: `http://www.bis.org/publ/rpfxf13fxt.pdf`.

[3]As two currencies are involved in each transaction, the sum of shares in individual currencies will total 200%.

Germany, Italy, the Netherlands and Spain joined the euro area on January 1st, 1999. The legacy exchange rates derived by the irrevocable conversion rate[4] are used for currencies of euro area member countries after they switched to the euro. Since we consider changes in exchange rates, the six countries have the same exchange rate returns after January 1999.

There are two data sources for interest rates: Datastream and GFD. Daily data of 1-month annual Eurorates are provided by Intercapital from Datastream. The data are the midpoint of the offer and bid rates. We construct the monthly series by using the interest rate on the last trading day. If the data period is relatively short from Datastream, monthly 1-month interbank interest rates and monthly 1-month T-bill rates from GFD are used instead. Original data are expressed at annual rates in percent, and we transform the annual rate into monthly rate as follows:

$$i_{1m} = 100 \times \left(\left(1 + \frac{i_{1y}}{100} \right)^{1/12} - 1 \right). \tag{1.9}$$

The data availability of interest rates is different for each country. The interest rates for euro area countries are the same after January 1st, 1999. Compared with the exchange rates, the sample periods for interest rates are shorter. As a result, the final sample periods used in the regressions are determined by the availability of interest rates. This is particularly true for emerging countries, which do not have long history of qualified interest rate data.

The economy fundamental data which are used to construct the long-run uncertainty measure are from World Development Indicators (WDI) of the World Bank. Out of all economic indicators, we pick real GDP growth and inflation. The reason that we pick these two indicators can be found in Chapter 2, where our theoretical model follows the standard pricing kernel framework in the literature and consider real GDP growth and inflation as exogenous stochastic processes. We can also extend the model and include some other indicators. But these two are enough to show the theoretical result qualitatively. The sample consists of annual data from as early as 1960 to 2014.

Table 1.1 and Table 1.2 summarize the descriptive statistics on monthly spot exchange

[4]The irrevocable conversion rates can be found from Bank of England (`http://www.bankofengland.co.uk/statistics/pages/iadb/notesiadb/Spot_rates.aspx`).

rate changes and interest rate differentials defined as the difference between the U.S. interest rate and foreign interest rate for both developed and emerging countries. From Table 1.1, we can see that the U.S. dollar has appreciated against half of the advanced economies over the sample period. The interest rate in the U.S. is relatively low during the sample years compared to other countries. It is also evident that the volatility of the exchange rate returns is significantly higher than the volatility of the interest differentials. On average, the interest rates in Japan and Switzerland are the lowest among the developed countries so that JPY and CHF are often used as funding currencies. On the other hand, AUD and NZD are often considered as high yielding currencies.

Table 1.2 gives the same descriptive statistics of the dataset for emerging countries. The volatility of exchange rate changes is still much larger than the one of interest rate differentials. However, there exists significant differences between developed and emerging countries. The first difference is that the U.S. dollar appreciates against all emerging currencies except for CZK. Several currencies of countries in the South America depreciate a lot during this period. Second, the interest rate in the U.S. is lower than all emerging countries. Brazil offers highest interest rate among all countries. Third, the standard deviations of both exchange rate changes and interest rate differentials are much higher for emerging countries than developed countries, where the averages of standard deviation for emerging countries are 18.32% and 5.42%, compared to 10.85% and 0.81% for developed countries.

1.5 Empirical Analysis

1.5.1 Fama Regressions over Different Groups of Countries

We begin with the individual country time series regression (1.6) using fixed sample periods,

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

Table 1.3 reports the estimation result with full data sample for developed currency pairs. The result is consistent with the literature; the slope estimates $\hat{\beta}$'s for developed countries are always below one, and most are negative. Out of the 15 currency pairs, only Italy, Spain and Sweden have positive slope estimates, none of which is significantly different from zero. The slope estimates for Japan, Netherlands and New Zealand are significantly smaller than 0. On the other hand, under the null hypothesis of UIP, the slope coefficient, β, is equal to 1. We can reject this null hypothesis for 10 out of the 15 currencies at 5% significance level, and two more are rejected at 10% significance level. Only Italy, Norway and Sweden cannot be rejected. This result confirms the usual finding that high interest rate currencies tend to appreciate and it is in opposition to what is predicted by the UIP. Notice that the slope coefficient of JPY is the largest in absolute value. JPY is often considered as the most puzzling currency partially because Japan has provided the lowest interest rate for the past several decades.

Table 1.4 shows the result for emerging countries. It is also consistent with the literature; the slope estimates $\hat{\beta}$'s for most emerging countries are positive and we cannot reject the UIP in general. Eight out of 15 countries have significantly positive $\hat{\beta}$. Only Indonesia, Korea and South Africa have negative slope estimates but they are insignificant at 10% level. In terms of null hypothesis that $\beta = 0$, three countries including Argentina, Brazil and South Africa can be rejected. However, only South Africa is rejected with the alternative hypothesis that $\beta < 1$.

Hence, by replicating the Fama regressions over whole sample period, we confirm the stylized facts that the FPP exists in developed countries but not in emerging ones.

Table 1.5 lists the estimation results for developed currencies when the 1975-2014 sample is broken into four equal sub-periods. Most slope estimates change signs across the four samples. They have negative signs in Panel I and III but positive signs in the other two. All the slope estimates are positive in Panel IV whose subsample covers the recent financial crisis. In fact, after the middle of 2007, the USD sharply appreciated against the high interest rate AUD several times while volatility increased. In particular, after the bankruptcy of Lehman Brothers in September 2008, the USD appreciated by more than 23% against the AUD in three months. At the same time, the relatively low interest rate USD depreciated 15% against the low interest rate JPY. For all the countries included in our sample, high interest rate currencies depreciated sharply during the 2008 financial crisis. The volatility during this period is extremely high compared with other periods. On the other hand, during the subsample period in Panel III, the currency market is relatively stable with low volatility. The 1997 Asian crisis does not affect the developed countries in the Europe and Northern America much. The evidence here shows that the FPP seems to exist during periods with crisis, but is not present in periods when economies are stable. As a result, the Fama regressions cannot help us understand the relationship between exchange rate returns in different economic states. We consider the rolling window estimation to evaluate the time varying behavior of the slope estimates in the Fama regressions in the next subsection.

1.5.2 Rolling Sample Regressions

In this subsection, we repeat the previous regression with different subsample periods. We conduct rolling window estimation to show the evolution of the slope estimates for all countries over the past 40 years.

With fixed rolling window size h years, we change the initial date from January 1975 (January 1982 for emerging countries) to $12 \times h$ months before June 2014. Notice that each country may have different initial date due to data availability. The initial date moves forward by 1 month at a time. For example, for $h = 10$, we estimate the first regression with a sample period from January 1975 to December 1984, resulting in a point estimate $\hat{\beta}_1$. Then we move the starting date forward by 1 month and get $\hat{\beta}_2$ with sample period from February 1975 to January 1985. We repeat this process until the starting date reaches $12 \times h$ months before June 2014, which is July 2004 and leaves 10 years of data for the last regression. Rolling estimations provide a good sense of how the UIP condition behaves in different time periods and which period is more biased. Therefore, for each currency, we have one series of estimated slopes and then 30 series in total.

We consider different window sizes h varying from 5 to 10 years in rolling window estimation for all countries. From Figure 1.4 to Figure 1.4, we report $h = 7$ and 10 for developed countries and $h = 5$ for emerging countries. We focus on shorter rolling windows for emerging group because their data samples are relatively shorter.

Figure 1.4 shows the results for all developed countries when the rolling window size $h = 7$. Each graph plots the series of slope estimates and their corresponding 95% confidence intervals. The thick red line represents slope estimates and the black dotted line indicates the 95% confidence interval. The horizontal dashed line represents the null hypothesis $\beta = 1$. The date 1982 means that the regression sample period is from January 1976 to December 1982. There are several interesting observations from these graphs. Let us take DEM for example. There is considerable variation over time in the point estimates of β. The estimated slope is significantly negative at the beginning and then insignificantly positive roughly between 1990 and 2000. It turns to be significantly negative again in early 2000s but significantly positive and even $\hat{\beta} > 1$ in recent subsamples which includes 2008 financial crisis. This pattern is similar for all currencies in developed economies except for JPY. Meanwhile, the 95% confidence interval is getting wider as the sample period covers recent financial crisis for some countries such as GBP and NZD.

In terms of the FPP, we also alternatively reject and do not reject the null that $\beta = 1$. The sample period around late 1970s rejects the UIP. Most early papers documenting the forward premium puzzle belong to this time. But there is a dominant US dollar appreciation during this time. The negative results of estimates may be due to the unique feature of that time. During the time around 2005, the UIP is also rejected. To sum up, the slope estimates vary during different sample periods. A clear observation is that if the sample period is extended to the recent financial crisis, the FPP does not exist.

Figure 1.4 shows the rolling estimation when $h = 10$ years. The result is similar as Figure 1.4 but look much smoother. For the positive part, the estimated slopes are less significant than shorter window size $h = 7$. The longer the window size is, we are more likely to reject the uncovered interest parity condition and observe the FPP. The intuition is also consistent with the stylized fact when we extend the sample to the whole sample period, which is reported in Table 1.3.

The result showed in Figure 1.4 for emerging countries is quite different from developed ones where all the developed economies behave similarly. We observe more individual characteristics from this group. The slope estimates for IDR, MYR and THB are positive around the 1997 Asian financial crisis. For MXN, the slope estimates are positive around 1995 when the peso crisis hit Mexico. Similarly, if we use larger rolling window size, the graph will look smoother.

Our question is: what drives the variation of β? We want to identify the periods when the slope estimates $\hat{\beta}$ tend to be positive. One common phenomenon we observe from both groups of economies is that $\hat{\beta}$ tend to be positive during financial crises when economies experience serious trouble. We follow the idea of Gourinchas and Tornell (2004) and Li and Tornell (2015) that investors' distorted belief about the importance of transitory and persistent interest rate differential shocks can determine the exchange rate in equilibrium. We match the periods when signs of the slope estimates become positive with a long-run uncertainty period defined in the next subsection, when investors have uncertainty about the long-run future economic fundamental performance.

1.5.3 Long-run Uncertainty Index

Since we want to investigate the relationship between fluctuation of estimated $\hat{\beta}$'s and investors' uncertainty about the economy, the definition of uncertainty is the key. Following Li and Tornell (2015), we consider that interest rate differential between two countries is driven by a persistent hidden state z_t and by a transitory observational noise v_t,

$$\begin{aligned} i_t - i_t^* &= z_t + v_t \\ z_t &= \rho z_{t-1} + w_t \end{aligned} \tag{1.10}$$

where i_t and i_t^* are domestic and foreign interest rates respectively, $v_t \sim N(0, \sigma_v^2)$, $w_t \sim N(0, \sigma_w^2)$ and they are independent under the data-generating process (DGP). The hidden state z_t itself is a AR(1) process. Investors are endowed with the baseline process but fear misspecification in some parts of the model. They balance the trade-off between negative impact of model misspecification and the cost of deviating from optimal result under DGP baseline model. Hence, investors solve a robust control optimization problem.

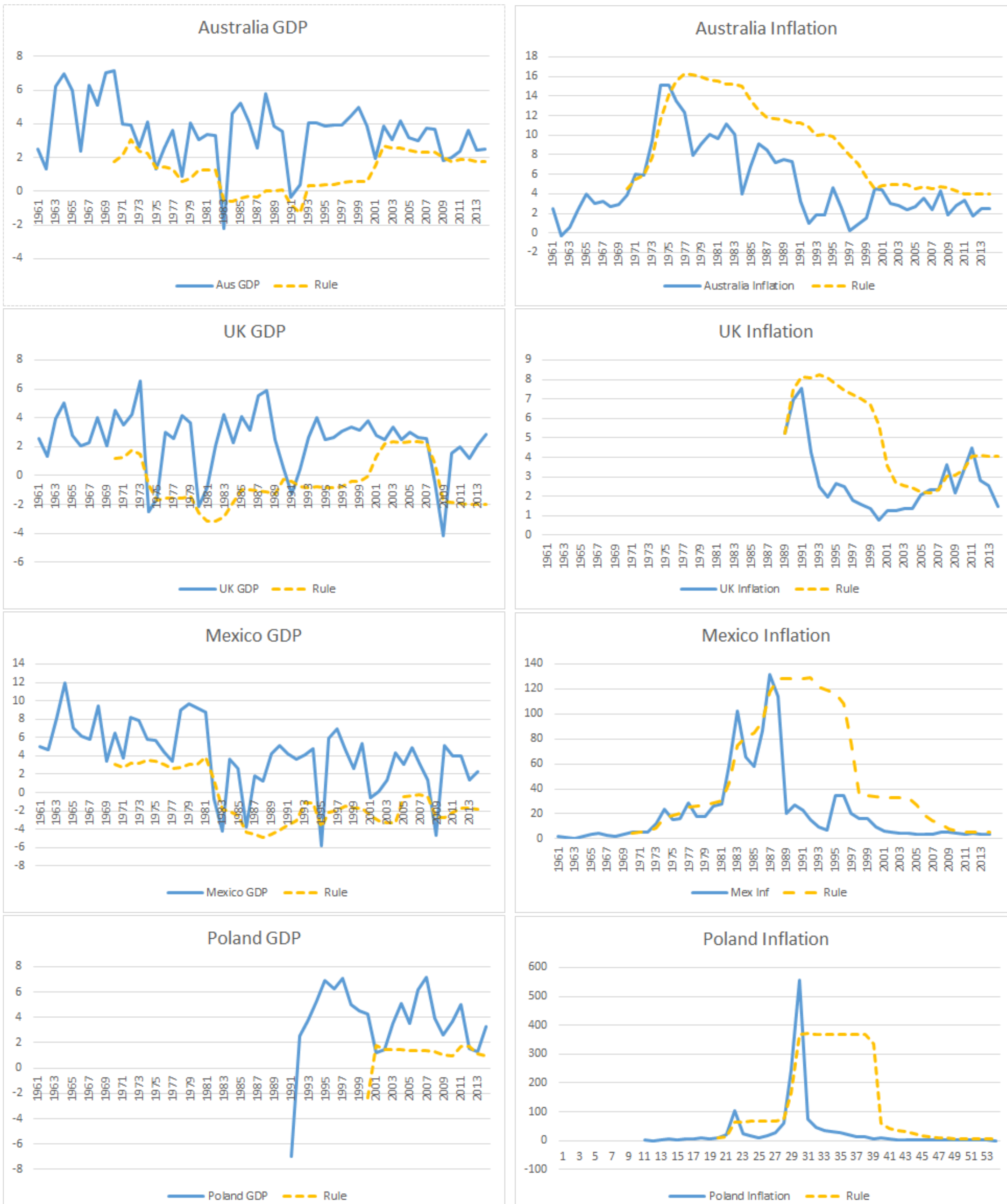
Here, we consider two possible types of structured model uncertainties. The first is observational uncertainty arising from investors' fear in the equation that links noisy observation and the unobservable persistent component, i.e., $i_t - i_t^* = z_t + v_t$. In this case, investors tend to put more weight on transitory shocks and underreact to news. Mathematically, they update their belief using the standard Kalman filter and the Kalman gain is smaller with distorted belief than under GDP process. It means overestimation of variance of the transitory shock, which is $v_t \sim N(0, \tilde{\sigma}_v^2)$ under investors' belief and $\tilde{\sigma}_v^2 > \sigma_v^2$, where tilde variable represents investors' own belief. The second one is called state uncertainty when investors fear model misspecification in the unobservable state equation, i.e., $z_{t+1} = \rho z_t + w_{t+1}$. Similarly, investors put more weights in variance of the persistent shock w_t and overreact to new information. Hence, when there exists a shock to the interest rate differential, the source of shock could be uncertain depending on investors' distorted belief about the model. Different uncertainties have different implication of the FPP. In short, observational uncertainty results in the FPP while state uncertainty does not. Details about the model setup can be found in Chapter 2. For the empirical part, we focus on the impact of state uncertainty to the FPP. The FPP disappears or the slope estimates become positive when investors have such uncertainty.

Given the two types of uncertainties, next question is: when and under what conditions would investors have structured state uncertainty instead of observational uncertainty about the model? Before we answer the question, first we map these two types of uncertainties of the model to economic fundamental variables. Suppose investors observe a negative shock in one economy, they are uncertain whether it is temporary or persistent. They also respond to the shock in a different way according to their beliefs. If investors believe it is transitory, they tend to underreact to the shock. Otherwise they overreact. Second, we specify which part of the economy is shocked. To simplify and represent the whole economy, we choose two important fundamental indicators: real GDP growth and inflation. The negative shock comes from either a huge drop of GDP growth or a sudden inflation hike. In terms of the model, if persistent, the shock is equivalent to the structured state uncertainty in the model because the persistent shock changes the state of the economy. We call this type of uncertainty in the economy as *long-run uncertainty*. Otherwise, we say the temporary shock is the same as observational uncertainty. Chapter 2 gives detailed derivation between the model and these two fundamental economic indicators.

Here we answer the previous question by setting up a reasonable rule that how investors can tell the difference between transitory and persistent shocks in the economy, and hence the difference between two types of uncertainties of the model. The rule is simply that investors have long-run uncertainty if there is a persistent negative economic shock coming from either real GDP growth or inflation. To be specific, a real GDP growth persistent shock is defined when real GDP growth falls 1.6 standard deviations or more below its h-year moving average, where h is the rolling window size. For example, we choose $h = 10$ to represent the long-run behavior. Both the standard deviation and h-year moving average are computed over a past h year rolling window. If investors make decisions at time t, then we use window from time $t - h$ to time t. The reason is that we cannot identify persistent shocks with future data from investors' perspective, conditional on information up to time t. It is similar but different from some ex post crisis definitions in the literature, such as Laeven and Valencia (2008), Laeven and Valencia (2012) and Reinhart and Rogoff (2009). An inflation persistent negative shock is computed in a similar way which is defined when inflation is higher than 1.6 standard deviations above its h-year moving average and, in addition, inflation is greater than 10 percent. Similar as Ranciere and Tornell (2015), we include the 10% inflation threshold in order to rule out false signals generated by an inflation increase from a low level, say from 1% to 1.5%.

We construct the long-run uncertainty index for each country based on the above rule. We call it as uncertainty index for simplification afterwards. The index is a time series of dummies, which equals one if the rule is satisfied for country i at particular year t, otherwise it is zero. Here in Figure 1.2 we show how to construct uncertainty indexes for two developed (Australia and the U.K.) and two emerging countries (Mexico and Poland) as examples. The solid blues lines mean the actual data of either real GDP or inflation for each country. The yellow dashed lines are computed with 10-year rolling window following the rule. For Australia, it is obvious that the real GDP growth rates drop significantly in 1983 and 2009, and the negative shocks which lower the growth are considered persistent. In our words, investors worry about the long-run performance of the growth and hold long-run

Figure 1.2: Long-run uncertainty index: construction



Note: The figure shows how to construct the long-run uncertainty. The solid blues lines mean the actual data of either real GDP or inflation for each country. The yellow dashed lines are computed with 10-year rolling window following the rule. For the growth series $\{y_t\}$, it means $\bar{y} - 1.6\sigma_{\bar{y}}$, where $\bar{y} = \frac{1}{h}\sum_{j=1}^{h} y_{t-j+1}$.

uncertainty during these years about Australia. On the other hand, inflation in Australia is below the rule except several years before 1975 with high inflation. But the sample starts from 1975 and excludes these uncertainties from our analysis. The long-run uncertainty exists either for the growth or inflation shocks. We combine these two parts and define the long-run uncertainty for Australia shown in Figure 1.3.

Figure 1.3: Long-run Uncertainty Index: developed and emerging Examples

We summarize the indexes in Figure 1.3 by combining GDP growth and inflation uncertainties. Blue bars represent real GDP growth uncertainties and orange bars indicate inflation uncertainties. At some years, these two coincide with each other and represented by darker bars. For developed countries, we can see that there is no inflation uncertainty during our sample period after 1975. This is intuitively true because we set a threshold for the inflation uncertainty which requires annual inflation to be higher than 10% which rarely happens in the developed economies. For developed countries, the uncertainty from real GDP growth clusters around two years: 1992 and 2008. These two years roughly match the starting dates of two positive periods of estimated $\hat{\beta}$'s in Figure 1.4. On the other hand, for emerging countries, investors have both real GDP growth and inflation uncertainties in our data sample. These uncertainties also happen at much higher frequencies compared to developed ones. Moreover, the periods for different countries have lower correlation and spread across different years. Details can be found in Table 1.6 which lists the uncertainty index for all countries in our data sample. Notice that due to the availability of WDI data, we might miss some uncertainties. For example, data for Czech Republic is available after 1991 for real GDP growth and 1994 for inflation. Therefore, our long-run uncertainty index can be understood as a lower bound about long-run uncertainty.

1.5.4 Long-run Uncertainty and the FPP

Here we investigate whether the uncertainty, which is short for long-run uncertainty when investors fear model misspecification in the unobservable state equation, has positive impact to the Fama regression slope coefficients.

To analyze the relationship between long-run uncertainty and the $\hat{\beta}$s, we use a similar two-step regression as Fama and MacBeth (1973). First, we estimate annual $\hat{\beta}$s with rolling subsamples for all 30 countries. For developed countries, we choose window size $h = 7$ years; while $h = 5$ years for emerging countries. We choose different window sizes because we consider developed countries as more stable economies compared to emerging countries. In this sense, investors may focus shorter periods for emerging countries. We try different window sizes $h = 5, 7, 10$ in the robustness subsection. Then we run regression of these $\hat{\beta}$s on the uncertainty index and its lags with both the following two different specifications:

$$\hat{\beta}_{i,t} = \alpha + \gamma(L) uncertainty_{i,t} + \epsilon_{i,t} \tag{1.11}$$

$$\hat{\beta}_{i,t} = \alpha_i + \gamma(L) uncertainty_{i,t} + \epsilon_{i,t} \tag{1.12}$$

where $\hat{\beta}_{i,t}$ are estimated from regression (1.6) with rolling window h for country i during time $t-h$ and t; $unvertainty_{i,t}$ are dummies and equals one if there exists uncertainty in country i at time t; $\gamma(L)$ is lag polynomial with order five in this case; α in (1.11) means that we pool all countries together; and α_i in (1.12) is the fixed effect when panel regressions are conducted. Other than pooled regressions, we also consider panel regressions with fixed effects because some countries might have stronger effects when the economy has no uncertainties. One example is JPY which has lowest $\hat{\beta} = -2.13$ for the whole sample period.

