Title
Deviations from Rational Expectations and Asset Prices

Author
Mokanov, Denis

Publication Date
2023

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Deviations from Rational Expectations and Asset Prices

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Management

by

Denis Murtaza Mokanov

2023

ABSTRACT OF THE DISSERTATION

Deviations from Rational Expectations and Asset Prices

by

Denis Murtaza Mokanov
Doctor of Philosophy in Management
University of California, Los Angeles, 2023
Professor Lars A. Lochstoer, Chair

In Chapter 1, I document a novel result regarding the uncovered interest rate parity (UIP) puzzle: investing in high interest rate currencies does not yield positive excess returns during recessions. That is, the UIP holds in bad times. This new finding is a challenge to existing rational expectations models that address the UIP puzzle. A model featuring investors whose interest rate expectations are distorted by extrapolation bias and time-varying stickiness is able to quantitatively account for this evidence when calibrated to available survey data. The model also generates predictions for bond return predictability, the profitability of time-series momentum in the foreign exchange and fixed income markets, and foreign exchange predictability during the post-2007 period, which are borne out in the data.

In Chapter 2 (with Gabriel Cuevas Rodriguez and Danyu Zhang), we document three stylized facts related to equity analysts' earnings expectations: (1) consensus earnings expectations underreact to news unconditionally, (2) the degree of underreaction declines during high-volatility periods, and (3) the degree of underreaction declines over our sample. To account for these findings, we develop a simple model featuring rational inattention. We show that our model is able to account for the unconditional profitability of momentum, momentum crashes, and the diminishing profitability of momentum over our sample. Based on the predictions of our model, we propose a trading strategy that mixes short-run and long-

run momentum signals and show that the resultant mixed momentum strategy outperforms conventional long-run momentum strategies. Finally, we use a machine learning algorithm to estimate the predictable component of earnings surprises and construct a portfolio that is long (short) on stocks with excessively pessimistic (optimistic) earnings expectations. The resultant trading strategy generates an annualized Sharpe ratio of about 1.16 and its returns are not explained by popular factor models.

The dissertation of Denis Murtaza Mokanov is approved.

Tyler Stewart Muir

Stavros Panageas

Mikhail Chernov

Lars A. Lochstoer, Committee Chair

University of California, Los Angeles

2023

To my family.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank to my doctoral committee members, Mikhail Chernov, Stavros Panageas, Tyler Muir, and especially my advisor, Lars Lochstoer, for their guidance. Other faculty members who have positively impacted my time at UCLA include Avanidhar Subrahmanyam, Valentin Haddad, Barney Hartman-Glaser, Bernard Herskovic, Francis Longstaff, William Mann, Ivo Welch, and William Zame.

I am grateful to my cohort mates, Paul Hübner, Edward Kim, Paul P. Momtaz, and James O'Neill for their camaraderie. I have also had the pleasure of working alongside a talented group of peers that includes Chady Gemayel, Mahyar Kargar, Geoffery Zheng, Yu Shi, Clinton Tepper, Salil Gadgil, Ljubica Georgievska, and Alice Meng.

Special thanks go out to my coauthors Gabriel Cuevas Rodriguez and Danyu Zhang who contributed heavily to the second chapter of this dissertation.

VITA

2012	B.B.A., University of Michigan, Ann Arbor, MI.
2016	M.Sc., Accounting and Financial Management, King's College London, London, UK.

CHAPTER 1

Deviations from Rational Expectations and the Uncovered Interest Rate Parity Puzzle

1.1 Introduction

According to traditional frameworks within macroeconomics, exchange rates evolve in a way that equalizes the expected returns of investing in risk-free assets denominated in different currencies. This prediction is known as the uncovered interest rate parity (UIP) condition. The consensus in the literature is that the UIP condition is violated in that borrowing in low interest rate currencies and investing in high interest rate currencies tends to deliver positive excess returns. This finding has given rise to one of the most widely-studied questions in finance: the UIP puzzle.[1]

In this paper, I provide novel evidence related to the UIP puzzle by conditioning the classic foreign exchange return predictability regression on the state of the business cycle. The conditional tests in this paper are motivated by recent findings regarding the conditional dynamics of stock (Gómez-Cram, 2022) and bond (Andreasen, Engsted, Møller, and Sander, 2020; Borup, Eriksen, Kjær, and Thyrsgaard, 2023) return predictability over the business cycle. The conditional tests reveal the following patterns in the data: higher interest rate currencies tend to carry high excess returns during expansions. However, the relation be-

[1]Hansen and Hodrick (1980) and Fama (1984) are examples of early work related to the UIP puzzle. See Engel (2014) for a comprehensive review of the literature.

tween interest rates and excess returns breaks down during recessions, i.e. investing in high interest rate currencies does not yield positive excess returns during economic downturns.

This novel result represents a challenge for workhorse rational expectations models (or straightforward extensions thereof). Rational expectations models fail to account for the conditional patterns in foreign exchange return predictability as they rely on the coexistence of pro-cyclical interest rates and counter-cyclical risk premia to explain the unconditional failure of the UIP.

Given the inability of existing models to account for the patterns of UIP violations over the business cycle, I consider an alternative explanation for the UIP puzzle based on Gourinchas and Tornell (2004): I hypothesize that UIP violations are driven by interest rate expectations that deviate from full-information rational expectations (FIRE). The idea that systematic expectation errors cause the UIP to fail has garnered renewed attention in the literature thanks to the proliferation of survey data that allows researchers to directly test the rational expectations hypothesis.[2]

I begin my analysis by showing that forecasters' expectations of short rates exhibit patterns that are qualitatively consistent with the return predictability patterns. In order to test the rationality of forecasters' expectations, I adopt the methodology developed by Coibion and Gorodnichenko (2015), which involves examining the ability of average (consensus) expectation revisions to predict consensus expectation errors. I use data from two unrelated sources: the FX4Casts survey and the Survey of Professional Forecasters (SPF) to show that expectations underreact to interest rate innovations during expansions and that FIRE cannot be rejected during recessions.

Next, I incorporate distorted interest rate expectations into a simple present-value model, discipline the extent of belief distortions using survey data and show that the model is capable of reproducing the patterns in foreign exchange return predictability presented in this paper.

[2]The idea that UIP violations are driven by deviations from rational expectations dates back to Froot and Thaler (1990). Recent papers that rely on deviations from rational expectations to explain the UIP puzzle include Candian and De Leo (2022), Valente, Vasudevan, and Wu (2022), and Molavi, Tahbaz-Salehi, and Vedolin (2021).

The interest rate expectations in the model are distorted by two effects, whose relative magnitude varies throughout the business cycle. The first effect, which, *ceteris paribus*, leads consensus expectations to underreact to interest rate news is rational inattention (Mankiw and Reis, 2002): some market participants choose to optimally remain inattentive to interest rate shocks, which generates expectation stickiness. I model the fraction of inattentive agents as a state-dependent variable with the stickiness of expectations declining significantly during recessions. Intuitively, recessions are associated with heightened uncertainty and heightened uncertainty generates stronger incentives to acquire information. The time-varying incentives to acquire information produce counter-cyclical expectation stickiness (similar to the models in Mäkinen and Ohl, 2015 and Flynn and Sastry, 2021).

The second effect that distorts expectations is (over)extrapolation. Extrapolation causes expectations to overreact to news. In my model, agents who incorporate an interest rate shock into their information sets expect the innovations to have a longer-lasting impact than they actually do (similar to Angeletos, Huo, and Sastry, 2021). I assume that extrapolation is driven by psychological biases that are unaffected by business cycle conditions.[3]

The combination of the two effects generates the following results: during expansions, the rational inattention effect dominates the extrapolation effect and we observe underreaction to interest rate news in the short-run (à la Brooks, Katz, and Lustig, 2018). The underreaction to interest rate news, in turn, drives the failure of the UIP condition. To gain an insight into the mechanism underlying the UIP violations, consider a positive domestic interest rate shock. In response to the shock, the domestic currency appreciates contemporaneously. However, the initial appreciation is only moderate due to the presence of inattentive agents. In subsequent periods, the domestic currency continues to appreciate driven by the increasing fraction of agents who incorporate the shock into their information sets. A strong rational inattention effect leads to the coexistence of high domestic interest rates and an appreciating domestic currency, thus generating violations of the UIP condition. Conversely, during recessions, the rational inattention effect becomes weaker and is roughly canceled out by the

[3]See Afrouzi, Kwon, Landier, Ma, and Thesmar (2020) for experimental evidence regarding extrapolative belief formation.

extrapolation effect. The convergence of consensus interest rate expectations to FIRE drives the lack of return predictability during recessions.

Having shown that my model generates results consistent with the patterns in foreign exchange return predictability, I explore several additional asset pricing implications of the model. First, I show that the model is consistent with the bond return predictability patterns identified by Andreasen et al. (2020), who show that the correlation between excess bond returns and yield spreads reverses from positive during expansions to negative during recessions. Second, I show that the predictions of the model regarding the profitability of time-series momentum (Moskowitz, Ooi, and Pedersen, 2012) are borne out in the data: time-series momentum strategies in the foreign exchange and fixed income markets fail to deliver statistically positive returns following recessions. Lastly, I show that the results in recent papers, which document a reversal of the sign of the UIP coefficient during the post-2007 period (e.g. Engel, Kazakova, Wang, and Xiang, 2021 and Bussiere, Chinn, Ferrara, and Heipertz, 2022), are an artifact of the inclusion of the Great Recession (2008-2009) and the COVID-19 crisis in the post-2007 sample. Generally speaking, the post-2007 period is not fundamentally different from other historical episodes and the differences documented by these papers are driven by the length and severity of the two recessions within the sample.

This paper contributes to several strands of the literature, both theoretical and empirical. In terms of documenting the state dependence of UIP violations, this paper is most closely related to Clarida, Davis, and Pedersen (2009) and Bansal and Dahlquist (2000). Clarida et al. (2009) condition the tests of the UIP on foreign exchange volatility and show that the sign of the foreign exchange return predictability coefficient flips during high volatility periods. Bansal and Dahlquist (2000) show that failures of the UIP condition are confined to advanced economy currencies and to states in which the US interest rate exceeds foreign interest rates. Bansal and Dahlquist (2000) also show that the violations of the UIP condition are less severe for economies with higher inflation uncertainty. My paper is also related to the empirical literature, dating back to Frankel and Froot (1987), that uses survey data on exchange rate expectations to study UIP violations.[4]

This paper also contributes to the empirical literature that documents state-dependent expectation stickiness. In related work, Coibion and Gorodnichenko (2015) show that the expectation stickiness of SPF respondents declines significantly during US Recessions. Loungani, Stekler, and Tamirisa (2013) show, for a large panel of advanced and emerging market economies, that forecasters increase the rate at which they incorporate news into their forecasts as the economy enters a recession. Andrade and Le Bihan (2013) document that the Great Recession is associated with increased attentiveness to unemployment, real GDP, and inflation news among professional forecasters surveyed by the European Central Bank.

In terms of its theoretical contributions, this paper is closely related to Gourinchas and Tornell (2004), who propose a model that features a boundedly-rational representative agent who misperceives the relative importance of transitory and persistent interest rate shocks and show that their model is capable of accounting for the UIP puzzle. More recently, Candian and De Leo (2022) and Valente et al. (2022) propose extensions of the Gourinchas and Tornell (2004) framework that feature delayed overreaction to interest rate news and show that their models are capable of accounting for several international finance puzzles. Molavi et al. (2021) develop a model, which features agents with limited information processing capacity and show that their model is capable of reproducing several asset pricing puzzles, including the UIP puzzle. Bacchetta and van Wincoop (2010) and Bacchetta and van Wincoop (2021) propose variations of a model in which UIP violations are driven by costly portfolio rebalancing. A common feature among the aforementioned models is that they aim to capture the unconditional failure of the UIP condition and lack mechanisms that generate state-dependent return predictability.