To differentiate the long-run uncertainty effects between developed and emerging countries, we estimate the above regression models with three groups: developed, emerging and all countries together. Table 1.6 shows the results of pooled two-step regressions. If we include all countries, the uncertainty dummy and its lags up to fourth order have positive significant impact on the value of $\hat{\beta}$s at 5% significance level. The average of all $\hat{\beta}$'s without any uncertainty is significantly negative. We can also see that the effects of long-run uncertainty is dominated by the developed countries.

The effects of long-run uncertainty on developed and emerging countries are different. First, developed countries have significantly negative $\hat{\beta}$ and emerging ones have insignificantly negative $\hat{\beta}$ if investors do not hold any long-run uncertainty. In addition, the average of β's over developed countries are more negative. Second, for developed countries, the uncertainty effect lasts for a long period and up to its fifth lag while the impact of the long-run uncertainty exists for up to two lags and decreases with higher lags for emerging ones. The intuition behind might be that investors have fewer long-run uncertainties about mature economies, which is true from our data sample where the long-run uncertainty about developed countries clusters around year 2008. However, once they have such uncertainty, it exists much longer than emerging countries. For the emerging countries with volatile economic situations, long-run uncertainty effect is shorter and decays faster.

Table 1.7 reports the estimation with fixed effect panel regression. The result for the long-run uncertainty effect is similar as pooled regression with some small changes on estimates. All individual fixed effects can be found in Table 1.8. Interestingly, only New Zealand has positive intercept for developed group. The rest 14 developed countries have significantly negative $\hat{\beta}$ if no long-run uncertainty exists. For emerging countries, Colombia, Korea and Poland have positive intercepts. The estimates are less significant compared to developed group.

In the traditional Fama regressions over the whole data sample, the stylized facts tell us that the FPP exists in developed but not emerging countries. The reason we may conclude from this subsection is that investors have more long-run uncertainties about emerging countries which have positive effects on $\hat{\beta}$. On average, the slope estimates tend to be positive. By contrast, even though long-run uncertainty has stronger positive effects on developed countries, investors have very few such uncertainty about performance of matured economies. Hence, on average, $\hat{\beta}$ tend to be negative and we observe the FPP over a longer sample period.

1.6 Robustness Test

In this subsection, we analyze the robustness of our estimation results to three different dimensions. First we use different rolling window sizes to estimate $\hat{\beta}$s in the first step of the two-step regressions. Second, we use different lags in the regression model (1.11), where the original $\gamma(L)$ is with order five there. Last, we construct the long-run uncertainty index by real consumption growth instead of the real GDP growth data.

1.6.1 Different Window Sizes and Index Lags

In previous subsection, we consider the rolling window to be $h = 7$ for developed countries and $h = 5$ for emerging countries. The uncertainty index lags are fixed at order of 5. Here we consider different combination of different window sizes ($h = 5, 7, 10$) and lags up to the rolling window sizes.

Table 1.9 to 1.11 report the results for rolling window size $h = 5, 7, 10$, respectively. For each h, different order of lags are estimated. We have similar results as before. First, the average of βs without long-run uncertainty is negative no matter what window size and lags are used. Second, all the effects of long-run uncertainty and its lags are positive up to certain lags. Third, we also notice that as we increase the window size, the effects for all different lags are decreasing.

1.6.2 Consumption Growth

Instead of defining the long-run uncertainty index from real GDP growth, we instead construct the new long-run uncertainty index with real consumption growth to better match the theoretical part in Chapter 2, where we model the real consumption growth as exogenous process. Table 1.12 lists the new long-run uncertainty index for all 30 countries. We observe less long-run uncertainties in this case. It is intuitive because consumption growth is smoothed and so less volatile than GDP growth.

We repeat the estimation in section 1.5.4 with the new long-run uncertainty index and Table 1.13 reports the result. Overall, effects from these two different long-run uncertainty indexes are similar. When we use real consumption growth, the uncertainty effects are smaller than estimation from GDP uncertainty. Moreover, the significance level becomes weaker.

1.7 Conclusion

This chapter studies empirical behaviors of the FPP during different subsamples instead of an average effect over the whole sample period as what is typically done in the literature. We find the estimated slope coefficients $\hat{\beta}$s from the Fama regressions vary considerably from period to period. The signs could be both positive and negative. Our goal is to find the reason which drives the variation of the slope estimates. We establish a link between the long-run uncertainty and variation of the slope estimates, and find that the long-run uncertainty index and its lags contribute to the positiveness of $\hat{\beta}$s.

The whole story is built based on Li and Tornell (2015) where they show that investors distorted belief about the importance of transitory and persistent interest rate differential shocks can determine the exchange rate in equilibrium. The subjective UIP still holds in their model. We map their model to the fundamental variables and empirically support their finding.

One important measure in our paper is the definition of the long-run uncertainty. We assume a simple and reasonable rule where investors worry about the long-run economic performance if a tail surprise happens. This can capture some aspects of difference between transitory and persistent shocks. It can be generalized to other complicated rule according to more theoretical work. Our work provides another way to look at all countries within a similar analysis framework.

1.8 Appendix

1.8.1 Tables and Figures

Table 1.1: Summary statistics: developed countries

			Δs_{t+1}		$i_t - i_t^*$	
Country	Currency	Entry Date	Mean	S.E	Mean	S.E.
Australia	AUD	7/31/1976	-0.72	11.63	-0.22	0.81
Belgium	BEF	10/31/1989	1.10	10.49	-0.03	0.62
Canada	CAD	1/31/1975	-0.16	6.75	-0.06	0.42
Denmark	DKK	1/31/1980	0.00	11.01	-0.08	0.80
France	FRF	1/31/1975	-0.26	10.85	-0.09	0.83
Germany	DEM	1/31/1975	1.25	11.15	0.10	0.74
Italy	ITL	6/30/1978	-1.40	10.84	-0.21	1.03
Japan	JPY	8/31/1978	1.76	11.56	0.23	0.73
Netherlands	NLG	1/31/1975	1.05	11.07	0.07	0.71
New Zealand	NZD	1/31/1985	2.12	12.23	-0.32	0.99
Norway	NOK	1/31/1979	-0.49	10.62	-0.16	0.87
Spain	ESP	1/31/1975	-1.95	10.93	-0.23	1.19
Sweden	SEK	1/31/1975	-1.29	11.16	-0.10	0.87
Switzerland	CHF	1/31/1975	2.62	12.16	0.22	0.88
U.K.	GBP	1/31/1975	-0.84	10.31	-0.15	0.67

Notes: The table presents summary statistics of monthly observations of exchange rate changes and interest rate differentials. Numbers are in percentage points. Means and standard deviations are annualized by multiplying the variables by 12 and $\sqrt{12}$. Entry date marks the year and month in which both exchange rate and interest rate date become available. The last sample date for all currencies is June 2014 and the choice of sample periods is mainly restricted by monthly interest rate data availability.

Table 1.2: Summary statistics: emerging countries

			Δs_{t+1}		$i_t - i_t^*$	
Country	Currency	Entry Date	Mean	S.E	Mean	S.E.
Argentina	ARS	1/31/1982	-56.18	49.38	-2.90	15.93
Brazil	BRL	1/31/1982	-75.76	38.44	-6.57	36.78
Chile	CLP	1/31/1982	-8.18	11.97	-0.74	2.88
Colombia	COP	1/31/1982	-10.63	9.55	-1.15	2.49
Czech Republic	CZK	4/30/1992	1.68	11.91	-0.17	1.01
India	IND	1/31/1993	-3.37	6.93	-0.39	0.56
Indonesia	IDR	1/31/1982	-8.97	21.56	-0.70	1.82
Korea	KRW	1/31/1982	-1.10	12.11	-0.28	1.07
Malaysia	MYR	1/31/1990	-0.71	7.19	-0.05	0.57
Mexico	MXN	1/31/1982	-19.08	25.21	-1.52	4.90
Philippines	PHP	1/31/1984	-3.63	12.10	-0.49	1.47
Poland	PLN	1/31/1983	-18.82	25.60	-0.77	3.73
South Africa	ZAF	1/31/1982	-7.38	15.84	-0.52	0.76
Thailand	THB	1/31/1992	-1.09	11.03	-0.21	1.13
Turkey	TRY	1/31/1982	-29.65	15.90	-2.79	6.13

Notes: The table presents summary statistics of monthly observations of exchange rate changes and interest rate differentials. Numbers are in percentage points. Means and standard deviations are annualized by multiplying the variables by 12 and $\sqrt{12}$. Entry date marks the year and month in which both exchange rate and interest rate date become available. The last sample date for all currencies is June 2014 and the choice of sample periods is mainly restricted by monthly interest rate data availability.

Table 1.3: UIP redux: developed countries

$s_{t+1} - s_t = \alpha + \beta(i_t - i_t^*) + \epsilon_{t+1}$ for 1975.1-2014.6

Country		α	$se(\alpha)$	$t_{\alpha=0}$	β	$se(\beta)$	$t_{\beta=0}$	$t_{\beta=1}$	N
Australia	AUD	-0.14	0.17	0.82	-0.36	0.53	0.68	**2.57*****	455
Belgium	BEF	0.07	0.18	0.40	-0.78	1.00	0.78	**1.78***	296
Canada	CAD	-0.08	0.11	0.71	-1.03	0.63	1.63	**3.22*****	473
Denmark	DKK	-0.07	0.17	0.40	-0.89	0.70	1.26	**2.68*****	413
France	FRF	-0.04	0.16	0.27	-0.25	0.75	0.33	**1.67***	473
Germany	DEM	0.16	0.17	0.98	-0.63	0.81	0.77	**2.01****	473
Italy	ITL	0.01	0.19	0.04	0.58	0.60	0.98	0.70	432
Japan	JPY	0.63	0.23	2.78	-2.13	0.70	**3.05*****	**4.48*****	430
Netherlands	NLG	0.21	0.16	1.27	-1.80	0.81	**2.22****	**3.45*****	473
New Zealand	NZD	-0.24	0.26	0.93	-1.29	0.59	**2.19****	**3.88*****	353
Norway	NOK	-0.06	0.19	0.30	-0.10	0.88	0.11	1.24	425
Spain	ESP	-0.16	0.17	0.94	0.01	0.47	0.02	**2.13****	473
Sweden	SEK	-0.07	0.16	0.44	0.39	1.06	0.37	0.58	473
Switzerland	CHF	0.48	0.22	2.16	-1.18	0.73	1.62	**3.00*****	473
U.K.	GBP	-0.26	0.18	1.38	-1.20	0.94	1.27	**2.33****	473

Notes: The table reports country-level results from regression (1.6). Δs_{t+1} is the change of log exchange rate and $i_t - i_t^*$ is the interest differential between the U.S. and other foreign countries. Standard errors in parentheses are Newey-West heteroscedasticity-consistent standard errors computed with an optimal number of lags according to Newey and West (1994). All variables are in percentage points. ***, ** and * represent the 1%, 5% and 10% significance levels, respectively. $t_\beta = 1$ denotes the t-statistics for null hypothesis $\beta = 1$. N refers to the number of observations in the regression. There are $N = 473$ observations from January 1975 to June 2014 (474 months). For those which have shorter sample periods, the starting date can be found in Table 1.1.

Table 1.4: UIP redux: emerging countries

$$s_{t+1} - s_t = \alpha + \beta(i_t - i_t^*) + \epsilon_{t+1} \text{ for 1982.1-2014.6}$$

Country		α	$se(\alpha)$	$t_{\alpha=0}$	β	$se(\beta)$	$t_{\beta=0}$	$t_{\beta=1}$	N
Argentina	ARS	0.16	0.72	0.22	1.67	0.40	**4.19*****	**1.67***	389
Brazil	BRL	-1.65	0.90	1.84	0.71	0.11	**6.56*****	**2.71*****	389
Chile	CLP	-0.04	0.22	0.20	0.86	0.25	**3.41*****	0.55	389
Colombia	COP	0.40	0.34	1.20	1.12	0.21	**5.37*****	0.57	389
Czech Republic	CZK	0.27	0.25	1.10	0.80	0.57	1.40	0.36	266
India	IND	0.43	0.33	1.28	1.81	0.90	**2.01****	0.90	257
Indonesia	IDR	-1.07	0.60	1.78	-0.47	1.12	0.42	1.31	389
Korea	KRW	-0.10	0.26	0.38	-0.02	0.64	0.04	1.59	389
Malaysia	MYR	-0.03	0.14	0.18	0.64	0.53	1.21	0.67	293
Mexico	MXN	-0.05	0.38	0.14	1.01	0.26	**3.89*****	0.04	389
Philippines	PHP	0.09	0.26	0.34	0.79	0.61	1.29	0.34	365
Poland	PLN	-0.11	0.57	0.20	1.88	1.06	**1.77***	0.83	377
South Africa	ZAF	-1.08	0.55	1.97	-0.91	1.01	0.90	**1.89***	389
Thailand	THB	0.16	0.25	0.64	1.20	1.67	0.72	0.12	269
Turkey	TRY	0.12	0.38	0.33	0.93	0.14	**6.77*****	0.52	389

Notes: The table reports country-level results from regression (1.6). Δs_{t+1} is the change of log exchange rate and $i_t - i_t^*$ is the interest differential between the U.S. and other foreign countries. Standard errors in parentheses are Newey-West heteroscedasticity-consistent standard errors computed with an optimal number of lags according to Newey and West (1994). All variables are in percentage points. ***, ** and * represent the 1%, 5% and 10% significance levels, respectively. $t_\beta = 1$ denotes the t-statistics for null hypothesis $\beta = 1$. N refers to the number of observations in the regression. There are $N = 473$ observations from January 1975 to June 2014 (474 months). For those which have shorter sample periods, the starting date can be found in Table 1.2.

Table 1.5: Four equal subsamples: developed countries

	Panel I: 75.1-84.12		Panel II: 85.1-94.12		Panel III: 95.1-04.12		Panel IV: 05.1-14.6	
	α	β	α	β	α	β	α	β
Australia	-0.29	1.36*	-0.46	-1.01	-0.57**	-5.57**	0.89	3.04
Belgium			0.94	2.31	0.43	-6.18***	0.00	4.08*
Canada	-0.33***	-1.38*	-0.44	-2.42**	0.14	-4.27***	0.18	2.95
Denmark	-1.47***	-2.71***	0.76*	0.93	0.25	-5.62***	0.04	3.09
France	-0.76**	-0.50	0.65	0.73	0.32	-5.17***	0.00	4.08*
Germany	0.24	-1.64	0.60*	1.55	0.44	-5.69***	0.00	4.08*
Italy	-1.43**	-0.80	1.30	2.89	-0.04	-2.33**	0.00	4.08*
Japan	0.88	-2.99**	1.01***	-2.07	0.89	-2.88	-0.12	1.03
Netherlands	0.39	-3.42***	0.65*	1.29	0.46*	-5.67***	0.00	4.08*
New Zealand			-0.77**	-1.89***	-0.81*	-4.85***	4.27***	17.01**
Norway	-0.80***	0.02	1.23*	2.53	-0.18	-3.08**	0.26	3.13
Spain	-1.38***	-1.02**	1.34*	2.27	-0.01	-2.94*	0.00	4.08*
Sweden	-0.69***	1.40	1.58**	3.84*	0.03	-5.35***	0.07	3.21
Switzerland	1.12	-2.25	0.57	0.50	1.37***	-6.06***	0.19	0.69
U.K.	-0.87**	-1.75*	0.04	-0.74	-0.21	-3.63*	0.49**	11.75**

Table 1.6: The Long-run uncertainty index

Country	Uncertainty Source	Years
Developed		
Australia	GDP Growth	1983, 2009
Belgium	GDP Growth	1993, 2009
Canada	GDP Growth	1982, 2009
Denmark	GDP Growth	2009
France	GDP Growth	1993, 2008, 2009
Germany	GDP Growth	1993, 2003, 2009
Italy	GDP Growth	1993, 2008, 2009
Japan	GDP Growth	1992, 1993, 1998, 2008, 2009
Netherlands	GDP Growth	1981, 2002, 2009
New Zealand	GDP Growth	2008
Norway	GDP Growth	1981, 1982, 2008, 2009
Spain	GDP Growth	1977, 1993, 2008, 2009
Sweden	GDP Growth	1977, 1980, 1991, 1992, 2008, 2009
Switzerland	GDP Growth	2009
U.K.	GDP Growth	1991, 2008, 2009

Notes: The table reports the uncertainty index constructed following the rule in section 1.5.3. We choose window size $h = 10$. Uncertainty source indicates whether the uncertainty comes from the growth or inflation.

Table 1.5: The Long-run uncertainty index (continued)

Country	Uncertainty Source	Years
Emerging		
Argentina	GDP Growth	1976, 1978, 2002
	Inflation	1975, 1976, 1984, 1985, 1989, 2002, 2014
Brazil	GDP Growth	1977, 1978, 1981, 2009
	Inflation	1988, 1989, 1990
Chile	GDP Growth	1975, 1999, 2009
Colombia	GDP Growth	1975, 1998, 1999
	Inflation	1977
Czech Republic	GDP Growth	2009
India	GDP Growth	1979, 1991
	Inflation	1991, 2009, 2010
Indonesia	GDP Growth	1982, 1997, 1998
	Inflation	1998
Korea	GDP Growth	1980, 1992, 1998, 2009
	Inflation	1975
Malaysia	GDP Growth	1975, 1985, 1997, 1998, 2009
Mexico	GDP Growth	1982, 1983, 1995, 2009
	Inflation	1977, 1982, 1983, 1987
Philippines	GDP Growth	1984, 1985, 2009
	Inflation	1984
Poland	GDP Growth	2001, 2012
	Inflation	1981, 1982, 1989, 1990
South Africa	GDP Growth	1977, 2009
	Inflation	1975, 1985, 1986, 2008
Thailand	GDP Growth	1997, 1998, 2008, 2009
Turkey	GDP Growth	1979, 1980, 1994, 2009
	Inflation	1977, 1978, 1979, 1980, 1994

Table 1.6: Long-Run uncertainty and the FPP: pooled

$$\hat{\beta}_{i,t} = \alpha + \gamma(L)uncertainty_{i,t} + \epsilon_{i,t}$$

	All Countries	Developed	Emerging
Intercept	-2.034***	-2.810***	-0.374
	(0.283)	(0.316)	(0.460)
LR-uncertainty$_t$	2.430***	2.086***	2.561***
	(0.436)	(0.455)	(0.701)
LR-uncertainty$_{t-1}$	2.371***	3.027***	1.161**
	(0.387)	(0.496)	(0.577)
LR-uncertainty$_{t-2}$	2.713***	3.319***	1.401***
	(0.435)	(0.610)	(0.534)
LR-uncertainty$_{t-3}$	2.638***	3.607***	0.963
	(0.490)	(0.659)	(0.666)
LR-uncertainty$_{t-4}$	3.082***	4.697***	0.479
	(0.643)	(0.628)	(1.119)
LR-uncertainty$_{t-5}$	1.430*	3.211***	-1.472
	(0.776)	(0.920)	(1.077)
Adjusted R^2	0.172	0.298	0.058

Notes: The table reports the result of pooled two-step regressions. There are 15 developed and 15 emerging countries, respectively, where $\gamma(L)$ has order of five. Standard errors in parentheses are Newey-West heteroscedasticity-consistent standard errors computed with an optimal number of lags according to Newey and West (1994). ***, ** and * represents the 1%, 5% and 10% significance levels, respectively.

Table 1.7: Long-Run uncertainty and the FPP: panel

$$\hat{\beta}_{i,t} = \alpha_i + \gamma(L)uncertainty_{i,t} + \epsilon_{i,t}$$

	All Countries	Developed	Emerging
LR-uncertainty$_t$	2.764***	2.551***	2.860***
	(0.492)	(0.606)	(0.751)
LR-uncertainty$_{t-1}$	2.573***	3.381***	1.300*
	(0.491)	(0.611)	(0.736)
LR-uncertainty$_{t-2}$	3.023***	3.708***	1.701**
	(0.498)	(0.615)	(0.757)
LR-uncertainty$_{t-3}$	2.890***	3.962***	1.189
	(0.487)	(0.601)	(0.738)
LR-uncertainty$_{t-4}$	3.264***	4.978***	0.654
	(0.479)	(0.594)	(0.721)
LR-uncertainty$_{t-5}$	1.659***	3.590***	-1.391
	(0.577)	(0.710)	(0.877)
Adjusted R^2	0.179	0.335	0.037

Notes: The table reports the result of two-step panel regressions with fixed effect. There are 15 developed and 15 emerging countries, respectively, where $\gamma(L)$ has order of five. ***, ** and * represents the 1%, 5% and 10% significance levels, respectively. Individual fixed effect for column 1 can be found in Table 1.8.

Table 1.8: Panel fixed effects

	f.e.	s.e.	t-value	p-value		f.e.	s.e.	t-value	p-value
ARS	-0.74	0.92	-0.80	0.42	AUD	-2.14	0.64	-3.32	0.00***
BRL	-1.35	0.92	-1.46	0.15	BEF	-2.56	0.85	-3.03	0.00***
CLP	-0.13	0.96	-0.14	0.89	CAD	-2.77	0.62	-4.44	0.00***
COP	0.15	1.04	0.15	0.88	CHF	-3.30	0.62	-5.31	0.00***
CZK	-1.18	0.87	-1.36	0.17	DEM	-3.59	0.63	-5.70	0.00***
IDR	-3.12	0.81	-3.87	0.00***	DKK	-2.19	0.67	-3.25	0.00***
IND	-0.86	1.09	-0.79	0.43	ESP	-1.75	0.63	-2.77	0.01***
KRW	0.01	0.85	0.01	0.99	FRF	-2.15	0.63	-3.42	0.00***
MXN	-2.90	0.69	-4.19	0.00***	GBP	-2.03	0.63	-3.22	0.00***
MYR	-0.51	0.82	-0.63	0.53	ITL	-1.46	0.67	-2.16	0.03**
PHP	-1.70	0.99	-1.71	0.09*	JPY	-5.54	0.69	-7.99	0.00***
PLN	1.52	0.90	1.69	0.09*	NLG	-4.27	0.63	-6.79	0.00***
THB	-3.23	0.88	-3.66	0.00***	NOK	-1.33	0.67	-1.98	0.05*
TRY	-0.85	0.99	-0.85	0.39	NZD	1.01	0.74	1.36	0.17
ZAF	-7.63	1.05	-7.30	0.00***	SEK	-2.64	0.74	-3.57	0.00***

Note. This table reports detailed fixed effects for panel regressions in the first column "All Countries" in Table 1.7. That panel regression includes all 30 countries in our data sample. First column f.e. is short for fixed effect. ***, ** and * represents the 1%, 5% and 10% significance levels, respectively.