Finally, the model in this paper shares insights with a large number of models featuring time-varying incentives to acquire costly information, e.g. Reis (2006) and Mackowiak and Wiederholt (2009).

[4]Recent examples of papers in Bussiere et al. (2022) and Kalemli-Özcan and Varela (2021)

The rest of the paper is organized as follows. In Section 1.2, I describe the data I use in the empirical section of this paper. In Section 1.3, I demonstrate that the UIP is rejected during expansions but not during recessions. I also show that forecasters' expectations underreact to interest rate news only during expansions. In section 1.4, I build a parsimonious present-value model that captures the patterns in foreign exchange return predictability. In Section 1.5, I discuss the implications of the model for the UIP puzzle, bond return predictability, time-series momentum, and post-2007 foreign exchange return predictability. Section 1.6 concludes the paper.

1.2 Data

1.2.1 Foreign Exchange Data

I retrieve daily spot and forward exchange rates from Barclays Bank International and Reuters (via Datastream). My sample includes the G10 currencies: Australian dollar (AUD), Canadian dollar (CAD), Danish krone (DKK), euro (EUR), Japanese yen (JPY), Norwegian krona (NOK), New Zealand dollar (NZD), Swedish krona (SEK), Swiss franc (CHF), pound sterling (GBP), and US dollar (USD). I let USD be the domestic currency and express all exchange rates in USD per unit of foreign currency. For the CAD, DKK, EUR (spliced with the German mark before 1999), JPY, NOK, SEK, CHF, and GBP, the sample begins in January 1983; for the AUD and NZD, the sample starts in January 1985 due to data availability. In order to avoid possible issues related to currency pegs, I drop the DKK from my analysis after January 1999 and the CHF between September 2011 and January 2015. I also obtain daily interbank rates from Datastream for the period between January 2008 and September 2020. I use OIS rates for the CAD, EUR, JPY, GBP, and USD, I use the LIBOR rate for the CHF and local interbank rates for the remainder of the currencies in my sample.

I convert the daily interest rate and exchange rate data into monthly data by using the data from the last trading day of each month. For the period between 1983 and 2007, I compute implied interest rate differentials using the covered interest rate parity (CIP) condition: $i_t^D = f_t - s_t$ where s_t denotes the log spot exchange rate and f_t denotes the log forward exchange rate. This is a standard practice in the literature because the forward market is deep and liquid. Due to the deviations from the CIP after the Great Recession documented by Du, Tepper, and Verdelhan (2018), I use actual interest rate differentials for the period between 01/2008 and 09/2020.

1.2.2 Survey Data

I utilize two sources of survey data to study forecast error predictability. The first source I use is survey data regarding the 3-month US T-Bill rate from the Survey of Professional Forecasters (SPF) conducted by the Philadelphia Fed. The SPF data is available at a quarterly frequency beginning in 1983:Q1. The SPF respondents forecast the average value of the 3-month T-Bill rate throughout a quarter.

The second source is survey data from FX4Casts. FX4Casts interest rate forecast data is available beginning in 10/2001. I restrict my sample to currencies for which interbank rate forecasts are available: USD, JPY, EUR, CHF, and GBP (LIBOR), SEK (Stibor), and NOK (Nibor). The FX4Casts data is available at the monthly frequency and contains the average interest rate forecasts of large financial institutions that are active participants in the foreign exchange market. Survey respondents forecast the value of interest rates at the end of each month. I construct interest rate differential forecasts by subtracting the interest rate forecast for each currency from the USD interest rate forecast.

1.2.3 Macroeconomic Variables

I use several macroeconomic variables to determine recession periods. I download the real-time US monthly industrial production index series from the Philadelphia Fed and the monthly Chicago Fed National Activity Index (CFNAI) data from the Chicago Fed.

I obtain the NBER recession dates, the Chauvet and Piger (2008) recession probabilities, and the OECD Major Seven and OECD and Non-member Economies recession indicators from the St. Louis Fed.

1.3 Empirical Findings

1.3.1 UIP Puzzle

Standard models in international finance imply that the returns of investing in risk-free assets denominated in different currencies are equalized. This condition, known as the uncovered interest parity (UIP) condition, implies that exchange rate depreciation offsets gains related to interest rate differentials. Therefore,

$$\mathbb{E}_t\left[s_{t+1}\right] - s_t = i_t - i_t^*$$

where s_t is the log exchange rate (in terms of domestic currency (USD) per unit of foreign currency), and i_t^* and i_t are the foreign and domestic interest rates, respectively. I define $i_t^D \equiv i_t - i_t^*$ as the interest rate differential.

1.3.2 Return Predictability Regressions

Following Fama (1984), the UIP condition is tested by examining whether excess foreign exchange returns are predictable using interest rate differentials:

$$r_{t+1} = \alpha + \beta^F i_t^D + \varepsilon_{t+1} \tag{1.1}$$

where the log excess return, r_{t+1}, is defined as:

$$r_{t+1} = s_{t+1} - s_t - i_t^D \tag{1.2}$$

A large number of papers reject the null of $\beta^F = 0$ in favor of a negative β^F, which implies that higher interest rates are associated with higher returns. The puzzle associated with $\beta^F < 0$ is called the "UIP puzzle" in the literature.

In this paper, I consider a version of the UIP regression which allows for time varying α and β^F coefficients. In particular, I focus on the case of coefficients that switch between recessions and normal times:[5]

$$r_{t+1} = \alpha_{\exp} + \alpha_{\Delta}\mathbf{1}_{\text{rec}|\text{t}} + \left(\beta^F_{\exp} + \beta^F_{\Delta}\mathbf{1}_{\text{rec}|\text{t}}\right) i^D_t + \varepsilon_{t+1} \tag{1.3}$$

where $\beta^F_{\exp}$ is the UIP coefficient during expansions and $\mathbf{1}_{\text{rec}|\text{t}}$ is an indicator variable which takes on the value of 1 if period t is a recession period and 0 otherwise. The UIP coefficient estimated during recessions, β^F_{rec}, is equal to $\beta^F_{\exp} + \beta^F_{\Delta}$.

Table 1.1 reports the results of regressions (1.1) and (1.3) estimated using a panel of developed-economy currencies. I adjust standard errors for heteroskedasticity, serial correlation, and cross-country correlation following Driscoll and Kraay (1998). I also report standard errors obtained using a block bootstrap. The UIP is usually estimated with currency fixed effects in the literature. In this paper, I report the results of tests without and without currency fixed effects.

Before I go into the specifics of my analysis, a note on how I define recessions is in order. Well-diversified financial institutions with global asset positions are likely to serve as marginal investors in the foreign exchange market (Haddad and Muir, 2021; He, Kelly, and Manela, 2017). Therefore, in a perfect world, all tests in this paper would use a real-time global recession indicator.[6] However, to my knowledge, a commonly accepted global recession indicator does not exist. Given the lack of a perfect recession proxy, I report the results from tests conducted using a number of recession proxies.

The first recession proxy I use is based on an indicator I construct using real-time US industrial production data: I follow the OECD methodology and construct a deviations-from-trend series by applying a one-sided Hodrick and Prescott (1997) filter to real-time industrial production data. A benefit of this approach is that it avoids the publication lag associated with recession proxies such as the NBER recession indicator. I set the recession threshold to $c = -2.2\%$.

[5] I use the terms normal times and expansions interchangeably throughout this paper.

[6] I favor recession indicators based on real-time data as real-time indicators are likely to better capture information available to market participants.

Table 1.1

Return Predictability Regressions

	Dependent variable:							
	r_{t+1}							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
i_t^D	-0.999^{***}	-1.085^{***}	-1.191^{***}	-1.284^{***}	-1.137^{***}	-1.222^{***}	-1.532^{***}	-1.660^{***}
	(0.234)	(0.261)	(0.250)	(0.288)	(0.240)	(0.272)	(0251)	(0.282)
	[0.213]	[0.246]	[0.246]	[0.292]	[0.222]	[0.261]	[0.300]	[0.352]
$i_t^D \cdot \mathbf{1}_{\text{rec}\mid\text{t}}$			1.435^{**}	1.451^{**}	1.716^{**}	1.690^{**}	0.999^{**}	1.036^{**}
			(0.600)	(0.596)	(0.814)	(0.797)	(0.404)	(0.406)
			[0.496]	[0.485]	[0.670]	[0.661]	[0.412]	[0.425]
Constant	0.001		0.000		0.001		0.001	
	(0.001)		(0.001)		(0.001)		(0.001)	
	[0.002]		[0.002]		[0.001]		[0.001]	
$\mathbf{1}_{\text{rec}\mid\text{t}}$			0.003	0.003	-0.002	-0.003	0.000	0.000
			(0.004)	(0.004)	(0.005)	(0.006)	(0.002)	(0.002)
			[0.004]	[0.004]	[0.006]	[0.005]	[0.002]	[0.002]
p-value ($H_0 : \beta^F_{\text{rec}} = 0$)	–	–	> 0.10	> 0.10	> 0.10	> 0.10	> 0.10	0.06
Observations	4,171	4,171	4,171	4,171	4,171	4,171	4,171	4,171
Rec. ind.	None	None	Ind. prod.	Ind. prod.	CFNAI	CFNAI	OECD	OECD
$P(\mathbf{1}_{\text{rec}} = 1)$	–	–	0.126	0.126	0.086	0.086	0.501	0.501
Currency FE	No	Yes	No	Yes	No	Yes	No	Yes
Adjusted R^2	0.016	0.015	0.019	0.018	0.020	0.019	0.019	0.019

This table reports the results for the UIP regression $r_{j,t+1} = \alpha + \beta^F i^D_{j,t} + \varepsilon_{j,t+1}$ and the modified UIP regression $r_{j,t+1} = \alpha_{\text{exp}} + \alpha_\Delta \mathbf{1}_{\text{rec}\mid\text{t}} + \left(\beta^F_{\text{exp}} + \beta^F_\Delta \mathbf{1}_{\text{rec}\mid\text{t}}\right) i^D_{j,t} + \varepsilon_{j,t+1}$, where the recession indicator is constructed using detrended industrial production data, the Chicago Fed National Activity Index (CFNAI) or is provided by the OECD. In parentheses are standard errors computed following Driscoll and Kraay (1998). In square brackets are bootstrapped standard errors. Significance at the 1%, 5%, and 10% is denoted by $***$, $**$, and $*$, respectively. The estimation is carried out using monthly data obtained from Barclays Bank International and Reuters and covers the period between 01/1983 and 09/2020. The panel consists of the exchange rate between the USD and the following 10 currencies: AUD, CAD, DKK, EUR, JPY, NOK, NZD, SEK, CHF, and GBP.

This threshold is chosen to maximize the correlation between the recessions identified using detrended industrial production and the NBER recession dates during the period between 01/1963 and 12/1982.

The second real-time recession proxy I use is based on the Chicago Fed National Activity Index (CFNAI). Using the methodology described in the previous paragraph, I set the recession threshold to $c = -0.60$. The correlation of the CFNAI recession proxy with the industrial production recession proxy is moderately high at 0.50 (t-statistic = 12.28).

I also consider a direct global recession proxy constructed by the Organization for Economic Co-operation and Development (OECD): the OECD and Non-member Economies (OECD) indicator, which covers 35 countries. A shortcoming of this proxy is that it identifies about 50% of my sample as recession periods.

In robustness tests, I use four additional recession proxies: the NBER recession dates, the Chauvet and Piger (2008) US recession probabilities. I also consider an alternative global recession indicator constructed by the OECD: the OECD M7 indicator (covers the following OECD economies: Canada, France, Germany, Italy, Japan, the United Kingdom, and the United States).[7]

1.3.3 Discussion of results

In the first two columns of Table 1.1, I replicate the well-established result regarding the violation of the UIP condition in unconditional tests. The UIP coefficient is estimated to be close to -1 and highly statistically significant.