Table 1.9: Different lags when $h = 5$

$$\hat{\beta}_{i,t} = \alpha + \gamma(L) uncertainty_{i,t} + \epsilon_{i,t}$$

	All Countries	Developed	Emerging	All Countries	Developed	Emerging
Intercept	-1.871***	-2.501***	-0.468	-2.467***	-3.373***	-0.374
	(0.340)	(0.430)	(0.445)	(0.335)	(0.390)	(0.460)
LR-uncertainty$_t$	3.331***	3.529***	2.612***	3.738***	4.207***	2.561***
	(0.512)	(0.678)	(0.715)	(0.515)	(0.684)	(0.701)
LR-uncertainty$_{t-1}$	2.498***	3.176***	1.180**	2.839***	3.698***	1.161**
	(0.514)	(0.744)	(0.565)	(0.538)	(0.773)	(0.577)
LR-uncertainty$_{t-2}$	3.2015***	4.093***	1.438***	3.790***	4.848***	1.401***
	(0.493)	(0.706)	(0.484)	(0.519)	(0.726)	(0.534)
LR-uncertainty$_{t-3}$	4.389***	6.290***	1.186*	3.738***	5.358***	0.963
	(0.742)	(1.073)	(0.626)	(0.724)	(1.034)	(0.666)
LR-uncertainty$_{t-4}$				4.187***	6.392***	0.479
				(0.920)	(1.191)	(1.119)
LR-uncertainty$_{t-5}$				1.290	2.472	-1.472
				(1.028)	(1.414)	(1.077)
Adjusted R^2	0.130	0.168	0.056	0.189	0.280	0.058

Table 1.10: Different lags when $h = 7$

$$\hat{\beta}_{i,t} = \alpha + \gamma(L)uncertainty_{i,t} + \epsilon_{i,t}$$

	All Countries	All Countries	All Countries
Intercept	-1.479***	-2.077***	-2.079***
	(0.281)	(0.282)	(0.299)
LR-uncertainty$_t$	1.379***	1.867***	1.848***
	(0.347)	(0.352)	(0.358)
LR-uncertainty$_{t-1}$	1.989***	2.286***	2.299***
	(0.347)	(0.364)	(0.367)
LR-uncertainty$_{t-2}$	1.886***	2.513***	2.514***
	(0.401)	(0.427)	(0.431)
LR-uncertainty$_{t-3}$	3.099***	2.682***	2.684***
	(0.481)	(0.455)	(0.459)
LR-uncertainty$_{t-4}$		3.540***	3.556***
		(0.491)	(0.497)
LR-uncertainty$_{t-5}$		2.284***	2.206***
		(0.626)	(0.631)
LR-uncertainty$_{t-6}$			0.432
			(0.437)
LR-uncertainty$_{t-7}$			-0.318
			(0.453)
Adjusted R^2	0.114	0.222	0.221

Table 1.11: Different lags when $h = 10$

$$\hat{\beta}_{i,t} = \alpha + \gamma(L)uncertainty_{i,t} + \epsilon_{i,t}$$

	All Countries	All Countries	All Countries	All Countries
Intercept	-1.098***	-1.450***	-1.578***	-1.708***
	(0.234)	(0.239)	(0.251)	(0.274)
LR-uncertainty$_t$	-0.094	0.182	0.252	0.341
	(0.329)	(0.336)	(0.338)	(0.345)
LR-uncertainty$_{t-1}$	0.549**	0.741***	0.779***	0.848***
	(0.269)	(0.285)	(0.288)	(0.291)
LR-uncertainty$_{t-2}$	0.645***	1.001***	1.081***	1.131***
	(0.277)	(0.293)	(0.297)	(0.303)
LR-uncertainty$_{t-3}$	1.544***	1.334***	1.415***	1.465***
	(0.329)	(0.318)	(0.325)	(0.332)
LR-uncertainty$_{t-4}$		1.916***	2.026***	2.096***
		(0.328)	(0.335)	(0.343)
LR-uncertainty$_{t-5}$		1.417***	1.405***	1.472***
		(0.522)	(0.523)	(0.531)
LR-uncertainty$_{t-6}$			0.683**	0.815**
			(0.350)	(0.386)
LR-uncertainty$_{t-7}$			1.009***	0.906***
			(0.353)	(0.331)
LR-uncertainty$_{t-8}$				0.935***
				(0.334)
LR-uncertainty$_{t-9}$				0.543
				(0.409)
Adjusted R^2	0.038	0.114	0.123	0.131

Table 1.12: The Long-run uncertainty index with consumption growth

Country	Years	Country	Years
Developed		Emerging	
Australia	2009	Argentina	1976, 2002, 2014
Belgium	1993	Brazil	1975, 1981
Canada	1982, 2009	Chile	1975, 1999, 2009
Denmark	1980, 2009	Colombia	1993, 1998,1999
France	1980, 2008	Czech Republic	2009, 2011, 2012
Germany	1982, 1993, 2002	India	2004
Italy	1993, 2008, 2009,2012	Indonesia	2000
Japan	1991, 1992, 1993, 1998, 2008	Korea	1980, 1998
Netherlands	1980, 1981, 1982, 2003, 2012	Malaysia	1985, 1986, 1998, 2009
New Zealand	1988, 2008, 2009	Mexico	1982, 1983, 1995, 2009
Norway	1981, 1982, 2008, 2009	Philippines	1983
Spain	1993, 2008, 2009, 2012	Poland	None
Sweden	1980, 1993	South Afric	1977, 1983, 2009
Switzerland	1992, 1993	Thailand	1997, 1998, 2009
U.K.	1991, 2005, 2008, 2009	Turkey	2001

Notes: The table reports the uncertainty index constructed following the rule in section 1.5.3. We choose window size $h = 10$. Instead of using real GDP growth, we construct the index with real consumption growth.

Table 1.13: Long-run uncertainty with consumption and the FPP

	Pooled Rregression			Panel Regression		
	All Countries	Developed	Emerging	All Countries	Developed	Emerging
Intercept	-1.382***	-2.172***	0.233			
	(0.322)	(0.353)	(0.547)			
LR-uncertainty$_t$	1.433***	1.060*	1.963**	1.915***	1.446**	2.580***
	(0.548)	(0.635)	(0.872)	(0.533)	(0.678)	(0.824)
LR-uncertainty$_{t-1}$	1.307***	1.873***	0.141	1.670***	2.161***	0.599
	(0.514)	(0.689)	(0.521)	(.527)	(0.679)	(0.805)
LR-uncertainty$_{t-2}$	1.373**	1.739**	0.183	1.784***	2.095***	0.642
	(0.614)	(0.851)	(0.622)	(0.523)	(0.706)	(0.820)
LR-uncertainty$_{t-3}$	1.373**	2.031***	-0.087	1.682***	2.250***	0.289
	(0.617)	(0.857)	(0.652)	(0.528)	(0.681)	(0.806)
LR-uncertainty$_{t-4}$	1.743**	2.683***	-0.351	1.975***	2.785***	0.019
	(0.748)	(0.846)	(1.192)	(0.521)	(0.670)	(0.792)
LR-uncertainty$_{t-5}$	0.634	1.888*	-2.312	1.123*	2.405***	-2.037**
	(0.782)	(0.961)	(1.086)	(0.577)	(0.721)	(0.920)
Adjust R^2	0.056	0.125	0.022	0.052	0.133	0.015

Figure 1.4: Rolling Regressions for Developed Countries with $h = 7$

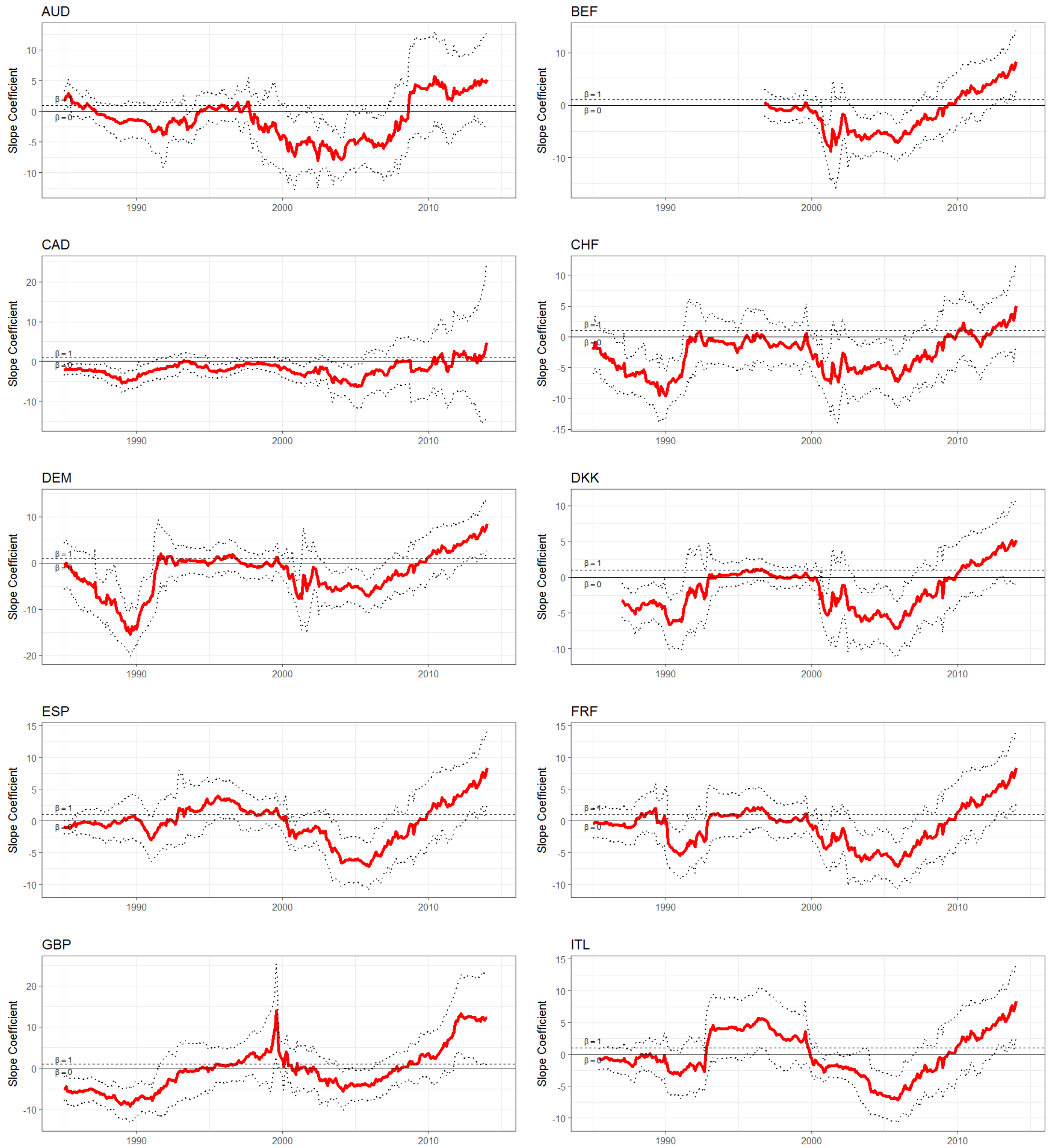
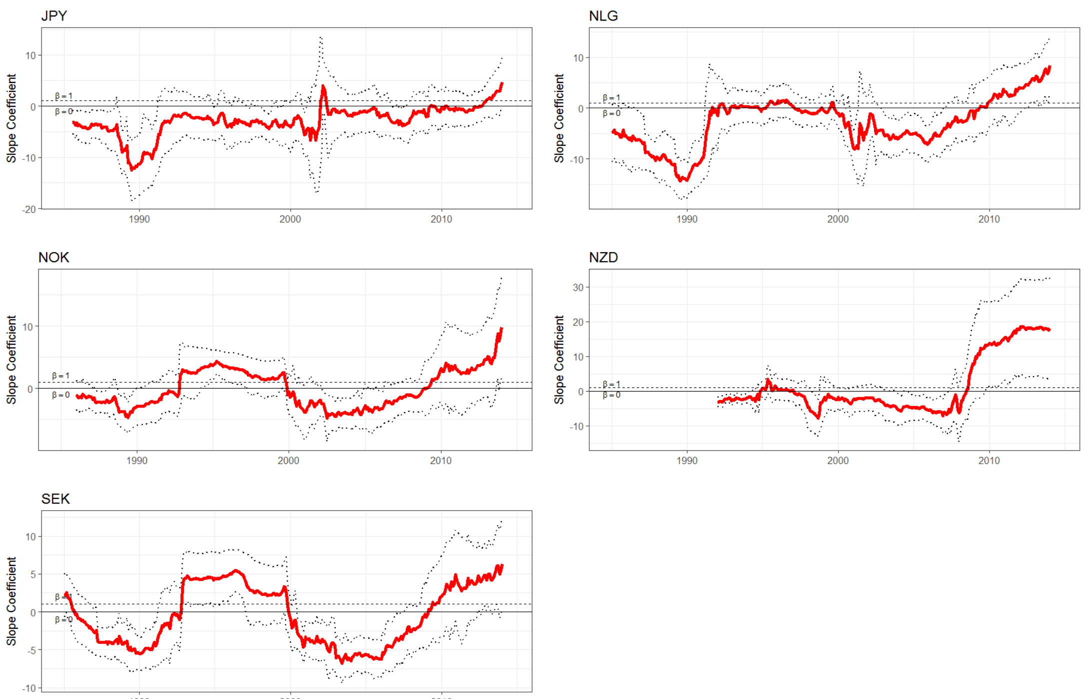


Figure 1.3: Rolling Regressions for Developed Countries with $h = 7$ (Continued)



Note: The figure shows the results of all developed countries when $h = 7$. Each graph plots the series of slope estimates and corresponding 95% confidence interval of rolling window estimations. The thick red line represents slope estimates and the black dotted line indicates the 95% confidence interval. The horizontal dashed line represents the null hypothesis $\beta = 1$. The date 1985 means that the regression sample period is from January 1976 to December 1982.

Figure 1.4: Rolling Regressions for Developed Countries with $h = 10$

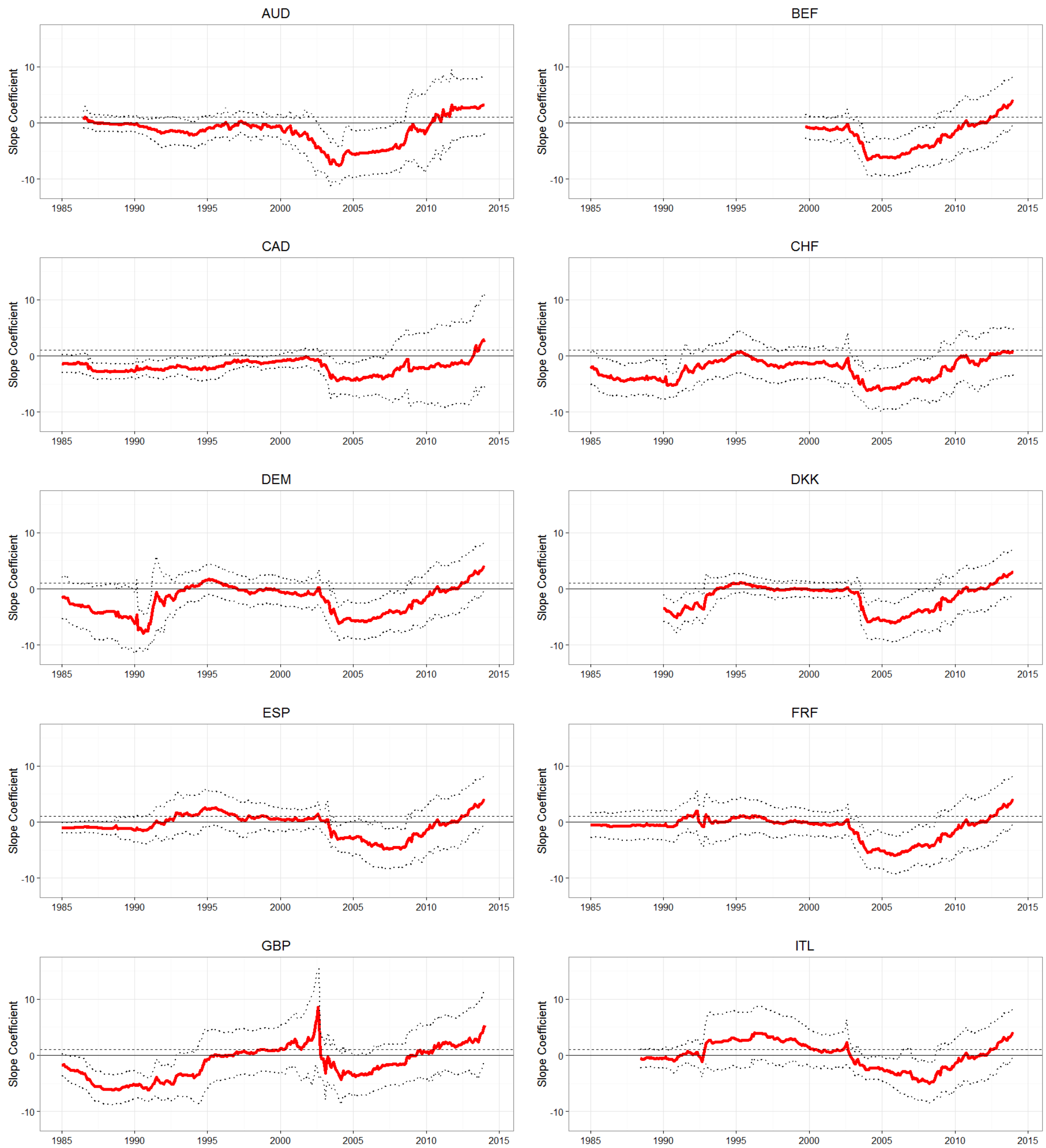


Figure 1.3: Rolling Regressions for Developed Countries with $h = 10$ (Continued)

Note: The figure shows the results of all developed countries when $h = 10$. Each graph plots the series of slope estimates and corresponding 95% confidence interval of rolling window estimations. The thick red line represents slope estimates and the black dotted line indicates the 95% confidence interval. The horizontal dashed line represents the null hypothesis $\beta = 1$. The date 1985 means that the regression sample period is from January 1976 to December 1985.

Figure 1.4: Rolling Regressions for Emerging Countries with $h = 5$

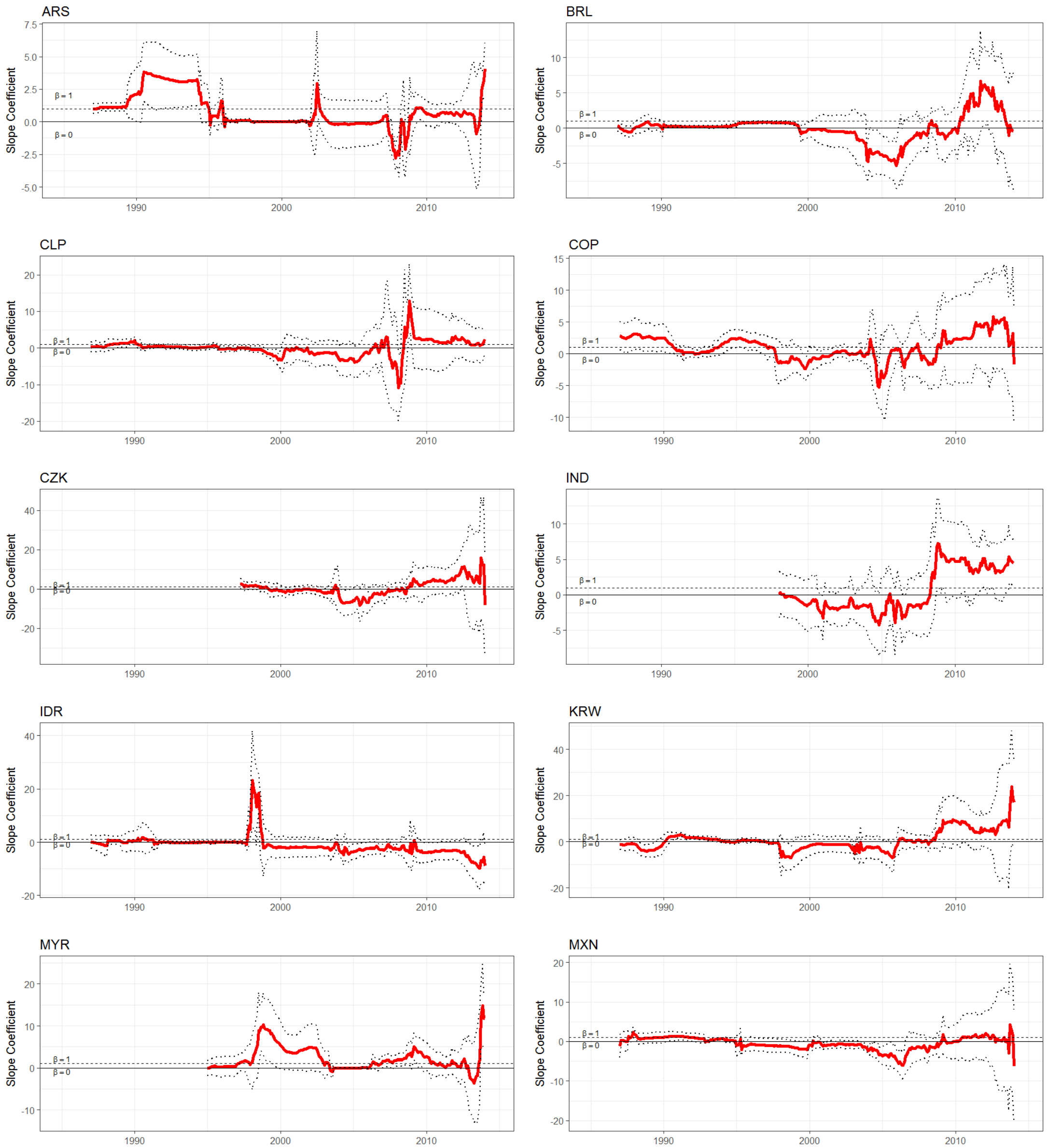
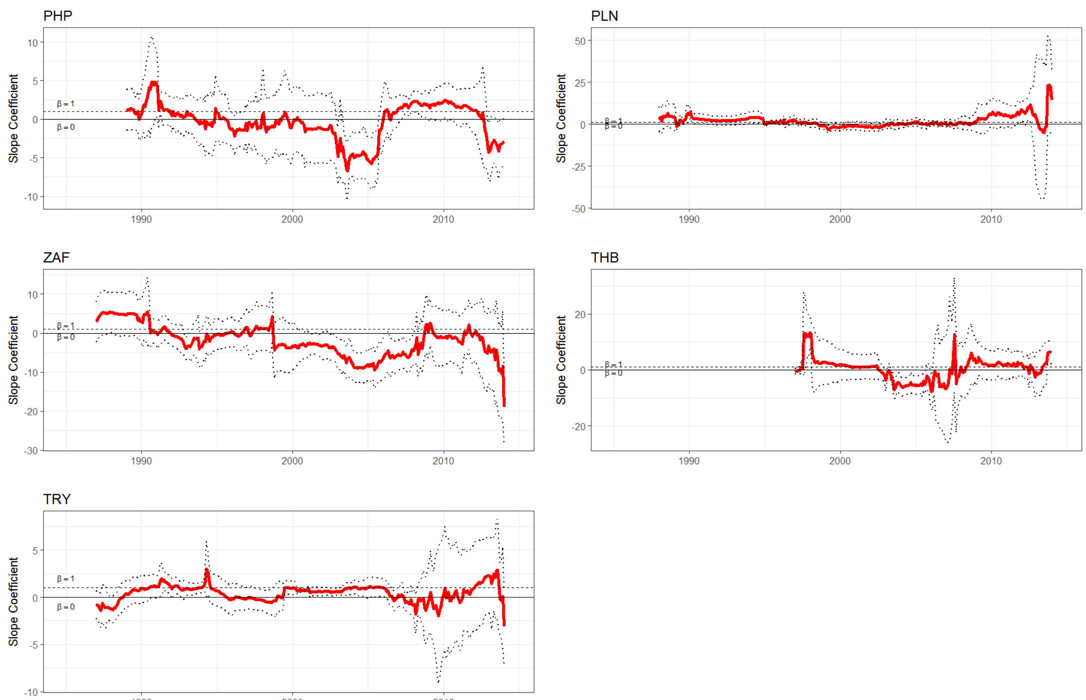


Figure 1.3: Rolling Regressions for Emerging Countries with $h = 5$ (Continued)



Note: The figure shows the results of all emerging countries when $h = 5$. Each graph plots the series of slope estimates and corresponding 95% confidence interval of rolling window estimations. The thick red line represents slope estimates and the black dotted line indicates the 95% confidence interval. The horizontal dashed line represents the null hypothesis $\beta = 1$. The date 1955 means that the regression sample period is from January 1991 to December 1995.