The results for the modified return predictability regression in Equation (1.3) are reported in columns (3) through (6). The β^F_{exp} coefficient is estimated to be less than -1 and statistically significant at the 1% level. This result implies that foreign exchange returns are forecastable using interest rate differentials during expansions. The novel result within Table 1.1 is that β^F_{Δ} coefficients are estimated to be positive (similar to the findings of Clar-

[7] I adopt the following rule to identify recessions using the Chauvet and Piger (2008) probabilities: recessions occur when the probability of a recession increases above 60% and last until the probability decreases below 30%.

ida et al., 2009) and significantly different from zero. The UIP coefficient estimated during recessions ($\beta^F_{\text{rec}} = \beta^F_{\text{exp}} + \beta^F_{\Delta}$) is estimated to be positive and insignificant in five out of the six specifications considered in the table (the β^F_{rec} coefficient in column (8) is only significant at the 10% level). The implication of these results is that the relation between excess returns and interest rate differentials breaks down during recessions, i.e. investing in risk-free assets denominated in high interest rate currencies does not produce positive excess returns.

1.3.4 Robustness of results

This subsection documents that the results reported in Table 1.1 are robust across a range of alternative specifications. The key robustness tests are shown in the body of the paper. Tests using individual currencies as test assets, tests based on sequentially omitting recession periods from my sample, and predictability regressions including alternative return predictors can be found in Appendix 1.E.[8]

1.3.4.1 Alternative Recession Proxies

A key choice in the modified return predictability regression is the choice of a recession indicator. Therefore, it is important to examine the degree to which the results reported in Table 1.1 are robust to this choice. I estimate the modified UIP regression using a number of alternative US-based and global recession indicators. As shown in Table 1.2, the recession probabilities of Chauvet and Piger (2008), the NBER recession dates, and the OECD M7 recession indicator produce results that are qualitatively similar to those reported in Table 1.1. The β^F_{Δ} coefficient is estimated to be positive and statistically significant for two out of the three recession proxies in Table 1.2. The magnitude of the β^F_{Δ} coefficient based on the NBER recession dates is similar to the magnitude of the coefficients reported in Table 1.1 but we cannot reject the hypothesis that $\beta^F_{\Delta} = 0$.

[8]The additional predictors considered are: real exchange rates, inflation differentials, the US variance risk premium, the VIX, and the He et al. (2017) intermediary factor.

Quantitatively, the results are somewhat weaker for the OECD recession proxies considered in Tables 1.1 and 1.2. This result is not unexpected given the fact that the OECD proxies identify about half of my sample as recession periods. It is also noteworthy that the β^F_{rec} coefficient based on the Chauvet and Piger (2008) recession probabilities is economically large even though the null of a zero β^F_{rec} cannot be rejected. In general, the six recession proxies considered in this paper produce qualitatively consistent results. Therefore, the choice of a specific recession indicator does not seem to be the driving force behind the findings reported in this section.

In appendix 1.E.3, I report the results of placebo tests based on recessions in small open economies and randomly drawn recession dates. The results in the appendix bolster the point that the global recession proxies considered in this section are special with respect to their impact on return predictability in the foreign exchange market.

Asynchronous Business Cycles

Carry trade involves taking a short position in one currency and a long position in another currency. Therefore, a natural question that arises is if the recessions in each individual country have an impact on foreign exchange return predictability above and beyond the impact of the global business cycle. In order to examine this question, I estimate the following regression:

$$r_{j,t+1} = \alpha_{\text{exp}} + \mathbf{1}^{\text{Global}}_{\text{rec|t}} + \mathbf{1}^{\text{Foreign}^\perp}_{\text{rec|t}} + \mathbf{1}^{\text{US}^\perp}_{\text{rec|t}} + \left(\beta^F_{\text{exp}} + \beta^{F,\text{Global}}_{\Delta} + \beta^{F,\text{Foreign}^\perp}_{\Delta} + \beta^{F,\text{US}^\perp}_{\Delta}\right) i^D_{j,t} + \varepsilon_{j,t+1} \tag{1.4}$$

where $\mathbf{1}^{\text{Global}}_{\text{rec|t}}$ is the OECD global recession indicator from Table 1.1 and $\mathbf{1}^{\text{US/Foreign}^\perp}_{\text{rec|t}}$ are recession indicators, which takes on the value of 1 if the US/the foreign country is in a recession and the global recession indicator is equal to 0. To ensure consistency across the recession proxies, I consider US and foreign recession indicators constructed by the OECD. The average probability of a foreign recession is about 0.11 and the probability of a US

Table 1.2
Return Predictability Regressions: Alternative Recession Indicators

	Dependent variable:			
	r_{t+1}			
	(1)	(2)	(3)	(4)
i_t^D	-0.999^{***}	-1.335^{***}	-1.096^{***}	-1.092^{***}
	(0.234)	(0.206)	(0.209)	(0.236)
	[0.213]	[0.246]	[0.223]	[0.224]
$i_t^D \cdot \mathbf{1}_{\text{rec}\vert\text{t}}$		0.739^{**}	5.103^{**}	1.606
		(0.270)	(2.232)	(1.294)
		[0.361]	[2.539]	[0.870]
Constant	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
	[0.002]	[0.002]	[0.001]	[0.002]
$\mathbf{1}_{\text{rec}\vert\text{t}}$		0.000	0.001	-0.001
		(0.002)	(0.007)	(0.006)
		[0.003]	[0.007]	[0.003]
Observations	4,171	4,171	4,171	4,171
Rec. ind.	None	OECD M7	CP (2008)	NBER
Currency FE	No	No	No	No
Adjusted R^2	0.016	0.017	0.024	0.018

This table reports the results for the modified UIP regression $r_{j,t+1} = \alpha_{\text{exp}} + \alpha_{\Delta}\mathbf{1}_{\text{rec}|\text{t}} + \left(\beta^F_{\text{exp}} + \beta^F_{\Delta}\mathbf{1}_{\text{rec}|\text{t}}\right) i^D_{j,t} + \varepsilon_{j,t+1}$, for a number of global (OECD Total and OECD Major 7) and US (Chauvet and Piger, 2008 and NBER recession dates) recession indicators. In parentheses are standard errors computed following Driscoll and Kraay (1998). In square brackets are bootsrapped standard errors. Significance at the 1%, 5%, and 10% is denoted by $***$, $**$, and $*$, respectively. The estimation is carried out using monthly data obtained from Barclays Bank International and Reuters and covers the period between 01/1983 and 09/2020. The panel consists of the exchange rate between the USD and the following 10 currencies: AUD, CAD, DKK, EUR, JPY, NOK, NZD, SEK, CHF, and GBP.

recession is about 0.09. The average correlation between the US recession indicator and the foreign recession indicators is about 0.17.

The results of the estimation are reported in Table 1.3. The evidence presented in the table provides support for the idea that global recessions are the driving force behind the dynamics of conditional return predictability documented in Table 1.1. The $\beta_{\Delta}^{F,\text{Global}}$ coefficient is positive and significantly different from 0 in all specifications considered in Table 1.3. In general, the addition of the local recession indicators does not have a significant quantitative impact on the conditional return predictability results: neither of the additional interaction terms is significant. The coefficient of the foreign recession interaction term is economically small and the coefficient of the interaction term for US recessions is estimated with a negative sign. The negative sign of the $\beta_{\Delta}^{F,US^{\perp}}$ is driven by the fact that the US recessions identified by the OECD lead the global business cycle.

1.3.4.2 Smooth Transition Between Recessions and Expansions

The specification in Equation (1.3) assumes that there is an immediate transition between expansions and recessions. In reality, the transition between recessions and expansions is more gradual. Therefore, modeling the transition as some smooth function may be informative. In order to test the robustness of my results with respect to the choice of transition speed between normal times and recessions, I modify the specification in Equation (1.3) by applying the logistic smooth transition regression (LSTR) of Teräsvirta (1994), which takes on the following form:

$$r_{t+1} = \alpha_{\text{exp}} + \beta_{\text{exp}}^{F} i_t^D + \left(\alpha_{\Delta} + \beta_{\Delta}^{F} i_t^D\right)\left(1 - G(z_t, \iota, c)\right) + \varepsilon_{t+1} \tag{1.5}$$

where $G(z_t,\, c) = \left(1+\exp\left(-\iota(z_t - c)\right)\right)^{-1}$, z_t is the variable used to identify recessions, and c is the value of z_t indicating a transition between an expansion and a recession. The ι coefficient controls the speed of transition between recessions and expansions and the specification in Equation (1.3) corresponds to $\iota \to \infty$. For the value of ι, I use 1.5, 2, and 3 to indicate a case of smooth transition and cases of more rapid transition.

Table 1.3

Return Predictability Regressions: Asynchronous Business Cycles

	Dependent variable:		
	r_{t+1}		
	(1)	(2)	(3)
i_t^D	-1.409^{***}	-1.557^{***}	-1.447^{***}
	(0.258)	(0.278)	(0.286)
	[0.294]	[0.284]	[0.298]
$i_t^D \cdot \mathbf{1}_{\text{rec}\vert\text{t}}^{\text{Global}}$	0.876^{**}	1.027^{**}	0.906^{**}
	(0.414)	(0.408)	(0.417)
	[0.439]	[0.420]	[0.451]
$i_t^D \cdot \mathbf{1}_{\text{rec}\vert\text{t}}^{\text{US}^\perp}$	-0.620		-0.673
	(0.592)		(0.609)
	[0.675]		[0.695]
$i_t^D \cdot \mathbf{1}_{\text{rec}\vert\text{t}}^{\text{Foreign}^\perp}$		0.056	0.136
		(0.351)	(0.329)
		[0.361]	[0.356]
Constant	0.001	0.001	0.001
	(0.002)	(0.001)	(0.002)
	[0.002]	[0.001]	[0.001]
$\mathbf{1}_{\text{rec}\vert\text{t}}^{\text{Global}}$	-0.000	0.000	-0.000
	(0.002)	(0.002)	(0.002)
	[0.002]	[0.002]	[0.002]
$\mathbf{1}_{\text{rec}\vert\text{t}}^{\text{US}^\perp}$	-0.004		-0.003
	(0.004)		(0.003)
	[0.004]		[0.004]
$\mathbf{1}_{\text{rec}\vert\text{t}}^{\text{Foreign}^\perp}$		-0.001	-0.001
		(0.001)	(0.001)
		[0.002]	[0.002]
Observations	4,171	4,137	4,137
Rec. ind.	Global+US$^\perp$	Global+Foreign$^\perp$	Global+US$^\perp$+Foreign$^\perp$
Currency FE	No	No	No
Adjusted R^2	0.020	0.019	0.020

This table reports the results for the modified UIP regression $r_{j,t+1} = \alpha_{\exp} + \alpha_\Delta^{\text{Global}} \mathbf{1}_{\text{rec}|\text{t}}^{\text{Global}} + \alpha_\Delta^{\text{Local}^\perp} \mathbf{1}_{\text{rec}|\text{t}}^{\text{Local}^\perp} + \left(\beta_{\exp}^F + \beta_\Delta^{F,\text{Global}} \mathbf{1}_{\text{rec}|\text{t}}^{\text{Global}} + \beta_\Delta^{F,\text{Local}^\perp} \mathbf{1}_{\text{rec}|\text{t}}^{\text{Local}^\perp}\right) i_{j,t}^D + \varepsilon_{j,t+1}$, for combinations of global and local recessions, as identified by the OECD. In parentheses are standard errors computed following Driscoll and Kraay (1998). In square brackets are bootsrapped standard errors. Significance at the 1%, 5%, and 10% is denoted by $***$, $**$, and $*$, respectively. The estimation is carried out using monthly data obtained from Barclays Bank International and Reuters and covers the period between 01/1983 and 09/2020. The panel consists of the exchange rate between the USD and the following 10 currencies: AUD, CAD, DKK, EUR, JPY, NOK, NZD, SEK, CHF, and GBP.

I use two of the recession proxies which are based on continuous variables in these tests: detrended industrial production and the CFNAI.

Table 1.4
Return Predictability Regressions: Smooth Transition

	Dependent variable:					
	r_{t+1}					
	Industrial Production			CFNAI		
	$\gamma = 1.5$	$\gamma = 2$	$\gamma = 3$	$\gamma = 1.5$	$\gamma = 2$	$\gamma = 3$
i_t^D	-20.508^{***}	-15.560^{***}	-10.617^{***}	-1.746^{***}	-1.513^{***}	-1.303^{***}
	(5.600)	(4.179)	(2.762)	(0.283)	(0.384)	(0.322)
$i_t^D \cdot (1-G)$	39.673^{***}	29.777^{***}	19.892^{***}	2.496^{**}	2.015^{*}	1.546^{*}
	(11.400)	(8.558)	(5.721)	(1.205)	(1.044)	(0.928)
Constant	-0.081^{**}	-0.060^{**}	-0.040^{**}	0.002^{**}	0.001^{**}	0.001^{*}
	(0.034)	(0.025)	(0.017)	(0.001)	(0.001)	(0.001)
$(1-G)$	0.166^{**}	0.125^{**}	0.083^{**}	-0.003^{*}	-0.002	-0.002
	(0.069)	(0.052)	(0.035)	(0.002)	(0.002)	(0.002)
Observations	4,171	4,171	4,171	4,171	4,171	4,171
Adj. R^2	0.018	0.018	0.018	0.018	0.017	0.017

This table reports the results for the modified UIP regression $r_{j,t+1} = \alpha_{\exp} + \beta^F_{\exp} i^D_{j,t} + \left(\alpha_\Delta + \beta^F_\Delta i^D_{j,t}\right)(1 - G(z_t, \gamma, c)) + \varepsilon_{j,t+1}$, where the recession indicator is an indicator constructed using real-time industrial production data or the Chicago Fed National Activity Index (CFNAI). I set the recession threshold $c = -0.022$ for industrial production and $c = -0.60$ for the CFNAI. In parentheses are standard errors computed following Driscoll and Kraay (1998). Significance at the 1%, 5%, and 10% is denoted by $***$, $**$, and $*$, respectively. The estimation is carried out using monthly data obtained from Barclays Bank International and Reuters and covers the period between 01/1983 and 09/2020. The panel consists of the exchange rate between the USD and the following 10 currencies: AUD, CAD, DKK, EUR, JPY, NOK, NZD, SEK, CHF, GBP.

The results reported in Table 1.4 demonstrate that the results in this section are robust to the use of the LSTR model: all $\beta^F_{\exp}$ coefficients are estimated to be negative and highly significant and all β^F_Δ coefficients are estimated to be positive, significant, and greater than $\beta^F_{\exp}$ in absolute value. Given the that using the parsimonious specification in Equation

(1.3) does not have a qualitative impact on my results, I continue to use the specification in Equation (1.3).

1.3.5 Rational Expectations Models and UIP Violations

Having presented empirical evidence on foreign exchange return predictability during the business cycle, I next explore the ability of existing asset pricing models to generate return predictability results similar to those reported in Section 1.3.2.

There are numerous studies that have attempted to explain the UIP puzzle and a thorough examination of all such models is beyond the scope of this paper. In this section, I limit my focus to two classes of representative agent models:[9]

1. A habit formation model based on Verdelhan (2010), which features time-varying risk aversion driven by investors' habit level of consumption. In the habits-based model, domestic agents' risk aversion increases following negative consumption shocks, which leads to lower domestic risk-free rates and higher risk premia.

2. A long-run risks model based on Bansal and Shaliastovich (2013), which features investors with Epstein-Zin preferences and a stochastic consumption volatility process. Following an increase in domestic consumption volatility, the domestic risk-free rate declines and risk-premia increase.

Detailed derivations and technical issues related to the two models are relegated to Appendix 1.A. In Appendix 1.A.2, I derive a closed-form solution for the β^F coefficient for the habits model and in Appendix 1.A.3, I derive the UIP coefficient for the long-run risks model.

As shown in Appendices 1.A.2 and 1.A.3, the canonical models in the literature are unable to account for the stylized fact presented in this paper. This result is not surprising given that these models rely on the coexistence of counter-cyclical risk premia and pro-cyclical interest rates to explain the failure of the UIP condition. This fact also makes it unlikely for

[9] I discuss the implications of heterogeneous agent models in 1.A.4.

straightforward extensions of these frameworks to replicate the evidence presented in this section.

In order for the rational expectations models to generate a positive β_{Δ}^{F}, they would need to carry either (but not both) of the following implications during recessions: (1) risk premia remain unchanged (or decrease) in response to negative shocks; (2) risk-free rates remain unchanged (or increase) in response to negative shocks. The risk premium channel is incompatible with the main idea underlying the models discussed in the section. Therefore, I focus on the risk-free rate channel.

An implication of the risk-free rate channel is that the intertemporal smoothing effect cancels out the precautionary savings effect and agents do not increase their savings in response to negative shocks during recessions. However, this implication is not borne out in the data. For instance, Mody, Sandri, and Ohnsorge (2012) use cross-sectional differences in the labor income uncertainty faced by households in different countries and show that during the Great Recession greater labor income uncertainty is associated with a higher saving rate. Further analysis related to the behavior of interest rates during recessions is presented in Appendix 1.A.

1.3.6 Deviations from Rational Expectations

1.3.6.1 Models Featuring Deviations from Rational Expectations

An alternative view of the UIP puzzle is that it arises from systematic distortions in investors' beliefs about the interest rate process (Froot and Thaler, 1990; Gourinchas and Tornell, 2004).

To gain some intuition regarding the mechanism underlying the UIP violations, consider a positive domestic interest rate shock at time t. If all investors incorporated the interest rate shock into their information sets immediately, arbitrage would force the domestic currency to appreciate up to the point at which future depreciation is equal to the interest rate differential. Now suppose that some investors fail to incorporate the interest rate news into their information sets. In this case, the domestic currency appreciates only moderately at time t and continues to appreciate in subsequent periods as an increasing fraction of the market participants incorporate the shock into their information sets.

In the next section, I examine if interest rate surveys produce results that are consistent with the patterns in foreign exchange return predictability.

1.3.6.2 Forecast Error Predictability Regressions

Testing the rational expectations hypothesis is challenging due to the fact that the econometrician does not observe the decision makers' full information set. Coibion and Gorodnichenko (2015) propose a solution to this problem that involves examining the ability of consensus forecast revisions to predict forecast errors.

Let $x_{t+h|t}$ be the time t expectation (forecast) of the value of variable x at time $t+h$ and x_{t+h} be the value of the variable realized at time $t+h$. The the forecast error is $FE_{t+h|t} = x_{t+h} - x_{t+h|t}$ Following the same notation, $x_{t+h|t-1}$ is the expected value of the variable at time $t-1$. Consequently, the h-period ahead forecast revision at time t is $FR_{t,h} = x_{t+h|t} - x_{t+h|t-1}$. The forecast revision captures agents' updating of their beliefs in response to news observed at time t. Consequently, the extent of over- or underreaction to time t news can be assessed by estimating the following regression:

$$FE_{t+h|t} = \alpha^{CG} + \beta^{CG} FR_{t,h} + \varepsilon_{t,t+h} \tag{1.6}$$

For brevity, I refer to this regression as the FE-on-FR regression and to β^{CG} as the FE-on-FR coefficient.

If an agent's information set includes all time t information, we would not be able to predict forecast errors using any information available at time time t, including forecast revisions, i.e. $\beta^{CG} = 0$ under the null of rational expectations. When forecasters underreact to new information, $\beta^{CG} > 0$. The mechanism underlying this result is the following: let us assume that agents receive a piece of positive news at time t. This implies that $FR_{t,h} > 0$.

However, if agents underreact to the news, the forecast errors will also be positive, on average, i.e. forecast revisions will be positively correlated with forecast errors.

Conversely, overreaction to news implies $\beta^{CG} < 0$. The mechanism underlying this result is the same as the one outlined in the underreaction case. The difference is that the forecast errors will be negative, on average, following positive news. That is, forecast errors are negatively correlated with forecast revisions in the case of overreaction.

Using consensus forecasts from the SPF, Coibion and Gorodnichenko (2015) show that professional forecasters' inflation expectations underreact to news. Coibion and Gorodnichenko (2015) also document that the expectation stickiness of the SPF respondents declines during US recessions.

I estimate the following regression in order to examine how business cycle conditions affect the stickiness of interest rate expectations:

$$FE_{t+h|t} = \alpha_{\text{exp}}^{CG} + \alpha_{\Delta}^{CG}\mathbf{1}_{\text{rec|t}} + \left(\beta_{\text{exp}}^{CG} + \beta_{\Delta}^{CG}\mathbf{1}_{\text{rec|t}}\right) FR_{t,h} + \varepsilon_{t,t+h} \tag{1.7}$$

where β_{exp}^{CG} is the FE-on-FR coefficient during expansions and $\mathbf{1}_{\text{rec|t}}$ is an indicator variable which takes on the value of 1 if period t is a recession period and 0 otherwise. The FE-on-FR coefficient during recessions, β_{rec}^{CG}, is equal to $\beta_{\text{exp}}^{CG} + \beta_{\Delta}^{CG}$.

The regressions in Equations (1.6) and (1.7) are estimated using data from the SPF. For the purposes of this estimation, I pool the forecast horizons of $h = 1$, $h = 2$, and $h = 3$ together. The results are reported in Table 1.5. The standard errors for the panel regressions are adjusted following Driscoll and Kraay (1998). The results for tests based on the individual forecast horizons are in Appendix 1.E.5.

1.3.6.3 Discussion of Results

In the first column of Table 1.5, I test rational expectations using Equation (1.6). The results indicate that unconditional tests reject rational expectations. The FE-on-FR coefficient is

Table 1.5
Forecast Error Predictability Regressions, SPF

	Dependent variable:			
	$FE_{t+h\|t}$			
	(1)	(2)	(3)	(4)
$FR_{t,h}$	0.563***	0.596***	0.544***	0.625**
	(0.110)	(0.137)	(0.128)	(0.247)
$FR_{t,1} \cdot \mathbf{1}_{\text{rec\|t}}$		−0.691***	−0.750**	−0.370
		(0.266)	(0.371)	(0.265)
Constant	−0.246***	−0.182**	−0.210***	−0.059
	(0.083)	(0.081)	(0.081)	(0.121)
$\mathbf{1}_{\text{rec\|t}}$		−0.688***	−0.991***	−0.445***
		(0.238)	(0.273)	(0.151)
p-value ($H_0:\ \beta^{CG}_{\text{rec}} = 0$)	–	> 0.10	> 0.10	0.03
Observations	441	441	441	441
Rec. ind.	None	Ind. prod.	CFNAI	OECD+NM
Adjusted R^2	0.109	0.164	0.154	0.181

This table reports the results for the forecast error-on-forecast revision regression $i_{t+h} - i_{t+h|t} = \alpha^{CG}_{\text{exp}} + \alpha^{CG}_{\Delta}\mathbf{1}_{\text{rec|t}} + \left(\beta^{CG}_{\text{exp}} + \beta^{CG}_{\Delta}\mathbf{1}_{\text{rec|t}}\right)\left(i_{t+h|t} - i_{t+h|t-1}\right) + \varepsilon_{t,t+h}$ for a pooled panel of $h = \{1, 2, 3\}$, where the recession indicator is an indicator constructed using detrended industrial production data or the Chicago Fed National Activity Index (CFNAI). In parentheses are standard errors computed following Driscoll and Kraay (1998). Significance at the 1%, 5%, and 10% is denoted by $***$, $**$, and $*$, respectively. The estimation is carried out using quarterly SPF data obtained from the Philadelphia Fed and covers the period between 1983:Q1 and 2020:Q3.

positive and significant at the 1% level. The adjusted R^2 is about 10%, which provides further evidence in favor of the ability of forecast revisions to predict forecast errors.