2. The Forward Premium Puzzle and Robust Control: Theory

2.1 Introduction

In Chapter 1, when there is a persistent shock either from real GDP growth drop or inflation hike in an economy, investors feel uncertain about the long-run performance of the economy. This long-run uncertainty and its lags contribute to the positiveness of estimated slope coefficients from the Fama regression and the effect lasts for up to several years depending on whether the country is developed or not. With this uncertainty, empirical results show us that the slope estimates from Fama regressions tend to be positive so that the FPP disappears.

This chapter provides a potential theoretical framework to understand the empirical facts described in Chapter 1 based on Li and Tornell (2015). They show that the robustness against model misspecification can generate both positive and negative Fama slope coefficients, depending on investors' beliefs about the relative importance of transitory and persistent interest rate shocks. If investors fear model misspecification in the state equation of the interest rate differential process, i.e., $z_{t+1} = \rho z_t + w_{t+1}$, positive β can be generated from the equilibrium, which means there is no FPP. On the other hand, the FPP exists if the misspecification lies in the observational equation i.e., $y_t = z_t + v_t$. But they miss one step linking the economic fundamentals to the assumed interest rate differential model. We fill the gap using the long-run risk model with two variables: real consumption growth and inflation. We map the persistent interest rate shocks to long-run shocks to either consumption growth or inflation, which matches the long-run uncertainty defined in Chapter 1. We then qualitatively explain the empirical facts of time-varying slope estimates.

Specifically, we use the long-run risk model following Bansal and Yaron (2004) and Bansal and Shaliastovich (2013) by modeling consumption growths as sum of a transitory shock and a persistent component in both domestic and foreign endowment economies. Inflation processes are also modeled in a similar way. The reason why we choose the long-run risk model is because the interest rate differential model used by Li and Tornell (2015) has similar state space form, and then we can map their model structure to economic underlying factors including consumption growth and inflation.

Suppose investors fear model misspecification in the structured persistent component in Li and Tornell (2015), we show that the fear comes from the uncertainty about the long-run part of either consumption growth or inflation process. To identify this type of uncertainty from data empirically, we employ a simple rule such that investors hold this long-run uncertainty beliefs when they observe a huge drop in consumption growth or inflation spike, which is captured by the long-run uncertainty index in Chapter 1. Hence, we map the state uncertainty in the interest rate differential model to the long-run uncertainty about the economic variables. The link is built within standard pricing kernel framework.

This chapter is organized as following. Section 2 reviews the literature of theoretical models trying to explain the FPP. Section 3 explains the key ingredients in Li and Tornell (2015). Section 4 shows the model linking the fundamental economic factors to the state space form model of interest rate differential. Section 5 concludes.

2.2 Literature

There is an extensive literature in economics and finance trying to explain the failure of the UIP condition. Two explanations are common in literature: time-varying risk premium and expectational errors. First and the most direct, many authors argue that the failure is because of risk premium from the exchange rates. Three types of general equilibrium models are extended from the closed economy equity premium puzzle: habit model (Abel (1990)), long-run risk model (Bansal and Yaron (2004)) and disaster model (Barro (2006)).

Following Backus, Foresi and Telmer (2001), in the currency market, the key is to solve the Euler equation for both domestic and foreign representative households, $S_{t+1}/S_t =$

M^*_{t+1}/M_{t+1} , where S_t is the nominal (real) exchange rates and M_{t+1} and M_{t+1} are the nominal (real) stochastic discount factors for foreign and domestic investors, respectively. When pricing kernels are conditionally log normal, risk premium boils down to differences in conditional variances of SDF. Verdelhan (2010) introduces habits into utility function following Campbell and Cochrane (1999). In this model, investors are more risk-averse when the consumption level is close to the external habit level, and so conditional variance of the SDF is large in bad times. To account for the UIP puzzle in this framework, real interest rates must be pro-cyclical, which arises endogenously in the model. Stathopoulos (2012) uses both external habits and home-biased preferences to reconcile the high degree of international risk sharing puzzle. Using a similar idea in Bansal and Yaron (2004), Bansal and Shaliastovich (2013) model the real consumption growth with a persistent long-run expected growth component, and use Epstein-Zin preference. In their models, short term shocks have different variance from the long-run component. The channel here is that periods of high volatility are associated with expected depreciation of the currency and low interest rate differentials. Colacito and Croce (2011, 2013) also apply the long-run risks models. Farhi and Gabaix (2016) develop a model in which the forward premium arises because certain countries are more exposed to rare global fundamental disaster events. Their model is calibrated to also match skewness patterns obtained from foreign exchange option prices. Other papers includes Lustig and Verdelhan (2007).

On the other hand, Froot and Thaler (1990) decompose biased forward discount into Gourinchas and Tornell (2004) interpret the bias as evidence of expectational errors. They show that both the forward premium puzzle and delayed-overshooting puzzle arise from a systematic distortion in investors beliefs about the interest rate process. Investors underreact to the persistence of short-run nominal interest rate changes. So the average change in exchange rates is smaller than the change in interest differential. Li and Tornell (2015) provide microfoundation of investors' distorted belief that leads them to under-react to news. Burnside, Han, Hirshleifer and Wang (2011) explain the FPP based upon investor overconfidence about future inflation. Yu (2013) proposes a sentiment-based model of the exchange rate to understand the FPP.

Besides the two main channels, there are also some other research trying to explain the FPP. Bacchetta and van Wincoop (2006) explain the puzzle from a microstructure approach where investors infrequently revise their portfolios.

The literature in finance focuses on explaining the carry trade profitability due to the failure of the UIP condition. Brunnermeier, Nagel and Pedersen (2008) provide a link between the currency carry and currency crash risk. However, Jurek (2014) shows that the crash risk account for at most one third of the excess returns from carry trades by constructing a crash-hedged portfolio with options. Burnside, Eichenbaum, Kleshchelski and Rebelo (2011) argue the on average positive payoffs from carry trades reflect a peso problem which means the effects caused by low-probability events that have not happened.

2.3 Li and Tornell (2015) Review

Here we provide a short review of this paper and present the key results which are related to our work. Notice that the results shown in this subsection do not include all necessary conditions in their paper.

In their model, the interest rate differential process consists of a persistent component z_t and a transitory component v_t,

$$\begin{aligned} i_t - i_t^* \equiv y_t &= z_t + v_t, \\ z_t &= \alpha z_{t-1} + w_t, \end{aligned} \tag{2.1}$$

where $v_t \sim N(0, \sigma_v^2)$ and $w_t \sim N(0, \sigma_w^2)$.

The investor is endowed with a baseline model which is the data generating process under probability measure θ'. But she fears model misspecification in the baseline model and behaves under her robustly distorted probability measure θ.

From the perspective of the investor, the equilibrium exchange rate can be presented in a recursive form as,

$$
\begin{aligned}
s_t =& E_t^{\theta} s_{t+1} - (i_t - i_t^*) \\
=& \bar{s}_t - (i_t - i_t^*) - \sum_{j=1}^{\infty} E_t^{\theta}(i_{t+j} - i_{t+j}^*)
\end{aligned}
$$

where $\bar{s}$ is the long run exchange rate, and E_t^{θ} denote the expectation under robust belief θ.

The investor forecasts future interest rate differentials using the Kalman filter under her own distorted belief. The Kalman gain plays an important role in determining the forecasting weight of the new information y_t,

$$
E_t^{\theta_t}(z_t) = \hat{z}_t^{\theta_t} = (1 - k_t^{\theta_t})\alpha \hat{z}_{t-1}^{\theta_{t-1}} + k_t^{\theta_t} y_t,
$$

where $k_t^{\theta_t}$ is the Kalman gain under belief θ.

The FPP arises in equilibrium if the investor underreacts to news, i.e., $k_t^{\theta_t} < k^{baseline(\theta')}$. On the other hand, if $k_t^{\theta_t} \geq k^{baseline(\theta')}$, the investor's robust forecasts are more sensitive to news and we cannot observe the FPP.

It is not trivial to generate small k and large k compared to the baseline model $k^{\theta'}$ if they do not specify the uncertainty structure. It will become a nonparametric distance between two probability measures θ' and θ. This paper specifies two types of structured model misspecification – uncertainty in the observation equation and uncertainty in the persistent component – to generate the above two different Kalman gains, respectively.

Consider the Kalman gain under baseline model,

$$
k^{\theta'} = \frac{\alpha^2 \xi^* + \sigma_w^2}{\alpha^2 \xi^* + \sigma_w^2 + \sigma_v^2},
$$

different types of uncertainties affect either σ_w^2 or σ_v^2 to have a robustly distorted Kalman gain.

If the investor fears misspecification in observation equation $y_t = z_t + v_t, v_t \overset{\theta'}{\sim} N(0, \sigma_v^2)$, then under her own distorted belief, the robust variance $\tilde{\sigma}_v^2$ is distorted upwards,

$$
v_t \overset{\theta}{\sim} N(0, \tilde{\sigma}_v^2) \text{ with } \tilde{\sigma}_v^2 > \sigma_v^2,
$$

which generates $k_t^{\theta_t} < k^{baseline(\theta')}$ because σ_v^2 appears in the denominator of the Kalman gain. Hence, the FPP arises and $\beta < 0$.

If the investor fears misspecification in persistent component $z_t = \alpha z_{t-1} + w_t, w_t \overset{\theta'}{\sim} N\left(0, \sigma_w^2\right)$, then the robust variance $\tilde{\sigma}_w^2$ is distorted downward,

$$w_t \overset{\theta}{\sim} N\left(0, \tilde{\sigma}_w^2\right) \text{ with } \tilde{\sigma}_w^2 > \sigma_w^2$$

which generates $k_t^{\theta_t} > k^{baseline(\theta')}$ and then $\beta > 1$, where we do not observe the FPP.

2.4 Model

In this model, we use a standard one-agent endowment economy to link pricing kernel, fundamental economic variables and interest rate differential. We can generate the baseline interest rate differential model in Li and Tornell (2015) from the standard pricing model in the asset pricing literature. Uncertainties about model misspecification arise from the investor's perspective about the long-run performance of economic underlying variables.

Given a rule of the investor about the economy, the investor either overreact or underreact to news, which generates positive and negative β's, respectively.

2.4.1 Pricing Kernels in Currency Market

Suppose M_{t+1} is the domestic stochastic discount factor or pricing kernel, for a gross return of any tradeable asset R_{t+1}, we have the basic pricing formula according to no arbitrage assumption[1],

$$E_t(M_{t+1}R_{t+1}) = 1. \tag{2.2}$$

The relationship can be also applied in any foreign market. We use $*$ to denote variables in a foreign country, and the pricing formula becomes $E_t(M_{t+1}^* R_{t+1}^*) = 1$. To build connection in different markets, an investor in the U.S. could either invest \$1 in the domestic market

[1] We skip the technique part of uniqueness condition of pricing kernel. Harrison and Kreps (1979) provide detailed conditions of uniqueness of pricing kernel M_t.

to get total return R_{t+1}, or convert the \$1 dollar to $1/S_t$ pounds and then invest in the British market to get gross return R^*_{t+1}. S_t here means dollar price per foreign currency, i.e., USD/GBP. After one period, she converts pounds back to dollars at exchange rate S_{t+1} and values the returns $R_{t+1} = (S_{t+1}/S_t)R^*_{t+1}$ with domestic pricing kernel.

Then we have

$$\begin{aligned} E_t(M^*_{t+1}R^*_{t+1}) &= E_t(M_{t+1}R_{t+1}) \\ &= E_t(M_{t+1}\frac{S_{t+1}}{S_t}R^*_{t+1}) \end{aligned} \tag{2.3}$$

which is satisfied if

$$M^*_{t+1} = M_{t+1}\frac{S_{t+1}}{S_t}. \tag{2.4}$$

We use lower case letters to denote logarithm variables, so we have the following relationship between pricing kernels of two countries,

$$m^*_{t+1} - m_{t+1} = s_{t+1} - s_t. \tag{2.5}$$

2.4.2 Preference

We consider a representative agent with the Epstein and Zin (1989) preference

$$\begin{aligned} U_t &= \{(1-\beta)C_t^\rho + \beta[\mu_t(U_{t+1})]^\rho\}^{\frac{1}{\rho}} \\ \mu_t(U_{t+1}) &= E_t(U_{t+1}^\alpha)^{\frac{1}{\alpha}} \end{aligned} \tag{2.6}$$

where β is the rate of time preference, $\rho < 1$ captures the intertemporal elasticity of substitution (IES) which is $1/(1-\rho)$, $\alpha \leq 1$ captures risk aversion which is $1-\alpha$, and C_t is the aggregate consumption. Notice that when $\alpha = \rho$, the preference collapse to the constant relative risk aversion (CRRA) preference $U_t^\rho = (1-\beta)\sum_{j=0}^\infty \beta^j E_t C_{t+j}^\rho$.

With this utility function, the real pricing kernel is

$$M_{t+1} = \beta\left(\frac{C_{t+1}}{C_t}\right)^{\rho-1}\left(\frac{U_{t+1}}{\mu_t(U_{t+1})}\right)^{\alpha-\rho} \tag{2.7}$$

Again, if $\alpha = \rho$, the last term equals one and we have the pricing kernel as generated by standard CRRA preference.

The previous pricing kernel depends on future utilities which are not observable. Epstein and Zin (1989) show how to make the unobservable term U_{t+1} to be an "observable" return on wealth, $R_{w,t+1}$. $R_{w,t+1}$ could be understood as the return of the aggregate portfolio which pays aggregate consumption as dividend. To be specific,

$$\begin{aligned} R_{w,t+1} =& \frac{W_{t+1}}{W_t - C_t} \\ =& \frac{P_{t+1} + C_{t+1}}{P_t} \end{aligned} \tag{2.8}$$

where W_t is the total wealth at time t, and P_t is the asset price per share. This is true as in Lucas (1978), where we normalize the supply of asset to be one and risk-free asset supply to be zero. The aggregate consumption in this economy is equal to the aggregate dividend.

With the new return on wealth $R_{w,t+1}$, the real pricing kernel is

$$M_{t+1} = \beta^{\frac{\alpha}{\rho}} \left(\frac{C_{t+1}}{C_t}\right)^{\frac{(\rho-1)\alpha}{\rho}} R_{w,t+1}^{\frac{\alpha-\rho}{\rho}} \tag{2.9}$$

and the logarithm form of pricing kernel is

$$m_{t+1} = \frac{\alpha}{\rho} \log \beta + \frac{(\rho-1)\alpha}{\rho} g_{t+1} + \frac{\alpha-\rho}{\rho} r_{w,t+1}, \tag{2.10}$$

where $r_{w,t+1}$ needs to be solved as function of state variables in order to fully specify the pricing kernel.

2.4.3 Economy

We model the domestic exogenous consumption growth process $g_{t+1} = \log(C_{t+1}/C_t)$ as containing a persistent hidden component x_t and a transitory shock with stochastic volatility,

$$\begin{aligned} g_{t+1} &= \mu + x_t + \sigma_t \eta_{t+1} \\ x_{t+1} &= \phi x_t + \sigma_x \epsilon_{t+1} \\ \sigma_{t+1}^2 &= \sigma^2 + \nu(\sigma_t^2 - \sigma^2) + \sigma_e e_{t+1} \end{aligned} \tag{2.11}$$

where ϕ governs the degree of persistence of the hidden long run mean of the consumption growth, σ^2 is the long run mean of the stochastic volatility, ν controls the speed of mean-reverting effect, and $\eta_{t+1}, \epsilon_{t+1}, e_{t+1}$ are independent standard normally distributed with mean zero and variance one.

Similarly, for the foreign country, we have

$$
\begin{aligned}
g_{t+1}^{*} &= \mu^{*} + x_t^{*} + \sigma_t^{*}\eta_{t+1}^{*} \\
x_{t+1}^{*} &= \phi^{*}x_t^{*} + \sigma_x^{*}\epsilon_{t+1}^{*} \\
\sigma_{t+1}^{*2} &= \sigma^{*2} + \nu^{*}(\sigma_t^{*2} - \sigma^{*2}) + \sigma_e^{*}e_{t+1}^{*}
\end{aligned}
\tag{2.12}
$$

This model is similar to Bansal and Yaron (2004) Case II and we extend it to a two-country model. There are some differences between their model and our setup. First, we drop the dividend process and focus on solving interest rates and their application on currency market. Second, Bansal and Yaron (2004) add the same stochastic volatility σ_t to both the hidden state equation x_{t+1} and consumption growth equation g_{t+1}. However, the shocks to conditional volatility of the consumption growth and volatility of its conditional expectation might be different. Here we only consider the conditional volatility of g_{t+1} is time-varying but not the conditional volatility of x_{t+1}. Hence we have different interpretation of this model. They consider σ_t as economic uncertainty, while we consider this model as a benchmark model and investors cannot tell the source of a negative shock, which might come from either $\sigma_t\eta_{t+1}$ or ϵ_{t+1}. Uncertainties about the model misspecification are then mapped to uncertainties about sources of shocks. Shocks from ϵ_{t+1} are considered as long-run shocks since they change the conditional mean of the growth, while shocks from $\sigma_t\eta_{t+1}$ are simply transitory.

2.4.4 State Space Form

Our goal is to solve the model and link consumption growth to interest rate differential, and eventually get the similar state space form of interest rate differential in equation (2.1). We list all the key equations in this subsection and details in the appendix.

Following Campbell and Shiller (1988) decomposition, equation (2.8) can be written as

$$
r_{w,t+1} = \kappa_0 + \kappa_1 pc_{t+1} - pc_t + g_{t+1}, \tag{2.13}
$$

where $r_{w,t+1} = \log R_{w,t+1}$, and pc_t is the price-consumption ratio.

In equilibrium, we assume the ratio is a linear function of all state variables, x_t and σ_t^2, where three parameters A_0, A_1 and A_2 need to be determined,

$$pc_t = A_0 + A_1 x_t + A_2 \sigma_t^2. \tag{2.14}$$

We know the pricing kernel holds for any return and so is $R_{w,t+1}$,

$$\begin{aligned} 1 =& E_t(M_{t+1} R_{w,t+1}) \\ =& E_t(\exp(m_{t+1} + r_{w,t+1})) \\ =& E_t\left(\exp\left(\frac{\alpha}{\rho}\log\beta + \frac{(\rho-1)\alpha}{\rho} g_{t+1} + \frac{\alpha}{\rho} r_{w,t+1}\right)\right) \end{aligned} \tag{2.15}$$

Given that g_{t+1} and $r_{w,t+1}$ are conditional normal distributed, we can compute the above equation by using property of log normal distribution, i.e., $E(e^X) = exp(\mu + \frac{1}{2}\sigma^2)$ if $X \sim N(\mu, \sigma^2)$. After collecting coefficients before x_t and σ_t^2, respectively, we solve the parameters in equation (2.14),

$$A_1 = \frac{\rho}{1 - \kappa_1 \phi} \tag{2.16}$$

$$A_2 = \frac{\alpha\rho}{2(1 - \kappa_1 \nu)}, \tag{2.17}$$

and A_0 can be found in appendix.