In columns (2) and (3), I test rational expectations by conditioning the FE-on-FR regression on the state of the business cycle. I use the same recession proxies I used in Section 1.3.2: detrended industrial production, the CFNAI, and the OECD recession indicator. The results in these columns indicate that rational expectations are rejected during expansions. The $\beta_{\exp}^{CG}$ coefficient is positive and significant for all three recession proxies. These results indicate that survey respondents' expectations underreact to interest rate innovations during expansions. The β_{Δ}^{CG} coefficients are estimated to be negative in all three specifications and statistically significant in two out of the three specifications. Based on the forecast error predictability tests in Table 1.5, we can conclude that the stickiness of interest rate expectations declines significantly during recessions. This decline is consistent with previous findings in the literature (e.g. Loungani et al., 2013 and Andrade and Le Bihan, 2013).

Additionally, the FE-on-FR coefficient associated with recessions, $(\beta_{\exp}^{CG} + \beta_{\Delta}^{CG})$, is estimated to be negative for two out of the three recession proxies. Therefore, the Coibion and Gorodnichenko (2015) methodology provides evidence in favor of overreaction rather than underreaction to interest rate shocks during recessions.

1.3.6.4 Robustness of results

The results of key robustness tests are shown in the body of the paper. Tests of the rationality of interest rate forecasts based on the methodologies developed in Mincer and Zarnowitz (1969) and Kohlhas and Walther (2021) can be found in Appendix 1.E.5.

Alternative recession definitions

In order to examine the degree to which the estimates reported in Table 1.5 are robust to the choice of a recession proxy, I use several alternative recession proxies. As shown in Table 1.6, the results are robust to the use of either US recession indicators (Chauvet and

Piger, 2008 recession probabilities and NBER recession dates) or the OECD M7 recession indicator.

Table 1.6

Forecast Error Predictability Regressions: Alternative Recession Indicators, SPF

	Dependent variable:			
	$FE_{t+h\|t}$			
	(1)	(3)	(4)	(5)
$FR_{t,h}$	0.563***	0.539**	0.613***	0.619***
	(0.110)	(0.212)	(0.133)	(0.144)
$FR_{t,h} \cdot \mathbf{1}_{\text{rec\|t}}$		−0.300	−0.886***	−0.925***
		(0.249)	(0.222)	(0.233)
Constant	−0.246***	−0.105	−0.213***	−0.218***
	(0.083)	(0.104)	(0.081)	(0.081)
$\mathbf{1}_{\text{rec\|t}}$		−0.468***	−0.871***	−0.961***
		(0.158)	(0.214)	(0.247)
Observations	441	441	441	441
Rec. ind.	None	OECD M7	CP (2008)	NBER
Adjusted R^2	0.109	0.166	0.150	0.156

This table reports the results for the forecast error-on-forecast revision regression $i_{t+h} - i_{t+h|t} = \alpha^{CG}_{\text{exp}} + \alpha^{CG}_{\Delta}\mathbf{1}_{\text{rec|t}} + \left(\beta^{CG}_{\text{exp}} + \beta^{CG}_{\Delta}\mathbf{1}_{\text{rec|t}}\right)\left(i_{t+h|t} - i_{t+h|t-1}\right) + \varepsilon_{t,t+h}$ for a pooled panel of $h = \{1, 2, 3\}$, for a number of global (OECD+NM and OECD Major 7) and US (Chauvet and Piger, 2008 and NBER recession dates) recession indicators. In parentheses are standard errors computed following Driscoll and Kraay (1998). Significance at the 1%, 5%, and 10% is denoted by $***$, $**$, and $*$, respectively. The estimation is carried out using quarterly SOF data obtained from the Philadelphia Fed and covers the period between 1983:Q1 and 2020:Q3.

In terms of the estimated sign and magnitude of the FE-on-FR coefficients, the results reported in Table 1.6 are qualitatively similar to the results reported in Table 1.5. However, the results are quantitatively weaker for the OECD M7 proxy considered in Table 1.6: the β^{CG}_{Δ} coefficient is not estimated to be significant at conventional levels, consistent with the results obtained using the broader OECD proxy.

Evidence from the FX4Casts survey of financial institutions

I conduct additional tests using interest rate differentials expectations from the FX4Casts survey to examine if the patterns uncovered in Table 1.5 can be replicated using an alternative survey.[10] The results of the tests using the FX4Casts survey are reported in Tables 1.7.

The forecast error predictability results based on the FX4Casts survey are qualitatively similar to those reported in Tables 1.5 and 1.6. The major difference is that rational expectations cannot be rejected unconditionally. The estimated β^{CG} is positive but not statistically different from zero.

Conditionally, the tests indicate that survey respondents' expectations underreact to interest rate innovations during expansions: all β^{CG}_{exp} coefficients are positive and statistically significant. The degree of expectation stickiness declines significantly during recessions, consistent with the results obtained using the SPF: the β^{CG}_{Δ} coefficients are estimated to be negative and statistically different from zero. The point estimates of the β^{CG}_{rec} coefficients are estimated to be negative in all specifications considered in Tables 1.7. These results indicate that the rejection of rational interest rate expectations during expansions is not a feature that is unique to the SPF. However, the results also indicate that the expectations of the respondents to the FX4Casts survey (financial institutions, which actively participate in the foreign exchange market) are, on average, less likely to underreact to interest rate innovations than the expectations of the professional forecasters surveyed by the Philadelphia Fed.

[10] I construct interest differential expectations by subtracting the foreign interbank rate forecast from the US LIBOR rate forecast.

Table 1.7

Forecast Error Predictability Regressions, FX4Casts

	Dependent variable:			
	$FE_{t+1\|t}$			
	(1)	(2)	(3)	(4)
$FR_{t,1}$	0.062	0.124**	0.142**	0.157*
	(0.042)	(0.061)	(0.071)	(0.082)
$FR_{t,1} \cdot \mathbf{1}_{\mathrm{rec\|t}}$		−0.192**	−0.177**	−0.163*
		(0.075)	(0.077)	(0.095)
Constant	−0.043	−0.057	−0.062	−0.045
	(0.069)	(0.059)	(0.059)	(0.068)
$\mathbf{1}_{\mathrm{rec\|t}}$		0.109	0.157	−0.008
		(0.121)	(0.130)	(0.083)
p-value ($H_0: \ \beta^{CG}_{\mathrm{rec}} = 0$)	–	> 0.10	> 0.10	> 0.10
Observations	1,256	1,256	1,256	1,256
Rec. ind.	None	Ind. prod.	CFNAI	OECD+NM
Adjusted R^2	0.005	0.022	0.026	0.014

This table reports the results for the forecast error-on-forecast revision regression $i^D_{j,t+1} - i^D_{j,t+1|t} = \alpha^{CG}_{\exp} + \alpha^{CG}_{\Delta} 1_{\mathrm{rec|t}} + \left(\beta^{CG}_{\exp} + \beta^{CG}_{\Delta}\mathbf{1}_{\mathrm{rec|t}}\right)\left(i^D_{j,t+1|t} - i^D_{j,t+1|t-1}\right) + \varepsilon_{j,t,t+1}$, where the recession indicator is an indicator constructed using detrended industrial production data or the Chicago Fed National Activity Index (CFNAI). In parentheses are standard errors computed following Driscoll and Kraay (1998). Significance at the 1%, 5%, and 10% is denoted by $***$, $**$, and $*$, respectively. The estimation is carried out using monthly data obtained from FX4Casts and covers the period between 10/2001 and 09/2020. The panel consists of forecasts of the interest rate differential between the the USD and the following six currencies: EUR, JPY, NOK, SEK, CHF, and GBP.

1.4 Model of Exchange Rate Determination

1.4.1 General Setup

1.4.1.1 Primitives

I treat the interest rate differential between the two countries as the primitive object of my analysis. The interest rate differential is exogenously specified and follows an AR(1) process with time-varying volatility:

$$i_t^D = \phi i_{t-1}^D + \sigma_{\varepsilon,t}\varepsilon_t \tag{1.8}$$

where ε is i.i.d. standard normal.

Agents observe realized interest rate differentials and understand that the data generating process is an AR(1).

1.4.1.2 Beliefs

I consider a model in which some investors do not have rational interest rate differential expectations. In particular, I use a version of the Mankiw and Reis (2002) sticky-information model to analyze the impact of deviations from rational expectations on asset prices.

In the Mankiw and Reis (2002) model a fraction λ of the market participants obtains new information about the state of the economy in each period. I extend the model in Mankiw and Reis (2002) by allowing the fraction of agents who update their expectations to depend on the state of the business cycle.

Let $\mathbb{E}^S[i_{t+h}^D]$ be the time t consensus (average) expectation of the time $t+h$ interest rate differential and λ_t be the fraction of agents who update their expectations during period t. Then, the consensus expectations can be written as:

$$\mathbb{E}_t^S[i_{t+h}^D] = \lambda_t \mathbb{E}_t^X[i_{t+h}^D] + (1-\lambda_t)\mathbb{E}_{t-1}^S[i_{t+h}^D] \tag{1.9}$$

where $\mathbb{E}_t^X[i_{t+h}^D]$ are the expectations updated at time t.

By iterating Equation (1.9) backward, I obtain the following expression for the time t consensus expectations:

$$\mathbb{E}_t^S[i_{t+h}^D] = \lambda_t \mathbb{E}_t^X\left[i_{t+h}^D\right] + \sum_{j=1}^{m-2} \lambda_{t-j}\mathbb{E}_{t-j}^X[i_{t+h}^D]\prod_{k=0}^{j-1}(1-\lambda_{t-k}) + \mathbb{E}_{t-m}^X[i_{t+h}^D]\prod_{j=0}^{m-1}(1-\lambda_{t-j}) \quad (1.10)$$

I assume that agents' expectations are distorted by an additional effect: (over)extrapolation. Agents perceive that the interest rate differential follows an AR(1) process with autocorrelation coefficient $\tilde{\phi} > \phi$, i.e. $\mathbb{E}_t^X[i_{t+h}^D] = \tilde{\phi}^h i_t^D$.

Extrapolative expectations generate overreaction to interest rate differential shocks (similar to Angeletos et al., 2021). I assume that the magnitude of the extrapolation effect is determined by psychological biases that are unaffected by the state of business cycle, i.e. market participants use a constant parameter $\tilde{\phi}$ when forming expectations.

Plugging the definition of extrapolative expectations into Equation (1.10) provides us with a simple model that relates current expectations to lagged interest rate differentials:

$$\mathbb{E}_t^S[i_{t+h}^D] = \lambda_t \tilde{\phi}^h i_t^D + \tilde{\phi}^h \sum_{j=1}^{m-2} \lambda_{t-j}\tilde{\phi}^j i_{t-j}^D \prod_{k=0}^{j-1}(1-\lambda_{t-k}) + \tilde{\phi}^{m+h} i_{t-m}^D \prod_{j=0}^{m-1}(1-\lambda_{t-j}) \quad (1.11)$$

Lower values of λ_t are associated with a greater weight placed on past observations (higher expectation stickiness). Given a persistent underlying process, expectations stickiness leads to an initial underreaction to news. On the other hand, higher values of $\tilde{\phi}$ lead to overreaction due to the data generating process being perceived to be too persistent.

Modeling recessions

An implicit assumption underlying the tests presented in Section 1.3 is that the world is either in a recession state or an expansion state. One of the principal effects of recessions discussed in Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016) is high return volatility. In order to incorporate the return volatility effect into my stylized model, I assume that the variance of the shock to the interest rate differential is higher in recessions than in expansions:

$\sigma_\varepsilon(r) > \sigma_\varepsilon(e)$. Further, I assume that the variance of the interest rate differential evolves according to a two-state Markov chain with a transition probability matrix Π:

$$\Pi = \begin{bmatrix} \pi_{ee} & 1-\pi_{rr} \\ 1-\pi_{ee} & \pi_{rr} \end{bmatrix} \tag{1.12}$$

Following the discussion above, I denote the fraction of agents who update their expectations in a given period by $\lambda_{s|t}$. The $\lambda_{s|t}$ notation represents that the fraction of agents that update their expectations only depends on the state prevailing at time t. In this two-state setting, $s = \{r, e\}$. The r subscript indicates recession periods and the e subscript indicates expansions. As shown in the equilibrium model developed in Appendix 1.C, higher uncertainty during recessions leads agents to endogenously choose a higher probability of updating their expectations, i.e. $\lambda_r > \lambda_e$. For the purposes of the simple model presented in this section, I take this result as a given and assume that λ_t evolves exogenously according to a two-state Markov chain with the transition probability matrix Π shown above.

A shortcoming of the expectation formation process in Equation (1.11) is that we cannot derive closed-form solutions for moments of interest. In order to rectify this shortcoming and ensure greater comparability with the existing literature, I approximate the sum in Equation (1.11) as:

$$\mathbb{E}_t^S\left[i_{t+h}^D\right] \approx \sum_{j=0}^{m-1} \lambda_{s|t}(1-\lambda_{s|t})^j \tilde{\phi}^{j+h} i_{t-j}^D + (1-\lambda_{s|t})^m \tilde{\phi}^{m+h} i_{t-m}^D \tag{1.13}$$

The approximation in Equation 1.13 replaces the λ realized at times $t-1$, $t-2$,..., $t-m$ with the value of λ realized at time t. For example, consider an economy with $m = 2$ which transitioned into a recession 1 period ago. Then, the sum in Equation (1.11) is:

$$\mathbb{E}_t^S\left[i_{t+h}^D\right] = \lambda_r \tilde{\phi}^h i_t^D + \lambda_e\left(1-\lambda_r\right)\tilde{\phi}^{h+1} i_{t-1}^D + \left(1-\lambda_e\right)\left(1-\lambda_r\right)\tilde{\phi}^{h+2} i_{t-2}^D$$

while the approximation is:

$$\mathbb{E}_t^S\left[i_{t+h}^D\right] = \lambda_r \tilde{\phi}^h i_t^D + \lambda_r\left(1-\lambda_r\right)\tilde{\phi}^{h+1} i_{t-1}^D + \left(1-\lambda_r\right)^2 \tilde{\phi}^{h+2} i_{t-2}^D$$

Equation (1.13) only serves as an approximation to the consensus expectations only during the first m periods following a transition between regimes. At all other times, the expressions in Equations (1.11) and (1.13) are equivalent.[11] Going forward, I denote $\lambda_{s|t}$ as λ_s to simplify notation.

1.4.2 Model Implications for Forecast Error Predictability

Having established intermediaries' expectation formation process, I study the implications of the process outlined in Equation (1.13) for forecast error predictability.

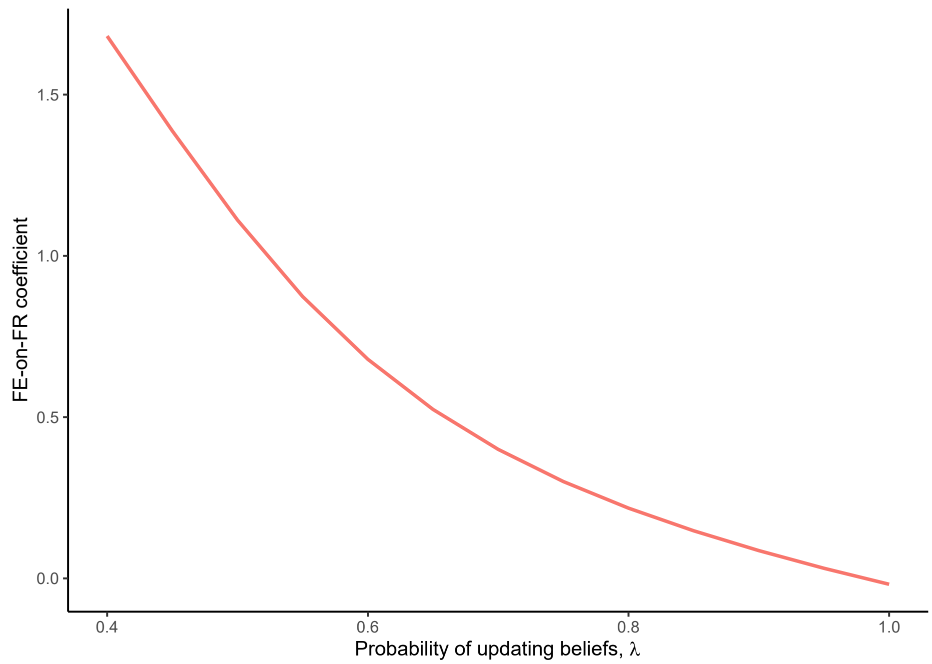
The functional form of the FE-on-FR coefficient implied by the model is not particularly intuitive. Therefore, the expression for β_s^{CG}, along with the derivation of the coefficient, are relegated to Appendix 1.B.

Figure 1.1 below illustrates the relation between the updating probability, λ and the FE-on-FR coefficient for empirically relevant values of λ for the SPF. The figure is based on the following parameters: $\phi = 0.94$, $\tilde{\phi} = 0.955$, and $m = 4$. The parameters are calibrated as follows: ϕ is the persistence of the 3-month T-Bill rate at the quarterly frequency and $\tilde{\phi}$ is calibrated to match β_{rec}^{CG}.

The FE-on-FR coefficient is a strictly decreasing function of λ. The intuition behind this result is straightforward: a higher λ implies that a larger fraction of the investors update their expectations in a given period. Therefore, interest rate news is incorporated into consensus expectations faster. If the fraction of agents updating their expectations is high enough, the constant extrapolation effect comes to dominate the expectation stickiness effect and expectations overreact to interest rate news (the FE-on-FR coefficient becomes negative).

[11] According to the equilibrium model in Appendix 1.C, the numerical errors associated with the approximation are relatively small.

Figure 1.1. FE-on-FR Coefficient



This figure illustrates the relation between the probability of updating beliefs, λ and the FE-on-FR coefficient calibrated according to the SPF. The parameter values used to generate the figure are $\phi = 0.94$, $\tilde{\phi} = 0.955$, $m = 4$.

The average probability of updating implied by the FE-on-FR regressions using the SPF is between 0.62 and 0.65 for the six recession proxies considered in this paper, i.e. during expansions about 35% to 38% of the professional forecasters fail to update their expectations in a given quarter. During recessions, the expectation stickiness effect essentially disappears, with implied $\lambda_r > 0.95$ for all recession proxies considered in this paper. The degree of expectation stickiness in the SPF survey is in line with stickiness estimated in empirical studies, e.g. Andrade and Le Bihan (2013) estimate the unconditional λ to be about 0.75.

The λ_e in the FX4Casts data is between 0.84 and 0.90, i.e. between 10% and 16% of the survey respondents fail to update their expectations during expansions.[12] The estimated updating probability during recessions is greater than 0.95, which is similar to the updating probability obtained using the SPF.

[12]The FX4Casts λ estimates are based on the following calibration: $\phi_{FX} = 0.92$, $\tilde{\phi}_{FX} = 0.93$, $m = 3$.

1.4.3 Equilibrium Exchange Rate

In order to solve for the equilibrium exchange rate, I start by taking the expectation of the expression in Equation (1.2) under measure $\mathbb{S}$:

$$\mathbb{E}_t^S[s_{t+1}] - s_t = i_t^D + \mathbb{E}_t^S[r_{t+1}] \tag{1.14}$$

Then, I iterate the expression in Equation (1.14) forward:[13]

$$s_t = -\sum_{j=0}^{\infty} \mathbb{E}_t^S[i_{t+j}^D] + \sum_{j=0}^{\infty} \mathbb{E}_t^S[r_{t+j}] + \lim_{T\to\infty} \mathbb{E}_t^S[s_{t+T}] \tag{1.15}$$

In order to emphasize the role of deviation from rational expectations, I abstract from time-varying subjective risk premia (I set $\mathbb{E}_t^S[r_{t+j}] = 0\ \forall j$).[14] I also assume that the exchange rate is stationary and that there are no misperceptions about the long-run level of the exchange rate (Gourinchas and Tornell, 2004). That is, $\lim_{T\to\infty} \mathbb{E}_t^S[s_{t+T}] = \lim_{T\to\infty} \mathbb{E}_t^P[s_{t+T}] = \bar{\mu} < \infty$.

Then the equilibrium exchange rate is given by:

$$s_t = -i_t^D - \frac{1}{1-\tilde{\phi}} \mathbb{E}_t^S[i_{t+1}^D] + \bar{\mu} \tag{1.16}$$

The expression in Equation (1.16) demonstrates that the exchange rate is the negative of the present discounted value of expected interest rate differentials. If we assume away the deviations from rational expectations, the exchange rate takes on the familiar form of: $-\frac{1}{1-\phi} i_t^D + \bar{\mu}$.

[13]Note that the law of iterated expectations holds under measure $\mathbb{S}$: $\mathbb{E}_t^S[\mathbb{E}_{t+1}^S[i_{t+2}^D]] = \mathbb{E}_t^S[i_{t+2}^D] = \tilde{\phi}\mathbb{E}_t^S[i_{t+1}^D]$.

[14]The equilibrium model in Appendix 1.C allows for the existence of subjective risk premia.

1.5 Asset Pricing Implications

1.5.1 Foreign Exchange Return Predictability

The UIP coefficient based on the specification in Equation (1.3) takes on the following form:

$$\beta_s^F = \frac{\text{cov}_s\left(\mathbb{E}_t^P[r_{t+1}], i_t^D\right)}{\text{var}_s\left(i_t^D\right)} \tag{1.17}$$

where the expectations operator $\mathbb{E}_t^P$ reflects the fact the econometrician uses returns under the physical measure $\mathbb{P}$ when estimating the UIP regression.

The expected returns under $\mathbb{P}$ can be written as:

$$\mathbb{E}_t^P[r_{t+1}] = \mathbb{E}_t^S[r_{t+1}] + \mathbb{E}_t^P[s_{t+1}] - \mathbb{E}_t^S[s_{t+1}] \tag{1.18}$$

According to (1.18), both risk premia under the subjective measure and expectation errors generate predictable excess returns under measure $\mathbb{P}$.

As shown in Appendix 1.B, the expectation error term, $\mathbb{E}_t^P[s_{t+1}] - \mathbb{E}_t^S[s_{t+1}]$, is equivalent to:

$$\left(1 + \frac{\tilde{\phi}\lambda_s}{1-\tilde{\phi}}\right)\left(\mathbb{E}_t^S[i_{t+1}^D] - \mathbb{E}_t^P[i_{t+1}^D]\right) \tag{1.19}$$

The sign and magnitude of the expectation error term are determined by the relative magnitude of the expectation stickiness and extrapolation effects. During expansions, the expectation stickiness effect dominates the extrapolation effect. As a result, consensus expectations underreact to interest rate news. The gap between the expectations under measures $\mathbb{S}$ and $\mathbb{P}$ is associated with return predictability and a negative UIP coefficient.

The mechanism which generates a negative UIP coefficient is as follows: in response to a positive domestic interest rate shock (an interest rate differential increase) expected interest rate differentials increase under both measures. The increase, however, is larger for the objective expectations because the positive shock is incorporated into the subjective expectations sluggishly.

The novel implication of the model developed in this paper is that periods of low expectation stickiness are associated with a less negative (or positive) UIP coefficient. The interplay between the weakening expectation stickiness effect and the constant extrapolation effect leads to a shrinking gap between expectations under measures $\mathbb{P}$ and $\mathbb{S}$. As expectations under the two measures converge, our ability to predict returns using interest rate differential disappears (we get a less negative UIP coefficient).