Then the pricing kernel is also a function of state variables x_t and σ_t^2,

$$\begin{aligned} m_{t+1} =& \bar{m} - (1-\rho)x_t - \frac{\alpha(\alpha-\rho)}{2}\sigma_t^2 \\ & - (1-\alpha)\sigma_t \eta_{t+1} + \frac{\alpha-\rho}{\rho}\kappa_1 A_1 \sigma_x \epsilon_{t+1} + \frac{\alpha-\rho}{\rho}\kappa_1 A_2 \sigma_e e_{t+1}, \end{aligned} \tag{2.18}$$

and conditional exception and variance are therefore

$$E_t(m_{t+1}) = \bar{m} - (1-\rho)x_t - \frac{\alpha(\alpha-\rho)}{2}\sigma_t^2, \tag{2.19}$$

$$Var_t(m_{t+1}) = (1-\alpha)^2\sigma_t^2 + \left(\frac{\alpha-\rho}{\rho}\kappa_1 A_1 \sigma_x\right)^2 + \left(\frac{\alpha-\rho}{\rho}\kappa_1 A_2 \sigma_e\right)^2. \tag{2.20}$$

Since the basic pricing formula (2.2) holds for any asset, including risk-free asset, we have the risk-free rate as

$$
\begin{aligned}
r_t =& -\log E_t e^{m_{t+1}} \\
=& -E_t(m_{t+1}) - \frac{1}{2}Var_t(m_{t+1}) \\
=& \bar{r} + (1-\rho)x_t + \left(\frac{\alpha(\alpha-\rho)}{2} - \frac{1}{2}(1-\alpha)^2\right)\sigma_t^2 \\
=& \bar{r} + (1-\rho)x_t - \frac{1}{2}(\alpha\rho - 2\alpha + 1)\sigma_t^2
\end{aligned}
\tag{2.21}
$$

where $\bar{r}$ collects all the constants, including $\bar{m}$.

The interest rate is the same as standard CRRA preference if we set $\alpha = \rho$, where coefficients before x_t and σ_t^2 become risk aversion parameter $1-\alpha$,

$$
r_t = -\log\beta + (1-\alpha)x_t - \frac{1}{2}(1-\alpha)^2\sigma_t^2.
$$

Similarly, the foreign interest rate is

$$
r_t^* = \bar{r}^* + (1-\rho^*)x_t^* - \frac{1}{2}(\alpha^*\rho^* - 2\alpha^* + 1)\sigma_t^{*2}. \tag{2.22}
$$

We assume symmetry between domestic and foreign countries, and then we have the interest rate differential as a state space form considering $x_t - x_t^*$ as the persistent component,

$$
r_t - r_t^* = (1-\rho)(x_t - x_t^*) + \bar{r} - \bar{r}^* + \frac{1}{2}(\alpha\rho - 2\alpha + 1)(\sigma_t^{*2} - \sigma_t^2) \tag{2.23}
$$

where

$$
x_{t+1} - x_{t+1}^* = \phi(x_t - x_t^*) + (\sigma_x\epsilon_{x,t+1} - \sigma_x^*\epsilon_{x,t+1}^*). \tag{2.24}
$$

and $\sigma_x\epsilon_{x,t+1} - \sigma_x^*\epsilon_{x,t+1}^* \sim N(0, \sigma_x^2 + \sigma_x^{*2})$.

Last, we consider the interpretation from risk premium channel. Consider carry trade when foreign interest rate is larger than domestic interest rate, the excess return of such trade is $rx_{t+1} = s_{t+1} - s_t + r_t^* - r_t$,

$$
E_t(rx_{t+1}) = \frac{1}{2}(1-\alpha)^2(\sigma_t^2 - \sigma_t^{*2}). \tag{2.25}
$$

and we can see that risk premium is determined by the transitory shock. This is not consistent with the empirical fact observed in chapter 1, where long-run risk instead of temporary shock can explain positive β.

2.4.5 Extension

We can easily extent the model with more factors. Here we use the exogenous inflation process similar as Bansal and Shaliastovich (2013),

$$\begin{aligned}
\pi_{t+1} =& \pi + x_{\pi,t} + \sigma_{\pi,t}\eta_{\pi,t+1} \\
x_{\pi,t+1} =& \phi_\pi x_{\pi,t} + \sigma_\pi \epsilon_{\pi,t+1} \\
\sigma^2_{\pi,t+1} =& \sigma^2_\pi + \nu_\pi(\sigma^2_{\pi,t} - \sigma^2_\pi) + \sigma_{\pi,e} e_{\pi,t+1}
\end{aligned} \tag{2.26}$$

where all the shocks are independent with the real economy shocks.

Given the inflation process, the nominal pricing kernel $m^\$_{t+1}$ is

$$m^\$_{t+1} = m_{t+1} - \pi_{t+1}. \tag{2.27}$$

We follow the previous steps to compute real interest rate and add inflation since the nominal inflation part does not affect the real pricing kernel, then the nominal interest rate is

$$\begin{aligned}
i_t &= -E_t(m^\$_{t+1}) - \frac{1}{2} Var_t(m^\$_{t+1}) \\
&= \bar{i} + (1-\rho)x_t + x_{\pi,t} - \frac{1}{2}(\alpha\rho - 2\alpha + 1)\sigma^2_t - \frac{1}{2}\sigma^2_{\pi,t}
\end{aligned} \tag{2.28}$$

where $\bar{i}$ is a constant.

The nominal interest rate differential is

$$\begin{aligned}
i_t - i^*_t =& (1-\rho)(x_t - x^*_t) + (x_{\pi,t} - x^*_{\pi,t}) \\
&+ \bar{r} - \bar{r}^* + \frac{1}{2}(\alpha\rho - 2\alpha + 1)(\sigma^{*2}_t - \sigma^2_t) + \frac{1}{2}(\sigma^{*2}_{\pi,t} - \sigma^2_{\pi,t})
\end{aligned} \tag{2.29}$$

which can be rewritten as a vector form

$$i_t - i^*_t = (1-\rho; 1)\begin{pmatrix} x_t - x^*_t \\ x_{\pi,t} - x^*_{\pi,t} \end{pmatrix} + \text{Variance term} + \text{const} \tag{2.30}$$

$$\begin{pmatrix} x_{t+1} - x^*_{t+1} \\ x_{\pi,t+1} - x^*_{\pi,t+1} \end{pmatrix} = \begin{pmatrix} \phi & 0 \\ 0 & \phi_\pi \end{pmatrix}\begin{pmatrix} x_t - x^*_t \\ x_{\pi,t} - x^*_{\pi,t} \end{pmatrix} + \text{error} \tag{2.31}$$

which still satisfies the framework used in Li and Tornell (2015).

Our setup here uses less assumptions than Bansal and Shaliastovich (2013) where they assume that there is negative correlation between consumption growth and inflation process, which they call inflation non-neutrality. The interpretation highly depends on this assumption because in their model, if the correlation is 0, inflation plays no role in determining the state variable pc_{t+1}. But their empirical result shows this correlation is not significant. Here we do not require such relationship and still can explain the empirical facts in Chapter 1.

Hence, we map the uncertainties in Li and Tornell (2015) to the fundamental economic factors.

2.5 Conclusion

To explain the empirical effect that the variation of the slope estimates is driven by the long-run uncertainty, we link the pricing kernel literature to Li and Tornell (2015) with two fundamental economic factors: consumption growth and inflation. Built on their story, we move one step forward to map the source of persistent interest rate shocks to the long-run part of either consumption growth or inflation process. Our model is also based on long-run risk literature but has different interpretation after we modify the setup. Bansal and Yaron (2004) do not focus on where to put stochastic volatility which appears in front of transitory and persistent shocks. We carefully specify the location of stochastic volatility for only transitory shocks and differentiate the effect with long-run part.

We show that our model can explain the variation of the slope estimates during different sample periods. In words, when the economy is stable, investors typically consider shocks as transitory and underreact to news about consumption growth or inflation, which eventually affects the interest rate differential. When the U.S. interest rate increases, the dollar appreciates less today but continues to appreciate in the following days, which generates negative slope estimates. However, on the contrary, if the economy is in serious trouble such as the long-run part of consumption growth or inflation is shocked, investors become worried and overreact. In this case, we observe positive slope estimates.

2.6 Appendix

Derivation of (2.7) and (2.9)

The general pricing kernel is defined as

$$M_{t+1} = \frac{\partial U_t/\partial C_{t+1}}{\partial U_t/\partial C_t}$$

For Epstein-Zin preference, we have

$$\begin{aligned}\frac{\partial U_t}{\partial C_{t+1}} &= \frac{1}{\rho} U_t^{1-\rho} \beta\rho[\mu_t(U_{t+1})]^{\rho-1} \frac{\partial \mu_t(U_{t+1})}{\partial U_{t+1}} \frac{\partial U_{t+1}}{\partial C_{t+1}} \\ &= U_t^{1-\rho} \beta [\mu_t(U_{t+1})]^{\rho-1} \frac{1}{\alpha} E_t(U_{t+1}^{\alpha})^{\frac{1}{\alpha}-1} \alpha U_{t+1}^{\alpha-1} \cdot \frac{1}{\rho} U_{t+1}^{1-\rho}(1-\beta)\rho C_{t+1}^{\rho-1} \\ &= U_t^{1-\rho} \beta [\mu_t(U_{t+1})]^{\rho-1} E_t(U_{t+1}^{\alpha})^{\frac{1}{\alpha}-1} U_{t+1}^{\alpha-1} \cdot U_{t+1}^{1-\rho}(1-\beta) C_{t+1}^{\rho-1} \\ &= U_t^{1-\rho} \beta [\mu_t(U_{t+1})]^{\rho-\alpha} U_{t+1}^{\alpha-\rho}(1-\beta) C_{t+1}^{\rho-1}\end{aligned}$$

and

$$\begin{aligned}\frac{\partial U_t}{\partial C_t} &= \frac{1}{\rho} U_t^{1-\rho}(1-\beta)\rho C_t^{\rho-1} \\ &= U_t^{1-\rho}(1-\beta) C_t^{\rho-1}\end{aligned}$$

Combining these two, we have equation (2.7) as

$$\begin{aligned}M_{t+1} &= \frac{\partial U_t/\partial C_{t+1}}{\partial U_t/\partial C_t} \\ &= \beta \left(\frac{C_{t+1}}{C_t}\right)^{\rho-1} \left(\frac{U_{t+1}}{\mu_t(U_{t+1})}\right)^{\alpha-\rho}.\end{aligned}$$

To further simplify the pricing kernel formula and make it "observable", we consider that the wealth of a representative agent is the value of discounted future aggregate consumption,

$$W_t - C_t = E_t(M_{t+1} W_{t+1}). \tag{2.32}$$

Next we want to show the relationship between the representative agent's utility and wealth is

$$U_t = W_t \frac{\partial F(C_t, \mu_t(U_{t+1}))}{\partial C_t} \tag{2.33}$$

where $F(C_t, \mu_t(U_{t+1})) = \{(1-\beta)C_t^\rho + \beta[\mu_t(U_{t+1})]^\rho\}^{\frac{1}{\rho}}$.

Equation (2.33) can be solved as

$$\begin{aligned} W_t =& U_t(U_t^{1-\rho}(1-\beta)C_t^{\rho-1})^{-1} \\ =& \frac{1}{1-\beta}U_t^\rho C_t^{1-\rho} \end{aligned} \tag{2.34}$$

Here we use guess and verify method by plugging equation (2.34) into equation (2.32) and the right-hand side of equation (2.32) is

$$\begin{aligned} E_t(M_{t+1}W_{t+1}) =& E_t\left(\beta\left(\frac{C_{t+1}}{C_t}\right)^{\rho-1}\left(\frac{U_{t+1}}{\mu_t(U_{t+1})}\right)^{\alpha-\rho}\frac{1}{1-\beta}U_{t+1}^\rho C_{t+1}^{1-\rho}\right) \\ =& \frac{\beta C_t^{1-\rho}}{1-\beta}\frac{E_t U_{t+1}^\alpha}{\mu_t(U_{t+1})^{\alpha-\rho}} \\ =& \frac{\beta}{1-\beta}C_t^{1-\rho}\mu_t(U_{t+1})^\rho \end{aligned}$$

and the left-hand side of equation (2.32) is

$$\begin{aligned} W_t - C_t =& \frac{1}{1-\beta}U_t^\rho C_t^{1-\rho} - C_t \\ =& \frac{\beta}{1-\beta}C_t^{1-\rho}\left(\frac{1}{\beta}U_t^\rho - \frac{1-\beta}{\beta}C_t^\rho\right) \\ =& \frac{\beta}{1-\beta}C_t^{1-\rho}\left(\frac{1-\beta}{\beta}C_t^\rho + \mu_t(U_{t+1})^\rho - \frac{1-\beta}{\beta}C_t^\rho\right) \\ =& \frac{\beta}{1-\beta}C_t^{1-\rho}\mu_t(U_{t+1})^\rho \end{aligned}$$

Hence, we confirm the equation (2.33) and also (2.33) is a solution to equation (2.32).

Last, we explicitly define the return to total wealth as

$$\begin{aligned} R_{w,t+1} =& \frac{W_{t+1}}{W_t - C_t} \\ =& \frac{1}{\beta}\left(\frac{C_{t+1}}{C_t}\right)^{1-\rho}\left(\frac{U_{t+1}}{\mu_t(U_{t+1})}\right)^\rho \end{aligned}$$

where the second equation is true by plugging W_{t+1} and we can solve ratio $\frac{U_{t+1}}{\mu_t(U_{t+1})}$ as function of $R_{w,t+1}$ and plug it into pricing kernel, which gives us equation (2.9),

$$M_{t+1} = \beta^{\frac{\alpha}{\rho}}\left(\frac{C_{t+1}}{C_t}\right)^{\frac{(\rho-1)\alpha}{\rho}} R_{w,t+1}^{\frac{\alpha-\rho}{\rho}}.$$

Derivatoin of (2.13)

Following Campbell and Cochrane (1999),

$$\begin{aligned}R_{t+1} =& \frac{P_{t+1} + C_{t+1}}{P_t} \\ =& \left(\frac{P_{t+1}}{C_{t+1}}\frac{C_{t+1}}{C_t} + \frac{C_{t+1}}{C_t}\right)/\frac{P_t}{C_t} \\ =& \left(\left(1 + \frac{P_{t+1}}{C_{t+1}}\right)\frac{C_{t+1}}{C_t}\right)/\frac{P_t}{C_t}\end{aligned}$$

and we take logarithm on both sides

$$\begin{aligned}r_{t+1} =& \log(1 + \exp(p_{t+1} - c_{t+1})) + g_{t+1} - (p_t - c_t) \\ \approx& \kappa_0 + \kappa_1 pc_{t+1} + g_{t+1} - pc_t\end{aligned}$$

Derivation of (2.16) and (2.17)

First, we compute the conditional mean and volatility for $g_{t+1}, x_{t+1}, \sigma^2_{t+1}$ and pc_{t+1},

$$\begin{aligned}g_{t+1}|\mathcal{F}_t &\sim N(\mu + x_t, \sigma_t^2), \\ x_{t+1}|\mathcal{F}_t &\sim N(\phi x_t, \sigma_x^2), \\ \sigma^2_{t+1}|\mathcal{F}_t &\sim N(\sigma^2 + \nu(\sigma_t^2 - \sigma^2), \sigma_e^2), \\ pc_{t+1}|\mathcal{F}_t &\sim N(A_0 + A_1\phi x_t + A_2(\sigma^2 + \nu(\sigma_t^2 - \sigma^2)), A_1\sigma_x^2 + A_2\sigma_e^2).\end{aligned} \tag{2.35}$$

Then we expand the expectation for log normal,

$$1 = E_t\left(\exp\left(\frac{\alpha}{\rho}\log\beta + \frac{(\rho-1)\alpha}{\rho}g_{t+1} + \frac{\alpha}{\rho}r_{w,t+1}\right)\right) \tag{2.36}$$

which gives

$$
\begin{aligned}
1 =& \frac{\alpha}{\rho}\log\beta + \frac{(\rho-1)\alpha}{\rho}E_t(g_{t+1}) + \frac{\alpha}{\rho}E_t(r_{w,t+1}) \\
&+ \frac{1}{2}\frac{(\rho-1)^2\alpha^2}{\rho^2}Var_t(g_{t+1}) + \frac{1}{2}\frac{\alpha^2}{\rho^2}Var_t(r_{w,t+1}) \\
&+ \frac{1}{2}\frac{(\rho-1)\alpha^2}{\rho^2}Cov_t(g_{t+1}, r_{w,t+1}) \\
=& \frac{\alpha}{\rho}\log\beta + \frac{(\rho-1)\alpha}{\rho}(\mu + x_t) \\
&+ \frac{\alpha}{\rho}\kappa_0 + \frac{\alpha}{\rho}\kappa_1[A_0 + A_1\phi x_t + A_2(\sigma^2 + \nu(\sigma_t^2 - \sigma^2))] \\
&- \frac{\alpha}{\rho}(A_0 + A_1 x_t + A_2\sigma_t^2) + \frac{\alpha}{\rho}(\mu + x_t) \\
&+ \frac{1}{2}\frac{(\rho-1)^2\alpha^2}{\rho^2}\sigma_t^2 + \frac{1}{2}\frac{\alpha^2}{\rho^2}[\kappa_1^2(A_1\sigma_x^2 + A_2\sigma_e^2) + \sigma_t^2] + \frac{1}{2}\frac{(\rho-1)\alpha^2}{\rho^2}\sigma_t^2
\end{aligned} \tag{2.37}
$$

Second we collect all coefficients for the same terms,

$$
\begin{aligned}
1 =& \frac{\alpha}{\rho}\log\beta + \frac{(\rho-1)\alpha}{\rho}\mu + \frac{\alpha}{\rho}\kappa_0 + \frac{\alpha}{\rho}\kappa_1[A_0 + A_2(\sigma^2 - \nu\sigma^2)] \\
&- \frac{\alpha}{\rho}A_0 + \frac{\alpha}{\rho}\mu + \frac{1}{2}\frac{\alpha^2}{\rho^2}[\kappa_1^2(A_1\sigma_x^2 + A_2\sigma_e^2)] \\
0 =& \frac{(\rho-1)\alpha}{\rho} + \frac{\alpha}{\rho}(\kappa_1 A_1\phi - A_1 + 1) \\
0 =& \frac{\alpha}{\rho}(\kappa_1 A_2\nu - A_2) + \frac{1}{2}\frac{(\rho-1)^2\alpha^2}{\rho^2} + \frac{1}{2}\frac{\alpha^2}{\rho^2} + \frac{(\rho-1)\alpha^2}{\rho^2}
\end{aligned}
$$

which gives

$$
A_1 = \frac{\rho}{1-\kappa_1\phi} \tag{2.38}
$$

$$
A_2 = \frac{\alpha\rho}{2(1-\kappa_1\nu)}, \tag{2.39}
$$

Derivation of (2.18)

The pricing kernel is

$$
\begin{aligned}
m_{t+1} =& \frac{\alpha}{\rho}\log\beta + \frac{(\rho-1)\alpha}{\rho} g_{t+1} + \frac{\alpha-\rho}{\rho} r_{w,t+1} \\
=& \frac{\alpha}{\rho}\log\beta + \frac{(\rho-1)\alpha}{\rho} g_{t+1} + \frac{\alpha-\rho}{\rho} \\
& \cdot \{\kappa_0 + \kappa_1(A_0 + A_1 x_{t+1} + A_2\sigma_{t+1}^2) - (A_0 + A_1 x_t + A_2\sigma_t^2) + g_{t+1}\} \\
=& \frac{\alpha}{\rho}\log\beta + \frac{\alpha-\rho}{\rho}[\kappa_0 + \kappa_1 A_0 - A_0 - A_1 x_t - A_2\sigma_t^2] + \\
& -(1-\alpha)g_{t+1} + \frac{\alpha-\rho}{\rho}\kappa_1 A_1 x_{t+1} + \frac{\alpha-\rho}{\rho}\kappa_1 A_2 \sigma_{t+1}^2 \\
=& \frac{\alpha}{\rho}\log\beta + \frac{\alpha-\rho}{\rho}[\kappa_0 + \kappa_1 A_0 - A_0 - A_1 x_t - A_2\sigma_t^2] \\
& -(1-\alpha)(\mu + x_t) + \frac{\alpha-\rho}{\rho}\kappa_1 A_1 \phi x_t + \frac{\alpha-\rho}{\rho}\kappa_1 A_2(\sigma^2 + \nu(\sigma_t^2 - \sigma^2)) \\
& -(1-\alpha)\sigma_t\eta_{t+1} + \frac{\alpha-\rho}{\rho}\kappa_1 A_1 \sigma_x \epsilon_{t+1} + \frac{\alpha-\rho}{\rho}\kappa_1 A_2 \sigma_e e_{t+1} \\
=& \bar{m} - (1-\rho)x_t - \frac{\alpha(\alpha-\rho)}{2}\sigma_t^2 \\
& -(1-\alpha)\sigma_t\eta_{t+1} + \frac{\alpha-\rho}{\rho}\kappa_1 A_1 \sigma_x \epsilon_{t+1} + \frac{\alpha-\rho}{\rho}\kappa_1 A_2 \sigma_e e_{t+1}
\end{aligned}
\tag{2.40}
$$

where $\bar{m}$ contains all the constants.

Derivation of CRRA

Consider CRRA utility,

$$U(C_t) = \frac{C_t^{1-\gamma}}{1-\gamma}$$

Pricing kernel is

$$
\begin{aligned}
M_{t+1} &= \beta\frac{u'(C_{t+1})}{u'(C_t)} = \beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} \\
m_{t+1} &= \log\beta - \gamma g_{t+1} \text{ where } g_{t+1} = \log\left(\frac{C_{t+1}}{C_t}\right)
\end{aligned}
$$

$$
\begin{aligned}
r_{f,t} =& -\log E_t e^{m_{t+1}} = -E_t(m_{t+1}) - \frac{1}{2}Var_t(m_{t+1}) \\
=& \log\beta + \gamma x_t - \frac{1}{2}\gamma^2\sigma_t^2
\end{aligned}
$$

Derivation of (2.25)

Given that the risk premium is $rx_{t+1} = s_{t+1} - s_t + r_t^* - r_t$, we have

$$\begin{aligned} E_t(rx_{t+1}) =& E_t(s_{t+1} - s_t + r_t^* - r_t) \\ =& E_t(m_{t+1}^*) - E_t(m_{t+1}) - E_t(r_t - r_t^*) \\ =& \frac{1}{2}Var_(m_{t+1}) - \frac{1}{2}Var_t(m_{t+1}^*) \\ =& \frac{1}{2}(1-\alpha)^2(\sigma_t^2 - \sigma_t^{*2}), \end{aligned} \tag{2.41}$$

where the second equality comes from the fact that $s_{t+1} - s_t = r_t - r_t^*$.

3. Exchange Rate Forecasting Using Big Data

3.1 Introduction

It is well known that economic models have little predictability power in exchange rates and usually cannot outperform the random walk model in out-of-sample forecasting (Meese and Rogoff (1983)). Another consensus in the literature is that models such as the purchasing power parity (PPP) and monetary models have more difficulty in prediction with horizon less than about 2 years. Can we use some other nontraditional factors to predict exchange rates, especially in the short run? Our answer is text mining.

There is an increasing interest in the subject of big data. One ongoing challenge within the world of "big data" analytics is how to analyze and use information from large text dataset archives such as tweets and news articles. This is called text mining or natural language processing (NLP). Although text mining is widely applied in computer science and machine learning fields, it is relatively new to economics and finance. However, text mining is worthy our attention since this new technique can access to additional information from many text sources which cannot be quantified by other methods.

An important example of a smaller text dataset for macroeconomists and practitioners in the financial industry is from the Federal Open Market Committee (FOMC), a committee within the Federal Reserve who are in charge of the open market operation and decide the US monetary policy. They release several types of documents after their regular meeting during the year such as statements, minutes, press conference transcripts etc. The reason why we choose the FOMC text documents is obvious. The Committee determine one of the most important benchmark interest rates – the federal fund rate, which is heavily monitored by the whole market. Academic research also shows stronger evidence of prediction power from the central bank related factors than other economic models. Molodtsova and Papel (2009) find that Taylor-rule predictors have predictive ability for exchange rate changes and Rossi (2013) confirms their finding after doing an extensive review of the literature. Because central banks follow certain version of the Taylor rule (Taylor (1993)), what the Federal Reserve says and reports to the public contains important information about their own forecast about several key macroeconomic aggregates for the near future, including exchange rates. Different from the traditional exchange rate forecasting with economic fundamentals, this chapter aims to forecast G10 currencies based on the information retrieved from the text documents of the FOMC meetings. Across the nine currencies and one portfolio including all currencies, in most cases our out-of-sample forecasts beat the random walk model for horizon less than one month. The longer the horizon, the less prediction power we have. We might conclude that the information is absorbed by the market gradually according to our empirical result.

In our empirical strategy, we first construct factors from the FOMC texts by using techniques in the text mining literature, since all the original qualitative resources should be converted into usable quantitative features. We mainly use three types of documents from the FOMC: the *statements* regarding the meeting's policy decisions, the *minutes* which summarize issues addressed at each FOMC meeting, and the *press conference transcripts* which record what the chairman says to the press and the Q&A session in quarterly press conference. We apply three different techniques to the documents, respectively, because of the unique characteristics in each type of text. The sample starts from 2010 to 2016, which includes 57 FOMC meetings. For the statements. we use *bag-of-words* to take every word's weight into account, which gives us a wide term-document matrix (small observations but large dimensions). We apply *Latent Semantic Analysis* to extract hidden themes from the minutes. We generate a time series of each theme across all meetings. The press conference transcripts are compared with the minutes for last meeting using *vector space model*, resulting a measure about content similarity.

Second, with all the constructed features or predictors from the FOMC text documents, we feed different combinations of features into a "black box" machine learning algorithm and train the models. We predict the direction of the exchange rate changes instead of point forecasts, and label the ups and downs as $\{1, -1\}$. In this case, it is called supervised learning in the machine learning literature. *Support vector machine* (SVM) is chosen as our candidate algorithm[1]. SVM is considered as one of the best out of the box classifier since it tents to avoid overfitting, and it performs well in handling text mining features.

Last, out of sample forecasting is conducted in subsamples starting from January 2014 and there are 24 meetings. We use all data up the end of 2013 as the first training sample set, including 23 meetings. Instead of using rolling window estimation, for one new FOMC meeting, we extend the training data with the new information. Besides nine individual currencies, we also construct a portfolio with $10,000 total risk, which means that the sum of absolute position equals $10,000. For those we predict to appreciate, we long $1,111 ($10,000/9) and short $1,111 for those to depreciate. Our forecast horizons include from 1 week to 6 weeks, with 1 week increment. The last 6-week horizon is chosen since two FOMC meetings are about six weeks apart. We only forecast the exchange rate changes between two meetings. For the best performance with 1 week horizon, out of the 216 forecasts (24 meetings for all nine currencies), we correctly forecast the directions 125 times and our result is significantly better than the random walk model with the weighted directional test.

The rest of the chapter is organized as follows. In Section 2, we discuss the related literature in empirical analysis. Section 3 reviews the FOMC meetings and their text documents. In Section 4, we construct the predictors from all the text documents with techniques in the text mining literature. Section 5 gives the result. Section 6 concludes.

3.2 Literature

Our paper is linked to several branches of the literature: text mining, the FOMC documents analysis and the exchange rate forecasting.

First, a number of recent papers use text mining techniques as an alternative way to empirically analyze a wide range of topics in both economics and finance. For example, traditionally, researchers measure sentiments based on either some market variables such as

[1]We also use other algorithms such as Adaboost and neural network with a small number of layers. Support vector machines gives the best out-of-sample forecasts.

trading volumes or survey data. Da, Engelberg and Gao (2015) construct a Financial and Economic Attitudes Revealed by Search (FEARS) index to measure investors' sentiment using daily Internet search volume. Specifically, they focus on search volume within the U.S. and find that FEARS predicts short-term asset prices, volatility and fund flows. Tetlock (2007) finds that high media pessimism predicts downward pressure on market prices followed by a reversion to fundamentals. Using a well-known automated quantitative content analysis program called General Inquirer, he constructs the proxy for investors sentiment or noninformational trading from the Wall Street Journal's "Abreast of the Market" column. GI can convert a text into a numeric value by counting the number of words that fall within some categories. Here the author chooses the Negative words and Weak words categories defined in Harvard IV-4 psychosocial dictionary.

Hoberg and Phillips (2016) classify firms into different industries based on text-based analysis of firm 10-K product descriptions instead of traditional SIC industry codes. Then they use the new industry classification to show that industry shocks change competition and product offerings. Tetlock (2011) tests whether investors can tell the difference between new and old information about firms. He uses an extensive public news archive and measures the similarity between two news stories. The similarity is computed by dividing the total number of words in the intersection of the two texts by the union of words in both texts. He finds that investors overreact to stale information. Loughran and McDonald (2011) design their own dictionaries and measure the tone of a text. Da, Engelberg and Gao (2011) simply use the search frequency in Google to predict stock price.

Second, few papers focus particularly on the Federal Open Market Committee (FOMC) text documents particularly. Actually, central bank communication is becoming a key tool for inflation expectation management. Boukus and Rosenberg (2006) apply a text mining technique called Latent Semantic Analysis to the minutes from the FOMC for the past two decades. They find that the information retrieved from the minutes are correlated with current and future economic fundamental variables such as Treasury yield. Hansena and McMahon (2016) measure the information from the FOMC by applying Latent Dirichlet Allocation (LDA) and dictionary methods to the communications (statements). They find that this type of information has effects on both the market and real economic fundamentals. Acosta (2015) also applies Latent Semantic Analysis to FOMC minutes from 1976 to 2008 and finds that the Fed continues to increase transparency which reflects a response to the requirement that the Fed should provide more details in reporting to its monetary policies.

The last branch of literature is about exchange rate forecasting. Meese and Rogoff (1983) find that economic models with fundamentals cannot beat the simple random walk model in out-of-sample exchange rate forecast. It is also called exchange rate disconnect puzzle. Engel and West (2005) show that in a rational expectations present-value model, exchange rates can be well approximated as random walks if two countries have roughly similar inflation rates. Molodtsova and Papel (2009) use Taylor rule based model to forecast one-month exchange rates, and their results are statistically significant better than the random walk model using Clark and West (2006) test. Kim, Liao and Tornell (2014) find that the speculators' net position in futures market help forecast exchange rates at short prediction horizon. Rossi (2013) provides a survey about exchange rate predictability.

3.3 The FOMC Meetings and Text Data

The Federal Open Market Committee is responsible for the Federal Reserve's open market operation. The FOMC holds eight regular meetings every year. "At these meetings, the Committee reviews economic and financial conditions, determines the appropriate stance of monetary policy, and assesses the risks to its long-run goals of price stability and sustainable economic growth."[2] The result of each FOMC meeting is heavily monitored by all participants in the financial market because one of the most important benchmark interest rates, the federal fund rate, is determined during the meeting.

However, the public does not have access to much information from the FOMC at the beginning. The Federal Reserve only published a summary of FOMC proceedings once a year to the Congress and these documents were confidential to the public in order to implement

[2] Details can be found from https://www.federalreserve.gov/monetarypolicy/fomc.htm.

the monetary policy. The FOMC made the minutes between 1936 and 1960 available to the public. The Freedom of Information Act in 1967 required the FOMC to make information about monetary policy to be public on a timely basis and the lag was 90 days after the meeting. From then on, the FOMC release more and more information about the process of monetary policy decisions. After the end of 2004, the FOMC publish minutes three weeks after each meeting. Since two meetings are about six weeks apart, minutes are released in the middle between two meetings. In addition, in January 2000 the Committee announced that it would issue a statement following each regularly scheduled meeting, regardless of whether there had been a change in monetary policy.

On April 17th, 2011, for the first time in the 98-year history of the nation's central bank, the chairman Ben Bennanke talked to the press after the quarter end FOMC meeting and discussed the monetary policy decision. Since then, the FOMC follows the same information disclosure procedure. A statement regarding its policy decisions is released immediately after the meeting. If the meeting is held at the end of the quarter, the chairman holds a press conference to discuss about why the decision is made and answer questions from the public. After three weeks, the Committee publish all the details in the minutes about what has happened during the last meeting. The purposes of these three types of documents are different from above introduction and so provide us different dimension of the information. They are not interchangeable. Hence, we analyze all the three types of text documents released by the FOMC.

We choose our sample period from 2010 to 2016 because this period is relatively stable compared to the recent 2008 financial crisis. During the sample period, the FOMC has had 56 meetings and one additional conference call in total. Then we have 57 statements but 55 minutes. The minutes are two less than the meetings. First, by the end of 2016 we cannot get the minutes for 20161214 meeting because minutes is released three weeks later which is 20170104. Second, there is one additional conference call on May 9, 2010 and the FOMC combine the minutes with next formal meeting on June 22, 2015. The press conference transcripts are available after 20110427 and there are 24 transcripts in total because the chairman of the Fed hold quarterly conference.

The exchange rates data are downloaded from Global Financial Data (GFD). We include all G10 currencies: the United States dollar, the euro, the Japanese yen, the British pound sterling, the Swiss franc, the Australian dollar, the New Zealand dollar, the Canadian dollar, the Swedish krona, and the Norwegian krone. Again, we consider the U.S. as domestic country, All the exchange rates have been converted to the dollar price of one unit of foreign currency. For each currency pair, the exchange rate is the daily rate from 2010 to 2016. For different forecast horizons, we compute corresponding returns.

3.4 Methods

In this section, we discuss how to construct features (or predictors, factors) from the FOMC text documents and the machine learning algorithm used to predict future exchange rates.

3.4.1 NLP Preprocessing

Before we construct useful factors or features from the text documents, we need to quantify or parse every text file first. A traditional simple way of doing this in the NLP literature is to convert every text file into a word vector by counting the occurrence of each word appearing in the file. For example, let us consider the following two sentences in two separate files,

1. John likes to learn economics. Mary likes economics too.

2. John also likes to watch football games.

Based on these two sentences, two word lists are represented as below,

$$\text{word list 1} = \begin{pmatrix} \text{John} & 1 \\ \text{likes} & 2 \\ \text{to} & 1 \\ \text{learn} & 1 \\ \text{economics} & 2 \\ \text{Mary} & 1 \\ \text{too} & 1 \end{pmatrix}, \ \text{word list 2} = \begin{pmatrix} \text{John} & 1 \\ \text{also} & 1 \\ \text{likes} & 1 \\ \text{to} & 1 \\ \text{watch} & 1 \\ \text{football} & 1 \\ \text{games} & 1 \end{pmatrix}.$$

We have already seen how to convert a text file to a word vector in a simple situation. However, in order to make every word vector consistent and comparable, there are several standard preprocessing steps in NLP. The first step is to eliminate capitalization, punctuation, symbols and numbers. The downside of capitalization removal is that it might obscure meaning of some words. For example, "Bank"s in "Bank of England" and "investment bank" are referring to different concepts. Second, we remove so-called stop-words such as the, a from the text, because they are superfluous and do not contribute to the meaning of the context. We use the stop-words list from Snowball[3]. Third, we apply Porter2[4] stemming algorithm to stem the remaining words, which means a crude heuristic process that chops off the ends of words. For example, `consumer` and `consumers` both end up with `consum`. In the appendix, we show how to do the preprocessing step by step to one paragraph from the statement in 20161214.

Compared to other texts on the website such as comments on Twitter or customer reviews on Amazon, the FOMC documents have high quality in the sense that the Committee uses words precisely and formats and contents are also very consistent. We will never see words such as yaaaaaaaaaaaay, hahahahahaha in the FOMC released materials. This feature makes our analysis easier and more accurate than other types of informal texts. The FOMC documents have good quality in the sense that words are used in stylistic purposes, and

[3]http://snowball.tartarus.org/algorithms/english/stop.txt.

[4]http://snowball.tartarus.org/algorithms/english/stemmer.html.

readers should not over-interpreting the meaning of specific word[5].

Besides the preprocessing steps in NLP, the FOMC documents have their own writing styles. Structure of the statements is the simplest. Several lines such as "Release Date:...", "For immediate release", "Voting for the..." have been ignored when the texts are fed into the parsing algorithm because these lines have irrelevant information.

In the minutes, only the content between line "Developments in Financial Markets ..." and line "The vote encompassed..." is parsed. The beginning section in the minutes contains information about participants of the meeting. The annual minutes (the one published in February) have more irrelevant information about organizational issues since the year end meeting has some routine administrative agenda to follow. At the end of each minute, all the words in the last statement are repeated and so this part is useless. There is no extra cleaning process for press conference transcripts.

3.4.2 Text Mining Techniques

Given the word vectors generated from the preprocessing steps, there are many text mining techniques which could be applied. Considering the properties of the FOMC documents, we use three methods from the NLP literature to different documents: bag-of-words (BOW), Latent Semantic Analysis (LSA) and vector space model.

Bag-of-Words BOW is a method which combines all word vectors column by column but consider all unique words as the row index. If all documents are arranged together as a corpus, we have a sparse term-document matrix $\{f_{t,d}\}$, where $t = 1, ..., T$ is the unique stemmed word and $d = 1, ..., D$ denotes document. let us consider the previous example

[5] See a short article "Background on FOMC Meeting Minutes" by Deborah J. Danker and Matthew M. Luecke

again. The term-document matrix is then a 11×2 matrix

$$\begin{pmatrix} \text{John} & 1 & 1 \\ \text{also} & 0 & 1 \\ \text{likes} & 2 & 1 \\ \text{to} & 1 & 1 \\ \text{learn} & 1 & 0 \\ \text{economics} & 2 & 0 \\ \text{Mary} & 1 & 0 \\ \text{too} & 1 & 0 \\ \text{watch} & 0 & 1 \\ \text{football} & 0 & 1 \\ \text{games} & 0 & 1 \end{pmatrix}.$$

Since we might put more weight to a small group of words which appear frequently, the simple BOW might overestimate their importance. For example, the word “like” might not give us too much information when we compare these two documents. What matters are “economics” and “football”. Even though this raw term-document matrix can be used as features in prediction directly. But usually we can put global weights to each word. The most common weighting method is term-frequency-inverse-document-frequency (TFIDF) as

$$tfidf_{t,d} = f_{t,d} \log \frac{D}{df_t},$$

where df_t is the number of documents in which term t appears. The first term is the previous term frequency and gives larger weight to frequent words. The second term adjusts the weighting scheme by giving less weight to words appear more frequently.

For the statements, we use BOW since statement is shorter in terms of number of words used, which means each word might be important to be analyzed. The Committee also is cautious the terms they used because the statement is the first text document available to the public after each meeting. The whole financial industry is heavily monitoring what the Committee says. For example, the Wall Street Journal (WSJ) has a Fed Statement Tracker[6]

[6] http://graphics.wsj.com/fed-statement-tracker/.

to compare any of two statements word by word in order to find some clues about the future economy. It is consistent with the spirit of BOW.

In our sample, the FOMC uses 17,276 words for all 57 statements and then the average is 303 words per each statement. There are 624 unique words for all statements, so $T = 624, D = 57$ and the term-document matrix is $f_{624,57}$. Figure 3.1 lists the top 25 frequent words from the raw term-document matrix year by year. It is obvious that "inflation" attracts much concern from the Committee since this word almost appears most frequently every year. Some intuitive topics are also within top 25: "employment", "economy", "price", "stability", etc. Interestingly, "mortgage-back" is frequently mentioned during 2012 and 2013. In 2016, "expect" is the top 2 word which might reflect the fact that the Committee has more expectation in year 2016.

Figure 3.2 lists the top 25 weighted words after we adjust the inverse-document-frequency. The focus totally shifts from the previous common topics for economy to some rare words. For example, in 2010, we see many words related to the currency market such as "dollar", "swiss", "japan". In 2011, the Committee started to talk about the mortgage-back securities problem. In 2012, the commodity market might be very important from these words like "oil" and "gasoline".

One type of research could rely on some important words picked by the reader and construct some sort of measures about each statement. However, for all the previous interpretation from BOG, it requires the reader's prior knowledge about economics and this might be a shortcoming because of subjective judgment. For example, in Figure 3.2 we randomly highlight some words for each year and interpret in our own way, which might not contain more information. Human interpretation might be missed by human readers who have their own prior knowledge and overlook some general pattern. In the machine learning algorithm that we introduce in later subsection, we feed every word into the algorithm instead of a preselected list of words.

Latent Semantic Analysis LSA is an algorithm which identifies common factors for a collection of text documents. Mathematically, LSA is a singular value decomposition on

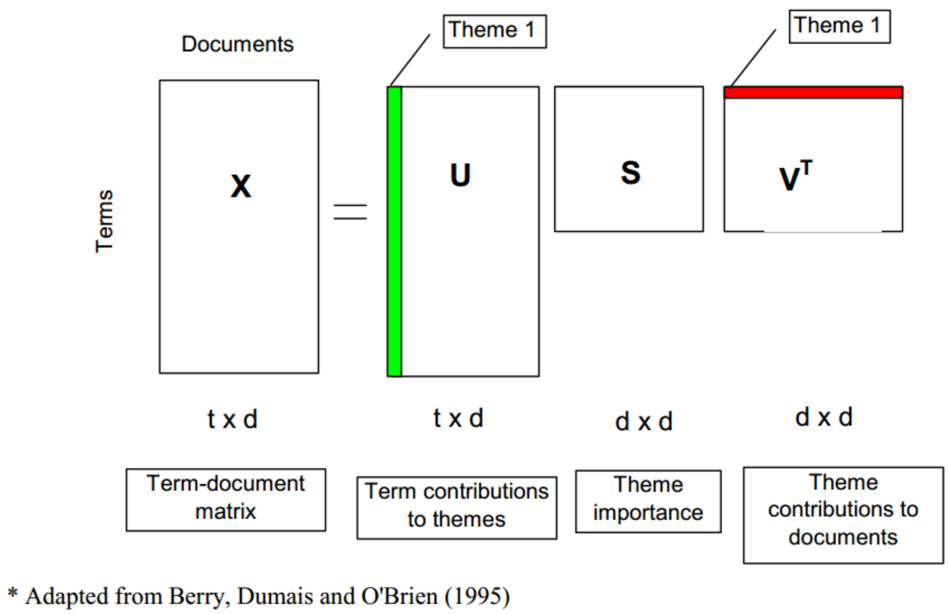
Figure 3.1: Top 25 frequent words in statements: raw count

2010		2011		2012		2013		2014		2015		2016	
continu	2.53%	inflat	3.38%	rate	2.57%	inflat	2.84%	inflat	3.36%	inflat	3.86%	inflat	3.69%
econom	2.27%	econom	2.39%	econom	2.31%	econom	2.07%	market	2.31%	market	2.39%	expect	2.70%
inflat	2.22%	continu	2.34%	inflat	2.26%	rate	1.97%	labor	2.02%	labor	2.39%	rate	2.41%
feder	2.07%	rate	2.08%	secur	2.15%	polici	1.94%	committe	1.57%	continu	1.67%	feder	2.41%
level	1.82%	price	1.93%	continu	1.80%	market	1.90%	econom	1.54%	percent	1.63%	econom	2.24%
market	1.77%	level	1.87%	remain	1.64%	purchas	1.70%	employ	1.51%	price	1.63%	market	2.20%
bank	1.72%	secur	1.77%	market	1.59%	secur	1.64%	condit	1.45%	rang	1.59%	fund	2.07%
rate	1.72%	consist	1.66%	agenc	1.54%	pace	1.64%	secur	1.43%	econom	1.55%	percent	1.99%
remain	1.46%	mandat	1.66%	consist	1.44%	continu	1.44%	rate	1.43%	employ	1.55%	labor	1.95%
secur	1.31%	expect	1.40%	price	1.44%	accommod	1.44%	polici	1.40%	expect	1.51%	condit	1.91%
fund	1.16%	hold	1.35%	mandat	1.33%	percent	1.40%	percent	1.34%	level	1.47%	price	1.58%
employ	1.11%	feder	1.30%	polici	1.28%	appropri	1.37%	purchas	1.31%	feder	1.43%	secur	1.33%
resourc	1.11%	recoveri	1.25%	expect	1.23%	agenc	1.34%	expect	1.28%	polici	1.39%	polici	1.33%
like	1.06%	remain	1.20%	feder	1.23%	expect	1.34%	remain	1.23%	rate	1.39%	remain	1.29%
recoveri	1.06%	anticip	1.09%	hold	1.23%	labor	1.34%	current	1.17%	remain	1.35%	indic	1.20%
financi	1.06%	employ	1.04%	mortgageback	1.13%	remain	1.17%	pace	1.14%	target	1.31%	declin	1.12%
expect	1.01%	low	1.04%	anticip	1.08%	month	1.17%	continu	1.11%	secur	1.27%	rang	1.12%
low	1.01%	market	1.04%	improv	1.03%	consist	1.17%	improv	1.11%	maximum	1.19%	strengthen	1.08%
stabil	1.01%	stabil	0.99%	employ	1.03%	employ	1.14%	maximum	1.08%	condit	1.15%	financi	1.04%
anticip	0.96%	outlook	0.88%	stabil	0.97%	longerterm	1.10%	longerterm	1.05%	near	1.11%	level	1.00%
price	0.96%	fund	0.88%	accommod	0.97%	mortgageback	1.07%	agenc	1.03%	fund	1.11%	continu	1.00%
pace	0.91%	dual	0.83%	month	0.92%	condit	1.04%	toward	1.03%	declin	1.07%	agenc	1.00%
condit	0.91%	pace	0.78%	dual	0.92%	committe	1.00%	month	1.00%	indic	1.07%	inform	1.00%
stabl	0.86%	inform	0.78%	financi	0.87%	price	0.97%	feder	1.00%	improv	1.00%	measur	1.00%
hous	0.86%	recent	0.78%	purchas	0.87%	longerrun	0.94%	asset	0.97%	agenc	0.96%	maintain	0.95%