In order to study the ability of the model to match the findings regarding foreign exchange return predictability presented in Section 1.3, I solve for the UIP coefficient. The UIP coefficient takes on the following form:

$$\beta_s^F = \left(1 + \frac{\tilde{\phi}\lambda_s}{1-\tilde{\phi}}\right)\left(\lambda_s\tilde{\phi}\frac{1-\left((1-\lambda_s)\tilde{\phi}\phi\right)^{m-1}}{1-(1-\lambda_s)\tilde{\phi}\phi} + (1-\lambda_s)^m\,\tilde{\phi}^{m+1}\phi^m - \phi\right) \tag{1.20}$$

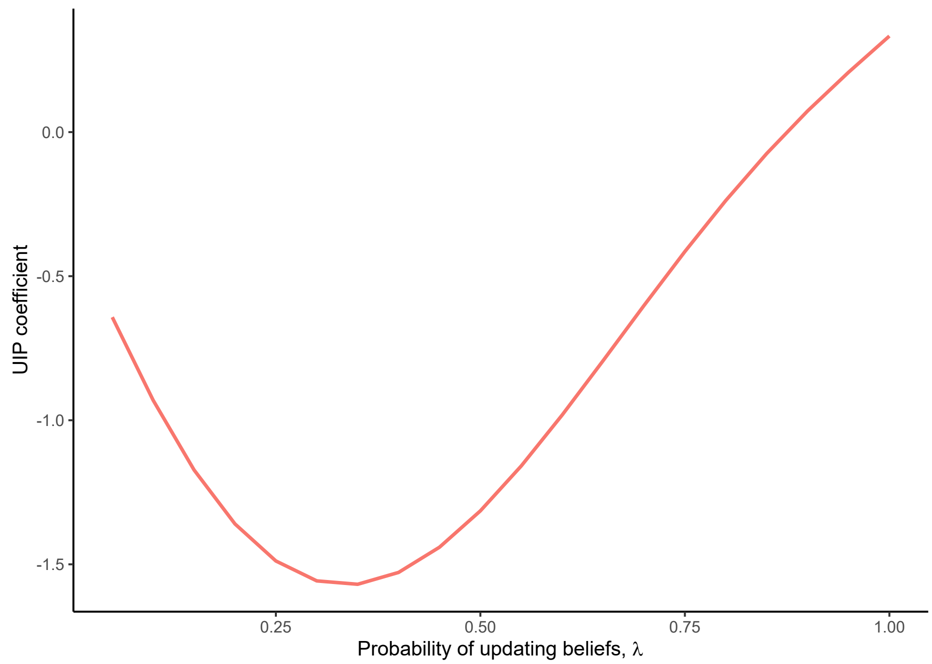
The derivation of the UIP coefficient is presented in Appendix 1.B.

Figure 1.2 demonstrates the model-implied UIP coefficient as a function of the updating probability λ. The ϕ, $\tilde{\phi}$, and m parameters are the same as those used to generate Figure 1.1.

Qualitatively, the model is capable of matching the patterns in the data. The model-implied UIP coefficient is negative for low values of the updating probability and is positive (or negative and close to zero) for high values of the updating probability. Intuitively, the UIP coefficient is negative as long as the degree of expectation stickiness is high enough ($\lambda_s = \lambda_e$). Expectation stickiness leads consensus interest rate expectations to be gradually revised upward following a positive domestic interest rate shock, i.e. high domestic interest rates coexist with a domestic currency that appreciates in the short run, which leads to the violation of the UIP condition. When expectation stickiness is low ($\lambda_s = \lambda_r$), consensus expectations incorporate the interest rate shock instantaneously and the high interest domestic currency depreciates in the short run, as predicted by the UIP condition.

Figure 1.3 shows the quantitative implications of the model for forecast error predictability. The results in the figure are based on the CFNAI recession proxy. Results based on the

Figure 1.2. UIP Coefficient



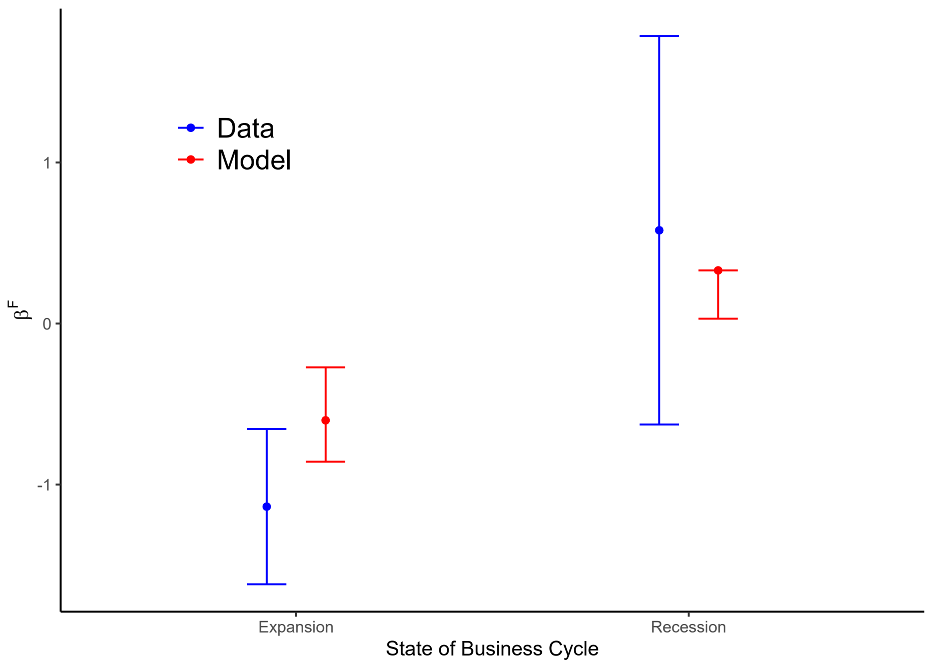
This figure illustrates the relation between the probability of updating beliefs, λ and the UIP coefficient for beliefs calibrated according to the SPF. The parameter values used to generate the figure are $\phi = 0.94$, $\tilde{\phi} = 0.955$, $m = 4$.

industrial production recession proxy are qualitatively similar. The 95% confidence intervals for β^F_{exp} and β^F_{rec} estimated in column (5) of Table 1.1 are plotted in blue. The 95% confidence intervals for the model-implied UIP coefficients are shown in red. The model-implied coefficients are based on the values of λ obtained from forecast error predictability regressions with a forecast horizon of one quarter (shown in Appendix 1.E.5). The confidence interval for λ_r is bounded by 1 from above.

Quantitatively, the model disciplined using the SPF does a good job in matching the patterns documented in Section 1.3 despite the fact that I do not use asset pricing data to calibrate the model. In particular, given the 95% confidence intervals for updating probabilities implied by the SPF, the model generates β^F_s coefficients that are within two standard errors of the point estimates of the β^F_s coefficients reported in column (5) of Table 1.1.[15]

[15]Note that the λ parameter used in this calibration provides a lower bound for stickiness, as it assumes that agents immediately rebalance their portfolios upon updating their expectations (Auclert, Rognlie, and Straub, 2020).

Figure 1.3. UIP Coefficients in the Model and in the Data



This figure illustrates the relation between the empirical and model-implied values of β^F for recessions and expansions. The parameter values used to generate the figure are $\phi = 0.94$, $\tilde{\phi} = 0.955$, $m = 4$. The error bars for β_s^F in the data represent the 95% confidence intervals around $\hat{\beta}^F$. The error bars for the model-derived coefficients represent the β_s^F implied by the values of λ_r and λ_e within the 95% confidence intervals. Recessions are identified using the CFNAI.

1.5.2 Additional Model Implications

1.5.2.1 Bond Return Predictability

Given the evidence regarding forecasters' 3-month T-Bill expectations presented in Section 1.3, the framework developed in Section 1.4 makes a clear prediction about bond return predictability.

The question of bond return predictability during expansions and recessions has been examined by Andreasen et al. (2020), who show that expected excess bond returns are positively correlated with the slope of the yield curve (yield spread) during expansions and that the correlation turns negative during recessions. In this section, I show that the simple model developed in this paper produces results that are consistent with the findings in Andreasen et al. (2020).

In order to zero in on the implications of deviations from rational expectations for bond return predictability, I abstract away from time-varying risk premia and impose that the expectations hypothesis holds under the subjective measure $\mathbb{S}$. Under the expectations hypothesis, the log price of an N-period zero-coupon bond is:

$$p_t^N = -\mathbb{E}_t^S \left[\sum_{j=1}^N i_{t+j-1} \right] = -i_t - \frac{1 - \tilde{\phi}^N}{1 - \tilde{\phi}} \mathbb{E}_t^S [i_{t+1}] \tag{1.21}$$

where $\mathbb{E}_t^S$ is the expectation operator under measure $\mathbb{S}$ defined in Section 1.4. Similarly, the log bond yield is:

$$y_t^N = \frac{1}{N} \mathbb{E}_t^S \left[\sum_{j=1}^N i_{t+j-1} \right] = \frac{1}{N} \left(i_t + \frac{1 - \tilde{\phi}^N}{1 - \tilde{\phi}} \mathbb{E}_t^S [i_{t+1}] \right) \tag{1.22}$$

In my analysis, I use the difference between the yield of an N-period bond and the yield of a 3-month (one-period) bond as the yield spread. Therefore, the coefficient of the bond return predictability regression (the bond predictability coefficient henceforth) can be written as:

$$\beta_s^{CS,N} = \frac{\text{cov}_s \left(\mathbb{E}_t^P [r_{t+1}], y_t^N - i_t \right)}{\text{var}_s \left(y_t^N - i_t \right)} \tag{1.23}$$

The excess return from holding an N-period bond for one period under measure $\mathbb{P}$ is:

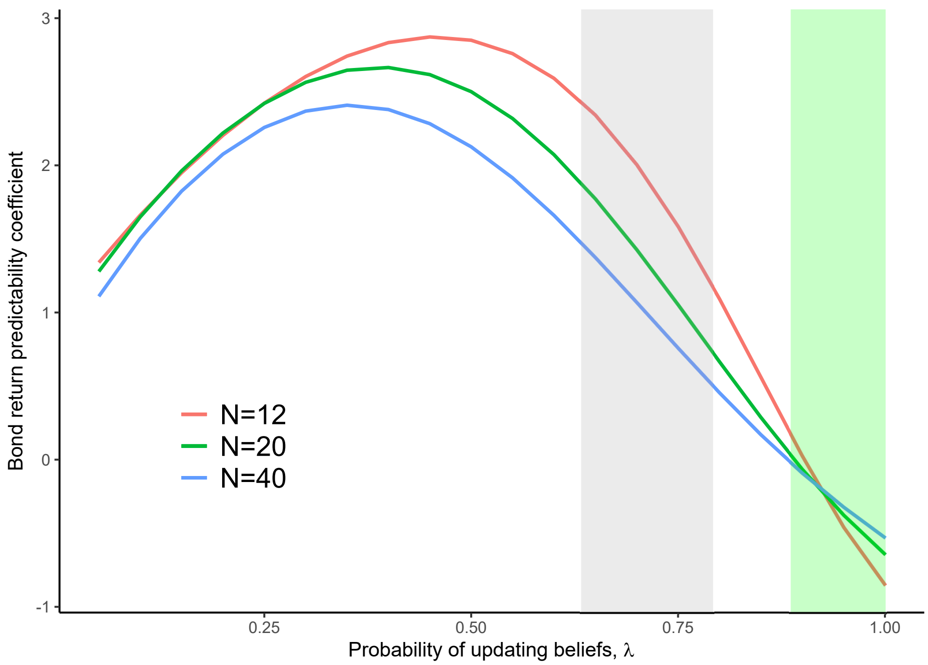
$$\mathbb{E}_t^P [r_{t+1}^N] = \mathbb{E}_t^P [p_{t+1}^{N-1}] - \mathbb{E}_t^S [p_{t+1}^{N-1}] + \mathbb{E}_t^S [r_{t+1}^N] \tag{1.24}$$

Given the assumption that the expectations hypothesis holds, $\mathbb{E}_t^S [r_{t+1}^N] = 0$. Therefore, all of the excess returns observed under measure $\mathbb{P}$ are driven by expectation errors regarding future bond prices.

The expression for the bond return predictability coefficient is derived in Appendix 1.B.