Figure 3.2: Top 25 frequent words in statements: TF-IDF weighting

2010		2011		2012		2013		2014		2015		2016	
dollar	1.60%	**agenc**	1.61%	purchas	2.34%	assess	1.91%	**increas**	1.69%	earlier	1.97%	still	2.66%
arrang	1.49%	**matur**	1.38%	pace	1.27%	larg	1.63%	guidanc	1.50%	monetari	1.76%	slow	2.47%
central	1.21%	measur	1.14%	year	1.23%	move	1.62%	per	1.44%	growth	1.59%	somewhat	2.46%
board	1.20%	under	1.14%	increas	1.17%	back	1.62%	billion	1.44%	**increas**	1.36%	rough	1.85%
leav	1.20%	**mortgageback**	1.07%	asset	1.17%	economi	1.45%	occur	1.37%	slow	1.35%	addit	1.84%
mandat	1.01%	oil	1.05%	billion	1.17%	sustain	1.38%	forward	1.23%	strengthen	1.33%	diminish	1.80%
consist	1.01%	year	0.98%	judg	1.14%	rather	1.28%	statement	1.23%	actual	1.33%	appear	1.79%
england	1.00%	well	0.98%	**oil**	1.12%	add	1.28%	believ	1.23%	adjust	1.33%	strong	1.78%
nation	1.00%	chain	0.98%	**gasolin**	1.12%	reduc	1.26%	lower	1.18%	futur	1.33%	begin	1.77%
swiss	1.00%	disrupt	0.85%	**longerrun**	1.02%	variat	1.22%	suggest	1.13%	stanc	1.32%	move	1.76%
japan	1.00%	although	0.83%	addit	1.01%	temporari	1.22%	declin	1.08%	occur	1.32%	unemploy	1.76%
loan	0.95%	growth	0.79%	pick	0.88%	part	1.19%	energi	1.00%	like	1.26%	nearterm	1.75%
facil	0.87%	recent	0.77%	earlier	0.85%	transitori	1.19%	marketbas	1.00%	near	1.23%	hous	1.75%
swap	0.80%	trend	0.76%	adjust	0.83%	increas	1.18%	month	0.99%	import	1.14%	export	1.75%
action	0.80%	quarter	0.76%	regular	0.83%	prepar	1.18%	somewhat	0.98%	lower	1.12%	sector	1.75%
european	0.80%	also	0.76%	review	0.83%	could	1.17%	patient	0.96%	decreas	0.99%	net	1.75%
canada	0.80%	come	0.76%	end	0.80%	paus	1.16%	previous	0.96%	steadi	0.98%	point	1.75%
author	0.80%	suppli	0.75%	quarter	0.77%	reaffirm	1.16%	also	0.86%	somewhat	0.97%	grow	1.74%
bank	0.79%	earlier	0.73%	main	0.77%	proceed	1.15%	abov	0.86%	chang	0.97%	pick	1.74%
purchas	0.68%	slower	0.71%	despit	0.77%	influenc	1.10%	chang	0.84%	press	0.96%	although	1.74%
liquid	0.67%	reflect	0.71%	late	0.77%	last	1.07%	paragraph	0.80%	show	0.96%	solid	1.74%
back	0.64%	declin	0.66%	gradual	0.77%	fall	1.01%	rather	0.72%	stay	0.95%	suggest	1.74%
hold	0.62%	pick	0.65%	subsequ	0.76%	factor	1.01%	add	0.72%	shown	0.95%	even	1.73%
avail	0.60%	program	0.62%	temporarili	0.76%	although	1.01%	make	0.72%	recoveri	0.94%	judg	1.71%
pdf	0.60%	higher	0.60%	somewhat	0.75%	becom	1.01%	recent	0.71%	strong	0.94%	seen	1.71%

Figure 3.3: LSA on a term-document matrix



term frequency matrix, either $\{f_{t,d}\}$ or $\{tfidf_{t,d}\}$, $t = 1, 2, \cdots, T$ and $d = 1, 2, \cdots, D$. This method tries to find linear combinations of terms that explain most of the variance of terms in documents, which can be understood as potential underlying and unobservable topics in the documents rather the words themselves. It is similar as principle component analysis when the target matrix is invertible. But the term frequency matrix is almost always singular since there are many more terms than the number of texts, i.e., $T \gg D$.

Singular value decomposition decomposes the term-document matrix X into the form of $USV^\top$. The columns of U and the columns of V are called the left-singular vectors and right-singular vectors of M, respectively. S is a diagonal matrix with all singular values arranged in a decreasing order. Figure 3.3 shows the interpretation of SVD in text mining. The columns of U can be understood as d hidden themes about the text corpus. For each theme, it is a combination of all terms and the element $u_{i,j}$ means the contribution of term i to theme j. Because the diagonal element of S is in decreasing order, all the themes are also in the order of decreasing importance, which means the first theme explains most variance of the original X and so on. On the other hand, matrix V captures the relationship between themes and text documents. The rows of $V^\top$ (columns of V) measure the importance of themes to each document. The element $v_{j,i}$ defines the contribution of theme i to document j. By arranging all text documents by time, a row of $V^\top$ is a time series of a theme, which is considered as our predictor and used to forecast exchange rates. For example, the highlighted row in V^T can be understood as the evolution of theme 1 during our sample period from 2010 to 2016.

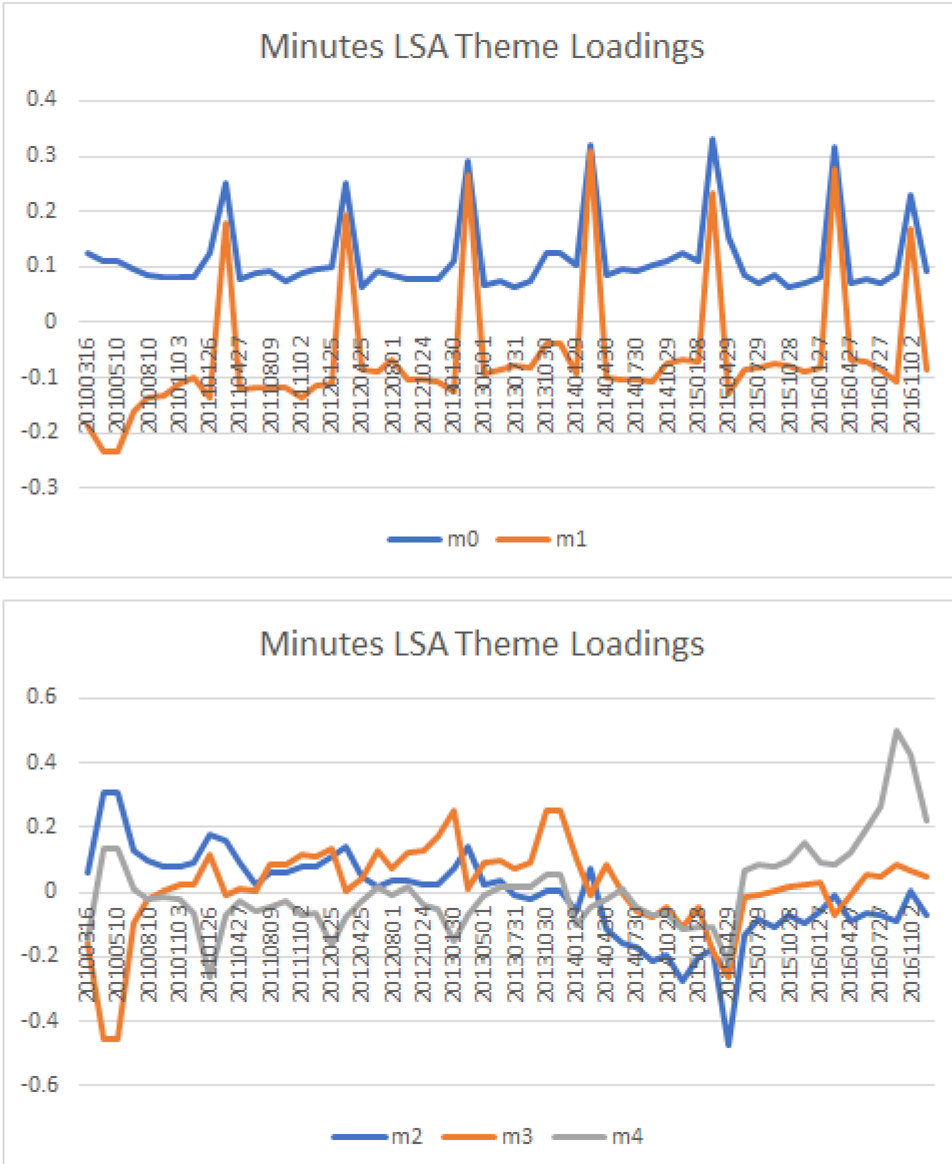
We apply LSA to extract the meanings from the minutes, which convey discussion of the Committee about why certain monetary policy is made and what the committee expects about the future. In this case, intuitively, LSA can extract several important hidden topics from the minutes. Meanings are more important than the words themselves. Because minutes are released three weeks after the meeting, we use minutes as predictor of the next meeting. Figure 3.4 shows the most five important themes extracted overtime after matching dates with statements. There is no easy interpretation about the underlying meaning of each theme. For the most two important themes, we might consider they capture the cyclic features of the economy. For other themes, they vary depending on the economic issues discussed by the Committee at certain point of time.

Vector Space Model Vector space model is used to measure the similarity between two text documents, which is the distance between two word vectors. Compared to simple Euclidean distance, it is more common to use cosine similarity to measure the similarity. Figure 3.5 illustrates the underlying reason. Suppose there are three word vectors from three text documents, which repeat only two terms a and b. Document 1 and 2 use term a and b in a similar proportion but d_1 is longer than d_2. Document 1 uses more term b but document 3 uses more term a. The Euclidean distance between d_1 and d_2 might be even larger than the distance between d_1 and d_3, which does not make much sense in text contents.

However, if we use cosine similarity, it can overcome such problem caused by the Euclidean distance. For the extreme case where document 1 only contains term a and document 3 contains only term b, the similarity is 0 according to cosine distance.

To leverage information from the press conference transcripts, we compare the content

Figure 3.4: LSA hidden themes of the minutes

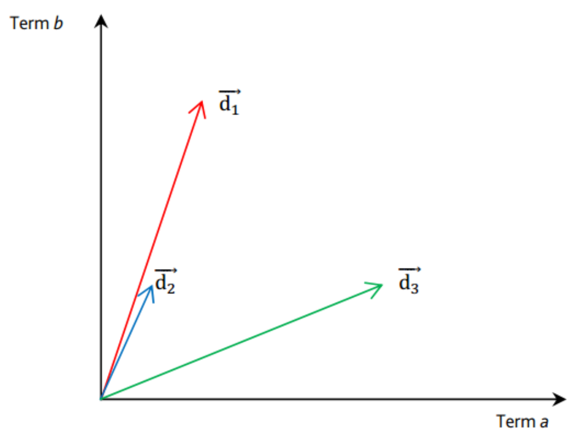


of press conference transcripts with the previous minutes and use this as one measure about mind change of the FOMC. The intuition is that journalists in the conference might ask questions related to last minutes because we cannot access to next minutes immediately and statement provides less information of the FOMC's discussion. The difference is calculated as cosine similarity between two documents vectors,

$$cos\left(\theta\right) = \frac{T_t M_{t-s}}{||T_t|| \cdot ||M_{t-s}||}, \tag{3.1}$$

where T_t denotes transcript vector at time t, M_{t-s} denotes the closest available minutes vector s period ago, and $|| \cdot ||$ is the L_2 norm.

Figure 3.5: Vector space model illustration



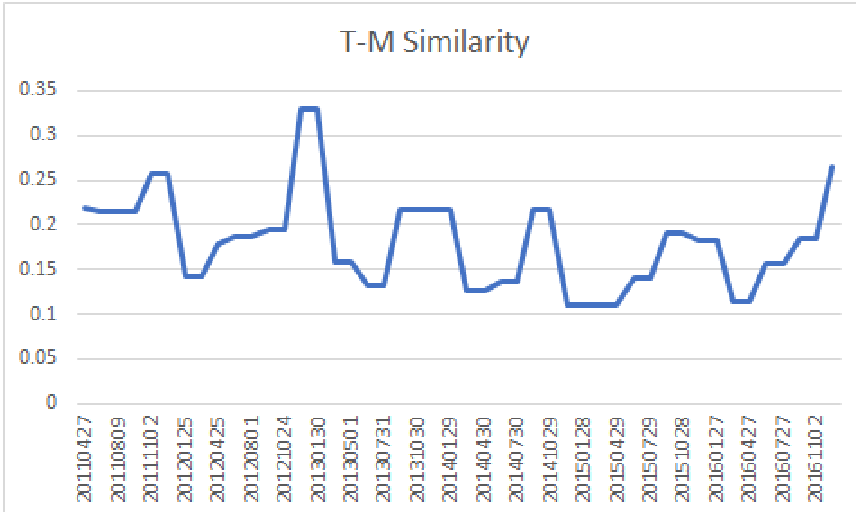
For dates before 20110427 when press conference transcripts are not available, $cos\left(\theta\right) = 0$. For days when there is no press conference, we use the previous similarity value to fill the gap, as shown in Figure 3.6.

3.4.3 FOMC Features

To summarize, based on different information in three types of FOMC text documents, we apply three methods in NLP to them, respectively.

We construct a term-document matrix with inverse-document-frequency adjustment weighting scheme from the statements, say X_1. We also extract the hidden themes from the minutes called X_2. The last feature is constructed from the comparison between press confer-

Figure 3.6: Similarity between minutes and transcripts



ence transcripts and the minutes, denoted as X_3. Our predictors are then $X = [X_1, X_2, X_3]$.

Before we feed the predictors to a machine learning algorithm, all features need to be normalized because they have different variability. Second, to eliminate extreme values, we winsorize all the features at 5% significance level, i.e., 2.5% for each tail. To be more specific, if the value is larger than 3, we fix it at 3 instead. It is also true for values less than -3.

3.4.4 Machine Learning Algorithms

In this paper, we only make directional forecast about the future exchange rates since we know it is not easy to forecast the exact number of change of returns. Exchange rate changes are labeled as either 1 for increasing or -1 for deceasing. The prediction problem can be viewed as a supervised classification.

The algorithm used is called support vector machine (SVM)[7]. SVM is often considered one of the best out of the box classifier since it tents to avoid overfitting. Roughly speaking, SVM is particularly useful when we have small sample size n but large predictor size p.

[7]Other types of machine learning algorithms such as Adaboost and neural network have been tested but do not outperform SVM.

SVM is indeed an optimization problem to locate the maximum margin hyperplane which is the separating hyperplane that is fastest from the training observations. We simply show the setup of SVM. For optimization details, we refer to Hastie, Tibshirani and Friedman (2009), Murphy (2012) and James et al. (2013). In short, we want to maximize the margin M

$$\max_{\beta_0,\beta_1,\cdots,\beta_p,\epsilon_1,\cdots,\epsilon_n} M \tag{3.2}$$

subject to

$$\sum_{j=1}^{p} \beta_j^2 = 1 \tag{3.3}$$

$$y_i f(x_i) \geq M(1-\epsilon_i) \tag{3.4}$$

$$\epsilon_i \geq 0 \tag{3.5}$$

$$\sum_{i=1}^{n} \epsilon_i \leq C \tag{3.6}$$

where $f(x_i)$ could be any function form including parametric or nonparametric, C is a non-negative tuning parameter, which can be determined by cross-validation, y_i is the labeled exchange rate changes, and X are the predictors we construct from the FOMC text documents.

3.5 Results

The whole sample contains 57 FOMC meetings and is divided into two subsamples: first period is from beginning to end of 2013 as the initial training sample and the testing sample starts from 2014. Training sample has 33 meetings and testing sample contains the rest 24 meetings. The out of sample test is based on expanding window, which means to in order to predict 20140319 return, we use whatever information up to that date. Our forecast horizon is from 1 week to 6 weeks with 1 week increment. We choose 6-week forecast horizon as the last because two meetings are about six weeks apart from each other and we forecast once for each meeting.

We forecast all nine individual currencies and also build one portfolio by assigning \$1,111 to "up" forecast and -\$1,111 to "down" forecast. By doing so, we limit our total risk to be \$10,000, which is the sum of absolute long and short dollar positions.

3.5.1 Forecast Success Ratios

Table 3.1 reports the forecast success ratios for all 9 currency pairs and the portfolio. We can see only one out of 9 currencies one has success ratio less than 50% for each forecast horizon up to three weeks. After that, the number of currencies with less than 50% success ratios increases. It might indicate that the information revealed from the FOMC documents becomes more apparent to the market as time passes. The portfolio has stable success ratios compared to individual currencies. Only 1-month forecast generates success ratio less than 50%. For the best performance with 1-week forecasting horizon, out of the 216 forecasts (24 meetings for all nine currencies), we correctly forecast the directions 125 times.

Table 3.1: Success ratios

	1w	2w	3w	4w	5w	6w
AUD	0.54	0.63	0.58	0.54	0.54	0.58
CAD	0.58	0.58	0.54	**0.42**	0.58	**0.46**
CHF	0.71	0.54	0.50	0.50	0.58	0.58
EUR	0.58	0.50	**0.38**	**0.42**	**0.46**	**0.46**
GBP	**0.42**	**0.46**	0.58	0.54	0.75	0.58
JPY	0.71	0.54	0.71	**0.46**	**0.33**	**0.42**
NOK	0.58	0.54	0.54	**0.46**	**0.46**	0.67
NZD	0.54	0.58	0.54	0.54	0.67	0.54
SEK	0.58	0.50	0.58	**0.42**	0.58	0.50
Portfolio	0.58	0.54	0.55	**0.48**	0.55	0.53

Notes: The success ratios are computed as the number of correct forecasts divided by the total number of out-of-sample forecasts. Here, we have 24 out-of-sample testing cases.

3.5.2 Evaluation of Forecasts

In this subsection, we formally test whether our forecasts are better than the random walk model or not. Here we use the weighted directional test proposed by Kim, Liao and Tornell (2014), which is more related to the profitability of our forecasts. Their test captures what George Soros observes: "Its not whether you're right or wrong, but how much money you make when you're right and how much you lose when you're wrong."

We consider the following test statistic:

$$T = \frac{1}{n - n_0 + 1} \sum_{t=n_0}^{n-1} D_{t,h}(P_{t+h} - P_t) \tag{3.7}$$

where n is the total number of sample points, n_0 is the number of initial training sample size, h is the forecast horizon, $D_{t,h}$ is the directional forecast at time t for horizon h, and P_t is the exchange rate at time t.

The forecast from the random walk model is P_t for any horizon h. Formally, the null hypothesis is

$$H_0 : E[D_{t,h}(P_{t+h} - P_t)] = 0 \tag{3.8}$$

which means that our directional forecasts are uncorrelated with future exchange rate changes. By martingale central limit theorem we have

$$\sqrt{n_1} V_T^{-1/2} T \to_d N(0, 1) \tag{3.9}$$

where V_T is the OS-LRV estimator whose formula can be found in the appendix. Since our out-of-sample testing size is not large, we need to take into consider the small sample effect and use OS-LRV estimator, which corrects this bias.

Figure 3.2 reports the result for weighted directional test. Even though eight out nine currencies have success ratios larger than 50% when the forecast horizon is less than 3 weeks, they have different performances under the weighted directional test. Six out of nine for 2-week forecasts, four out of nine for both 3-week and 5-week forecasts are significantly better than random walk model. The best result can be found in 1-week horizon forecast where eight out of nine currencies can beat the random walk model at 10% significance level. On the other hand, the portfolio performs well up to 3-week forecast horizon and also 5-week forecasting horizon. The portfolio for 1-week forecasting horizon has the second largest test statistic.

Table 3.2: Directional weighted test result

	1w	2w	3w	4w	5w	6w
AUD	2.12**	3.22***	1.97*	0.69	2.39**	1.59
CAD	2.91***	2.50***	-0.34	-1.47	2.91***	0.74
CHF	3.41***	2.37**	0.67	1.92*	-1.33	2.36**
EUR	1.85*	-1.20	-2.11	-0.60	0.95	0.75
GBP	-0.52	2.03*	2.78***	1.77*	3.06***	2.85***
JPY	1.90*	2.16**	2.00*	-2.31	-2.88	-1.63
NOK	1.73*	-0.93	-1.24	-0.61	-0.07	2.05*
NZD	2.21**	3.33***	-0.86	-0.95	3.25***	1.02
SEK	2.87***	-0.27	2.24**	-2.58	1.32	-0.03
Portfolio	3.04***	3.02***	2.51***	-1.06	3.29***	1.14

Notes: The table reports the t-statistics from the weighted directional test for each currency and each forecasting horizon. ***, ** and * represent the 1%, 5% and 10% significance levels, respectively.

3.5.3 Maximum Drawdowns

One other measure used in the financial industry is to compute the maximum drawdown for a investment strategy. A maximum drawdown is the maximum loss from a peak to a trough of a portfolio, before a new peak is attained. This measures the downside risk over the whole investment horizon.

Figure 3.7 shows the returns and maximum drawdown for each forecast after the FOMC meeting with forecast horizon $h = 1$ week. To illustrate the performance of our strategy,

Figure 3.7: Maximum drawdown for $h = 1$ and benchmark

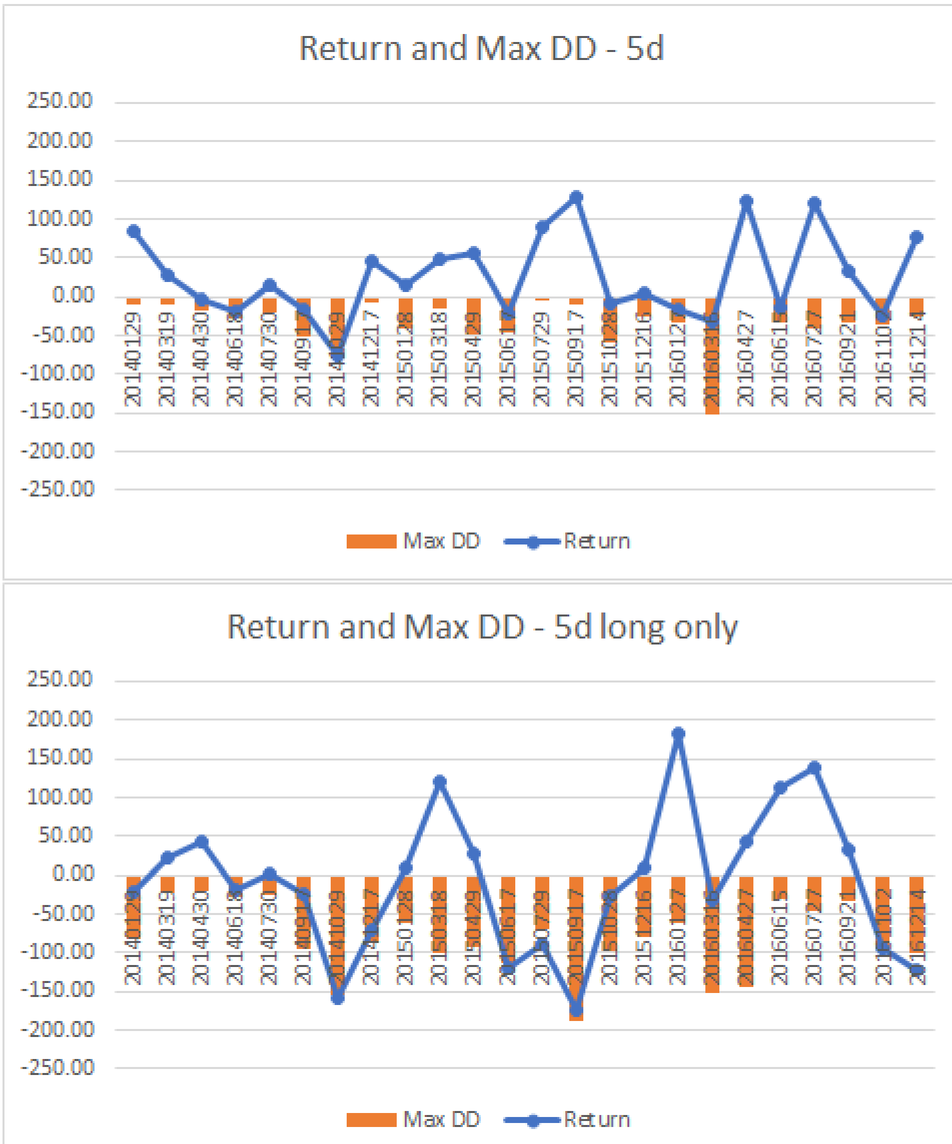
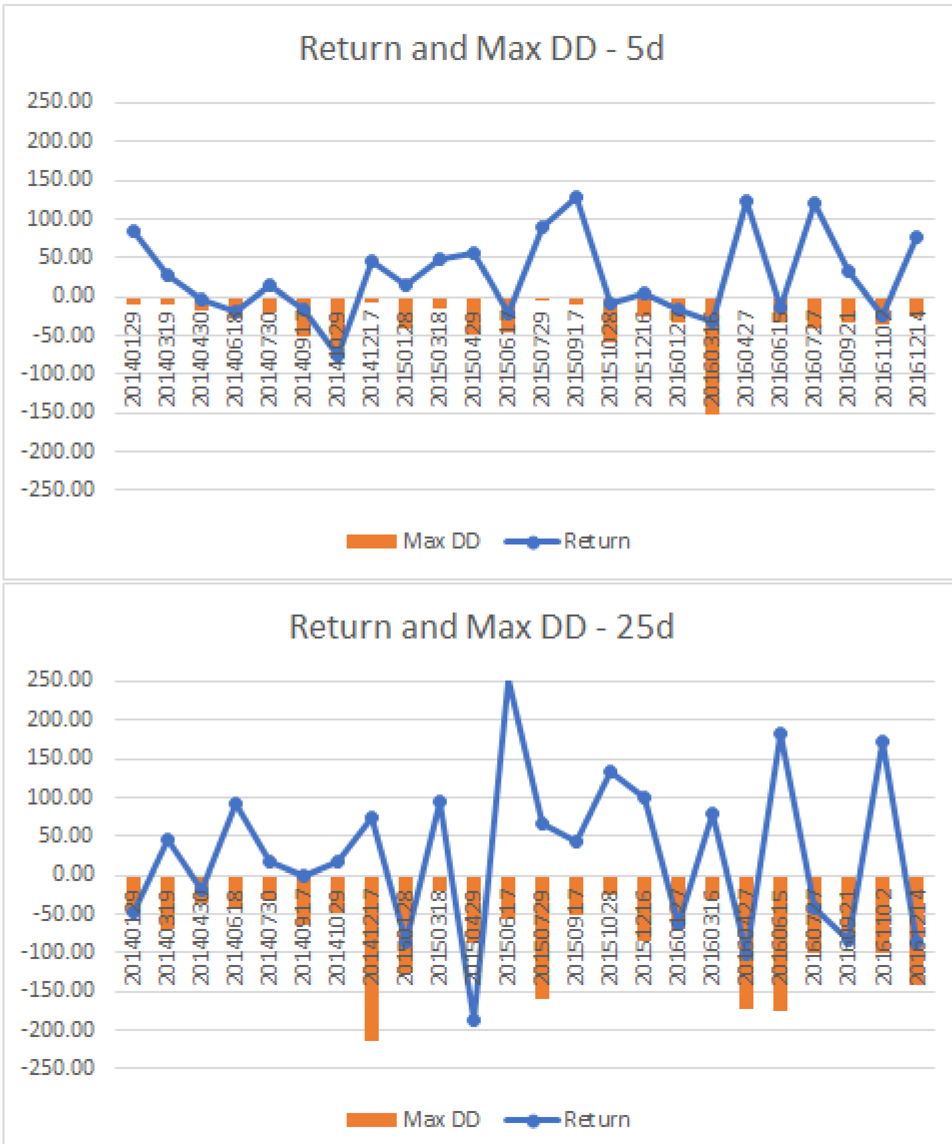


Figure 3.8: Maximum drawdown for $h = 1$ and $h = 5$



we construct a naive investment strategy which longs every currency equally for the same forecast horizon. This might be understood as the random walk model since we do not have any information about the future and there is half probability that the exchange rate going either direction. It is obvious not only our returns are more stable but the downside risk is much smaller than the simple naive strategy.

Another interesting result in Table 3.2 is our portfolio performs well for 5-week forecast horizon. We also compare its maximum drawdown with 1-week forecasting. Figure 3.8 shows the 5-week forecast has larger average return but maximum drawdown is larger for 5-week horizon than 1-week horizon

3.6 Conclusion

This chapter combines text mining techniques and machine learning algorithm to extract some nontraditional information from the FOMC text documents to forecast the future exchange rates among G10 currencies. Our result shows that this type of method can help us forecast in the very short-term period, which is different from the exchange rate forecast literature where we normally can forecast exchange rates in relatively longer horizons.

3.7 Appendix

3.7.1 Preprocessing Example

Let us consider the following example. Here is the second paragraph from the FOMC statement in 20161214, which says,

"Consistent with its statutory mandate, the Committee seeks to foster maximum employment and price stability. The Committee expects that, with gradual adjustments in the stance of monetary policy, economic activity will expand at a moderate pace and labor market conditions will strengthen somewhat further. Inflation is expected to rise to 2 percent over the medium term as the transitory effects of past declines in energy and import prices dissipate and the labor market strengthens further. Near-term risks to the economic outlook appear roughly balanced. The Committee continues to closely monitor inflation indicators and global economic and financial developments."

After the first two steps, it becomes

consistent statutory mandate seeks foster maximum employment price stability expects gradual adjustments stance monetary policy economic activity expand moderate pace labor market conditions strengthen somewhat further inflation expected rise percent medium term transitory effects past declines energy import prices dissipate labor market strengthens further nearterm risks economic outlook appear roughly balanced continues closely monitor inflation indicators global economic financial developments

The last step would require us to stem words. `strengthens` and `strengthen` become `strengthen` after `s` is removed,

consist statutori mandat seek foster maximum employ price stabil expect gradual adjust stanc monetari polici econom activ expand moder pace labor market condit strengthen somewhat inflat expect rise percent medium term transitori effect past declin energi import price dissip labor market strengthen nearterm risk econom outlook appear rough balanc continu close monitor inflat indic global econom financi develop

3.7.2 OS-LRV Estimator

For a weakly dependent process $\{W_t\}_{t=1}^n$, we define

$$\Lambda_{2m-1} = \frac{1}{n}\sum_{t=1}^{n}\phi_{2m-1}(\frac{t}{n})W_t \tag{3.10}$$

$$\Lambda_{2m} = \frac{1}{n}\sum_{t=1}^{n}\phi_{2m}(\frac{t}{n})W_t \tag{3.11}$$

for $m = 1, 2, \cdots, M/2$, where M is any fixed even integer and

$$\phi_{2m-1} = \sqrt{2}\cos(2m\pi x) \text{ and } \phi_{2m} = \sqrt{2}\sin(2m\pi x). \tag{3.12}$$

The OS-LRV estimator is defined as

$$\Sigma_n(M) = \frac{1}{M}\sum_{m=1}^{M/2}(\Lambda_{2m-1}^2 + \Lambda_{2m}^2). \tag{3.13}$$

We have the result that

$$\Sigma_n^{-1/2} n^{-1/2}\sum_{t=1}^{n} W_t \to_d t(M). \tag{3.14}$$

For more details, we refer to Kim, Liao and Tornell (2014), Phillips (2005) and Sun (2014).

4. References

Abel, Andrew B. 1990. "Asset Prices under Habit Formation and Catching up with the Joneses." *American Economic Review* 80(2):38–42.

Acosta, Miguel. 2015. "FOMC Responses to Calls for Transparency." *Working Paper* .

Alvarez, Fernando, Andrew Atkeson and Patrick J. Kehoe. 2009. "Time-Varying Risk, Interest Rates, and Exchange Rates in General Equilibrium." *Review of Economic Studies* 76(3):851–878.

Bacchetta, Philippe and Eric van Wincoop. 2006. "Can Information Heterogeneity Explain the Exchange Rate Determination Puzzle?" *American Economic Review* 96(3):552–576.

Backus, David K., Silverio Foresi and Chris I. Telmer. 2001. "Affine Term Structure Models and the Forward Premium Anomaly." *Journal of Finance* 56(1):279–304.

Bansal, Ravi and Amir Yaron. 2004. "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles." *Journal of Finance* 59(4):1481–1509.

Bansal, Ravi and Ivan Shaliastovich. 2013. "A Long-Run Risks Explanation of Predictability Puzzles in Bond and Currency Markets." *Review of Financial Studies* 26(1):1–33.

Bansal, Ravi and Magnus Dahlquist. 2000. "The Forward Premium Puzzle: Different Tales From Developed and Emerging Economies." *Journal of International Economics* 51(1):115–144.

Barro, Robert J. 2006. "Rare Disasters and Asset Markets in the Twentieth Century." *Quarterly Journal of Economics* 121(3):823–866c.

Bekaert, Geert. 1996. "The Time Variation of Risk and Return in Foreign Exchange Markets: A General Equilibrium Perspective." *Review of Financial Studies* 9(2):427–470.

Bilson, John F. O. 1981. "The "Speculative Efficiency" Hypothesis." *Journal of Business* 54(3):435–451.

Boukus, Ellyn and Joshua V. Rosenberg. 2006. "The Information Content of FOMC Minutes." *Working Paper* .

Brunnermeier, Markus K., Stefan Nagel and Lasse H. Pedersen. 2008. "Carry Trades And Currency Crashes." *NBER Macroeconomics Annual* 23:313–347.

Burnside, Craig, Bing Han, David Hirshleifer and Tracy Y. Wang. 2011. "Investor Overconfidence and the Forward Premium Puzzle." *Review of Economic Studies* 78:523–558.

Burnside, Craig, Martin Eichenbaum, Isaac Kleshchelski and Sergio Rebelo. 2006. "The Returns to Currency Speculation." *NBER Working Paper No. 12489* .

Burnside, Craig, Martin Eichenbaum, Isaac Kleshchelski and Sergio Rebelo. 2011. "Do Peso Problems Explain the Returns to the Carry Trade?" *Review of Financial Studies* 24(3):853–891.

Campbell, John Y. and John Cochrane. 1999. "Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior." *Journal of Political Economy* 107(2):205–251.

Campbell, John Y. and Robert J. Shiller. 1988. "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors." *Review of Financial Studies* 1(3):195–228.

Chaboud, Alain P. and Jonathan H. Wright. 2005. "Uncovered Interest Parity: It Works, But Not for Long." *Journal of International Economics* 66(2):349–362.

Chinn, Menzie D. 2006. "The (partial) rehabilitation of interest rate parity in the floating rate era: Longer horizons, alternative expectations, and emerging markets." *Journal of International Money and Finance* 25(1):7–21a.

Chinn, Menzie D. and Guy Meredith. 2005. "Testing Uncovered Interest Parity at Short and Long Horizons during the Post-Bretton Woods Era." *Working Paper* .

Clark, Todd E. and Kenneth D. West. 2006. “Using out-of-sample mean squared prediction errors to test the martingale difference hypothesis.” *Journal of Econometrics* 135:155–186.

Colacito, Riccardo and Mariano M. Croce. 2011. “Risks for the Long Run and the Real Exchange Rate.” *Journal of Political Economy* 119(1):153–181.

Colacito, Riccardo and Mariano M. Croce. 2013. “International Asset Pricing with Recursive Preferences.” *Journal of Finance* 68:2651–2686.

Da, Zhi, Joseph Engelberg and Pengjie Gao. 2011. “In Search of Attention.” *Journal of Finance* 66(5):1461C1499.

Da, Zhi, Joseph Engelberg and Pengjie Gao. 2015. “The Sum of All FEARS Investor Sentiment and Asset Price.” *Review of Financial Studies* 28(1):1–32.

Eichenbaum, Martin and Charles Evans. 1995. “Some Empirical Evidence on the Effects of Shocks to Monetary Policy on Exchange Rates.” *Quarterly Journal of Economics* 110(4):975–1009.

Engel, Charles. 1996. “The Forward Discount Anomaly and the Risk Premium: A Survey of Recent Evidence.” *Journal of Empirical Finance* 3(2):123–192.

Engel, Charles. 2014. “Exchange Rates and Interest Parity.” *Handbook of International Economics* 4:453–522.

Engel, Charles. 2016. “Exchange Rates, Interest Rates, and the Risk Premium.” *American Economic Review* 106(2):436–474.

Engel, Charles and Kenneth D. West. 2005. “Exchange Rates and Fundamentals.” *Journal of Political Economy* 113(3):485–517.

Epstein, Larry G. and Stanley E. Zin. 1989. “Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework.” *Econometrica* 57(4):937–969.

Fama, Eugene F. 1984. “Forward and Spot Exchange Rates.” *Journal of Monetary Economics* 14(3):319–338.

Fama, Eugene F. and James D. MacBeth. 1973. “Risk, Return, and Equilibrium: Empirical Tests.” *Journal of Political Economy* 81(3):607–636.

Farhi, Emmanuel and Xavier Gabaix. 2016. “Rare Disasters and Exchange Rates.” *Quarterly Journal of Economics* 131(1):1–52.

Faust, Jon and John H. Rogers. 2003. “Monetary Policy’s Role in Exchange Rate Behavior.” *Journal of Monetary Economics* 50(7):1403–1424.

Froot, Kenneth A. 1990. “Short Rate and Expected Asset Returns.” *NBER Working Paper No. 3247* .

Froot, Kenneth A. and Richard H. Thaler. 1990. “Anomalies: Foreign Exchange.” *Journal of Economic Perspectives* 4(3):179–192.

Gourinchas, Pierre-Olivier and Aaron Tornell. 2004. “Exchange Rate Puzzles and Distorted Beliefs.” *Journal of International Economics* 64(2):303–333.

Hansena, Stephen and Michael McMahon. 2016. “Shocking Language: Understanding the Macroeconomic Effects of Central Bank Communication.” *Journal of International Economics* 99:S114–S133.

Harrison, J. Michael and David M Kreps. 1979. “Martingales and Arbitrage in Multiperiod Securities Markets.” *Journal of Economic Theory* 20(3):381–408.

Hastie, Trevor, Robert Tibshirani and Jerome Friedman. 2009. *The Elements of Statistical Learning*. Springer.

Hoberg, Gerard and Gordon Phillips. 2016. “Text-Based Network Industries and Endogenous Product Differentiation.” *Journal of Political Economy* 124(5):1423–1465.

Huisman, Ronald, Kees Koedijk, Clemens Kool and Franqois Nissen. 1998. “Extreme Support for Uncovered Interest Iarity.” *Journal of International Money and Finance* 17(1):211–228.

James, Gareth, Daniela Witten, Trevor Hastie and Robert Tibshirani. 2013. *An Introduction to Statistical Learning with Applications in R.* Springer.

Jurek, Jakub W. 2014. “Crash-Neutral Currency Carry Trades.” *Journal of Financial Economics* 113(325-347).

Kim, Young Ju, Zhipeng Liao and Aaron Tornell. 2014. “Speculators’ Positions and Exchange Rate Forecasts: Beating Random Walk Models.” *Working Paper* .

Laeven, Luc and Fabian Valencia. 2008. “Systemic Banking Crises: A New Database.” *IMF Working Paper* .

Laeven, Luc and Fabian Valencia. 2012. “Systemic Banking Crises Database: An Update.” *IMF Working Paper* .

Lewis, Karen K. 1995. “Puzzles in International Financial Markets.” *Handbook of International Economics* III:1913–1971.

Li, Ming and Aaron Tornell. 2015. “Exchange Rates Under Robustness: An Account of the Forward Premium Puzzle.” *Working Paper* .

Lothian, James R. and Liuren Wu. 2011. “Uncovered Interest-Rate Parity over the Past Two Centuries.” *Journal of International Money and Finance* 30(3):448–473.

Loughran, Tim and Bill McDonald. 2011. “When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks.” *Journal of Finance* 66(1):35–65.

Lucas, Robert E. 1978. “Asset Prices in an Exchange Economy.” *Econometrica* 46(6):1429–1445.

Lustig, Hanno and Adrien Verdelhan. 2007. “The Cross Section of Foreign Currency Risk Premia and Consumption Growth Risk.” *American Economic Review* 97(1):89–117.

Mark, Nelson C. 1995. “Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability.” *American Economic Review* 85:201–218.

Meese, Richard and Kenneth Rogoff. 1983. “Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?” *Journal of International Economics* 14:3–24.

Molodtsova, Tanya, Alex Nikolsko-Rzhevskyy and David H. Papell. 2008. “Taylor Rules with Real-Time Data: A Tale of Two Countries and One Exchange Rate.” *Journal of Monetary Economics* 55:S63–S79.

Molodtsova, Tanya and David H. Papel. 2009. “Out-of-Sample Exchange Rate Predictability with Taylor Rule Fundamentals.” *Journal of International Economics* 77(2):167–180.

Murphy, Kevin P. 2012. *Machine Learning A Probabilistic Perspective*. The MIT Press.

Newey, Whitney K. and Kenneth West. 1994. “Automatic Lag Selection in Covariance Matrix Estimation.” *Review of Economic Studies* 61(4):631–653.

Phillips, Peter C. B. 2005. “HAC Estimation by Automated Regression.” *Econometric Theory* 21(1):116–142.

Ranciere, Romain and Aaaron Tornell. 2015. “Why Do Reforms Occur in Crises Times?” *Working Paper* .

Reinhart, Carmen M. and Kenneth S. Rogoff. 2009. *This Time Is Different: Eight Centuries of Financial Folly*. Princeton University Press.

Rossi, Barbara. 2013. “Exchange Rate Predictability.” *Journal of Economic Literature* 51(4):1063–1119.

Stathopoulos, Andreas. 2012. “Asset Prices and Risk Sharing in Open Economies.” *Working Paper* .

Sun, Yixiao. 2014. “Lets fix it: Fixed-b asymptotics versus small-b asymptotics in heteroskedasticity and autocorrelation robust inference.” *Journal of Econometrics* 178:659C677.

Taylor, John B. 1993. "Discretion Versus Policy Rules in Practice." *Carnegie-Rochester Conference Series on Public Policy* 39:195–214.

Tetlock, Paul C. 2007. "Giving Content to Investor Sentiment: The Role of Media in the Stock Market." *Journal of Finance* 62(3):1139–1168.

Tetlock, Paul C. 2011. "All the News Thats Fit to Reprint: Do Investors React to Stale Information?" *Review of Financial Studies* 24:1481–1512.

Verdelhan, Adrien. 2010. "A Habit-Based Explanation of the Exchange Rate Risk Premium." *Journal of Finance* 65(1):123–146.

Verdelhan, Adrien. 2013. "The Share of Systematic Risk in Bilateral Exchange Rates." *Working Paper* .

Yu, Jianfeng. 2013. "A Sentiment-based Explanation of the Forward Premium Puzzle." *Journal of Monetary Economics* 60(4):474–491.