Figure 1.4 shows the bond return predictability coefficient as a function of λ_s for a three-year ($N = 12$), a five-year ($N = 20$), and a ten-year ($N = 40$) bond. The figure is generated using the following parameters: $\phi = 0.94$, $\tilde{\phi} = 0.955$, $m = 4$. The figure is based on recession periods identified using the CFNAI proxy. The 95% confidence interval for λ estimated during expansions is represented by the gray rectangle in the figure. The green rectangle represents The 95% confidence interval for λ estimated during recession, bounded by 1.

Figure 1.4. Bond Return Predictability Coefficient



This figure illustrates the relation between the probability of updating beliefs, λ and the bond return predictability coefficient for three different bond maturities: 12 quarters ($N = 12$), 20 quarters ($N = 20$), and 40 quarters ($N = 40$). The parameter values used to generate the figure are $\phi = 0.94$, $\tilde{\phi} = 0.955$, $m = 4$. The gray rectangle represents the 95% confidence interval for λ_e. The green rectangle represents the 95% confidence interval for λ_r , bounded by 1. Recessions are identified using the CFNAI.

The model delivers bond return predictability coefficients that match the findings in Andreasen et al. (2020). The model delivers a positive bond return predictability coefficients during expansions and captures the sign reversal of the coefficient observed during recessions.

To gain some intuition regarding the mechanism that allows the model to match the patterns in the data, consider a positive interest rate shock, which leads to lower bond prices. During periods characterized by high expectation stickiness, the interest rate news is incorporated into consensus expectations gradually and bond prices continue to decline in the short run. Therefore, declining yield spreads are associated with negative returns during low λ_s periods.[16] Conversely, during periods with low information stickiness, bond prices decline too much in response to a positive interest rate shock, which leads to positive expected returns and a negative correlation between excess returns and the yield spread.

1.5.2.2 Time-Series Momentum

The framework developed in Section 1.4 makes predictions regarding the profitability of time-series momentum strategies involving foreign exchange and fixed income instruments (Moskowitz et al., 2012).[17] In this section, I emphasize the intuition underlying the implications of the model for time-series momentum. All derivations are relegated to Appendix 1.B.4.

Consider a positive domestic interest rate shock during period $t-1$, characterized by high inattention. Following the shock, the domestic currency appreciates and investors experience positive domestic currency returns ($r_{t-1} < 0$). Fraction $1 - \lambda_e$ of the participants fails to incorporate the time $t-1$ news into their information sets. Consequently, the domestic currency continues to appreciate in expectation, thus delivering positive expected returns ($\mathbb{E}^P_{t-1}[r_t] < 0$). Therefore, the slow diffusion of interest rate news is associated with the presence of time-series momentum (Hong and Stein, 1999).

Conversely, following periods characterized by low expectation stickiness, the time $t-1$ shock is incorporated into expectations promptly and the domestic is not expected to further appreciate at time t. Therefore, the model in Section 1.4 predicts lower time-series momentum returns following high λ periods.[18]

In order to empirically test the predictions of the model, I compare the average time series momentum factor returns following expansions ($\bar{r}_{\text{exp}}$) to the average returns following recessions ($\bar{r}_{\text{rec}}$). The month t momentum returns are identified as following a recession if

[16] Note that $\frac{\partial(y_t^N - i_t)}{\partial i_t} < 0$ for a large enough $N (\approx 5)$ regardless of the value of λ_s.

[17] The correlation between the returns of the two momentum strategies is relatively low at 0.13 (t-statistic = 2.66).

[18] Episodes of low returns (crashes) have been identified in the cross-sectional momentum literature (Cooper, Gutierrez Jr., and Hameed, 2004 and Daniel and Moskowitz, 2016).

at least one of the h months preceding t is identified as a recession using the CFNAI.[19] I download the time-series momentum factor returns for the period between 01/1985 and 09/2020 from AQR Capital Management.[20] The average returns for $h = \{1, 3, 6\}$ are reported in Table 1.8.

The results in Table 1.8 show that the returns of time series momentum strategies are lower following periods that contain a recession. The average factor returns following expansions are positive (between 0.92% and 1.80% per month) with t-statistics greater than 3.8. Conversely, the performance of the time-series momentum strategy is dramatically different following recessions. Neither of the momentum strategies generates statistically positive returns following recessions. In fact, $\bar{r}_{\text{rec}}$ for fixed income momentum is negative in all specifications. The difference between $\bar{r}_{\text{exp}}$ and $\bar{r}_{\text{rec}}$, $\Delta\bar{r}$, is negative and significant at conventional levels in five out of the six specifications considered in the table. The significance of $\Delta\bar{r}$ reported in the table is based on one-tailed tests of the null hypothesis of $\bar{r}_{\text{rec}} \geq \bar{r}_{\text{exp}}$. In general, the evidence presented in Table 1.8 is consistent with the time-varying inattention model developed in Section 1.4.

1.5.2.3 Post-2007 Foreign Exchange Return Predictability

A recent finding that has attracted attention in the international economics literature (Bussiere et al., 2022; Engel et al., 2021) is that the UIP coefficient has attenuated or even reversed following the Great Recession, during which period interest rates have been close to or at the zero lower bound (ZLB).

Based on the analysis presented in this paper, I hypothesize that the two severe recessions during the post-2007 period drive the results reported in these papers. To test this hypothesis, I conduct several tests whereby I start with the full post-2007 sample and sequentially drop observations associated with the Great Recession and the COVID-19 recession. The

[19] Alternative recession proxies produce similar results.

[20] The data is available at https://www.aqr.com/Insights/Datasets/Time-Series-Momentum-Factors-Monthly.

Table 1.8
Time Series Momentum

	$r_{\mathrm{t,exp}}$	$r_{\mathrm{t,rec}}$	Δr_t
Panel A: Foreign Exchange Momentum			
	$h = 1$		
Mean profit	0.92***	0.09	−0.82
(t-statistic)	(3.89)	(0.10)	(−1.25)
	$h = 3$		
Mean profit	0.95***	0.12	−0.83*
(t-statistic)	(4.03)	(0.18)	(−1.55)
	$h = 6$		
Mean profit	1.05***	0.14	−0.92**
(t-statistic)	(4.51)	(0.27)	(−1.74)
Panel B: Fixed Income Momentum			
	$h = 1$		
Mean profit	1.64***	−1.19	−2.83**
(t-statistic)	(3.96)	(−0.83)	(−2.25)
	$h = 3$		
Mean profit	1.80***	−0.90	−2.69***
(t-statistic)	(4.47)	(−0.88)	(−3.03)
	$h = 6$		
Mean profit	1.80***	−0.05	−1.85***
(t-statistic)	(4.61)	(−0.06)	(−2.44)

This table reports the average returns of the foreign exchange (Panel A) and fixed income (Panel B) factors following expansions ($\bar{r}_{\mathrm{exp}}$) and recessions ($\bar{r}_{\mathrm{rec}}$). The difference between the average returns is denoted by $\Delta\bar{r}$. Significance at the 1%, 5%, and 10% is denoted by $***$, $**$, and $*$, respectively. The significance of $\Delta\bar{r}$ is based on one-tailed tests of the null hypothesis of $\bar{r}_{\mathrm{rec}} \geq \bar{r}_{\mathrm{exp}}$. Recession periods are determined using the Chicago Fed National Activity Index with recession threshold of -0.60 and lookback periods (h) of one month, three months, and six months. The time series momentum data is obtained from AQR Capital Management and covers the period between 01/1985 and 09/2020. Returns are expressed in percentage points per month.

tests are based on the following regression:

$$r_{t+1} = \alpha + \alpha_\Delta \mathbf{1}_{\mathrm{t}\in[t_0,T]} + \left(\beta^F + \beta^F_\Delta \mathbf{1}_{\mathrm{t}\in[t_0,T]}\right) i^D_t + \varepsilon_{t+1} \tag{1.25}$$

where the indicator function $\mathbf{1}_{\mathrm{t}\in[t_0,T]}$ takes on the value of 1 if period t falls within: (1) the post-2007 period; (2) the post-2007 period excluding the Great Recession and the COVID-19 crisis; (3) the post-2007 period excluding 06/2007-12/2009 and 03/2020-09/2020; (4) the period during which the 90-day USD LIBOR rate is below 1%. The results are reported in Table 1.9.

In columns (1) and (2), I replicate the results reported by Bussiere et al. (2022) and Engel et al. (2021) and find evidence in favor of a sign reversal. The β^F_Δ coefficient is positive and statistically significant. Additionally, the $\beta^F + \beta^F_\Delta$ coefficient is estimated to be positive. If we were to end the analysis at this point, we would erroneously conclude that we face a "new Fama puzzle" that involves a reversal of the sign of the UIP coefficient.

In columns (3) and (4), I exclude the Great Recession and the COVID-19 recession from my sample. The β^F_Δ coefficient is estimated to be positive but is statistically insignificant. Therefore, we cannot reject the hypothesis that the UIP coefficient during the post-2007 is identical to the full sample β^F once the two recessions are excluded from the sample. Additionally, the $\beta^F + \beta^F_\Delta$ estimated during this sample is negative albeit economically smaller than the full sample UIP coefficient. These results imply that the post-2007 period is not dissimilar from previous historical episodes and that the differences in unconditional tests are primarily driven by the fact that the Great Recession lasted for 18 months.

The theoretical analysis in Appendix 1.C suggests that the increased updating probability during recessions should carry over to periods immediately following a recession. Additionally, an extended period of heightened uncertainty preceded the Great Recession (e.g. the VIX shot up during the second half of 2007 and did not drop below 20 until early 2010).

Table 1.9

Return predictability regressions: Post-2007 Sample

	Dependent variable:							
	r_{t+1}							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
i_t^D	-1.109^{***}	-1.136^{***}	-1.032^{***}	-1.091^{***}	-0.965^{***}	-1.047^{***}	-0.997^{***}	-1.085^{***}
	(0.214)	(0.233)	(0.225)	(0.249)	(0.245)	(0.268)	(0.241)	(0.270)
	[0.225]	[0.244]	[0.219]	[0.245]	[0.211]	[0.241]	[0.214]	[0.244]
$i_t^D \cdot 1_{\mathrm{t}\in[t_0,T]}$	1.966^{**}	1.979^{*}	0.896	0.787	-0.257	-0.528	-0.186	-0.726
	(0.997)	(1.143)	(0.739)	(0.842)	(0.835)	(0.842)	(0.883)	(0.983)
	[0.954]	[1.033]	[0.749]	[0.797]	[0.938]	[1.002]	[1.088]	[1.269]
Constant	0.001		0.001		0.001		0.001	
	(0.001)		(0.001)		(0.001)		(0.001)	
	[0.001]		[0.001]		[0.001]		[0.001]	
$\mathbf{1}_{\mathrm{t}\in[t_0,T]}$	-0.001	-0.001	-0.001	-0.001	-0.003	-0.003	-0.002	-0.003
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.003]
Observations	4,171	4,171	4,171	4,171	4,171	4,171	4,171	4,171
$1_{\mathrm{t}\in[t_0,T]} = 1$	Post-2007		Post-2007 excl. rec.		Post-2007 excl. rec.+		ZLB	
Currency FE	No	Yes	No	Yes	No	Yes	No	Yes
Adjusted R^2	0.019	0.018	0.016	0.015	0.017	0.016	0.016	0.015

This table reports the results for the modified UIP regression $r_{j,t+1} = \alpha + \alpha_\Delta \mathbf{1}_{\mathrm{t}\in[t_0,T]} + \left(\beta^F + \beta^F_\Delta \mathbf{1}_{\mathrm{t}\in[t_0,T]}\right) i^D_{j,t} + \varepsilon_{j,t+1}$, where the indicator $\mathbf{1}_{\mathrm{t}\in[t_0,T]}$ takes on the value of 1 for: the post-2007 period, the post-2007 period excluding the Great Recession and the COVID-19 crisis, the post-2007 period excluding the periods between 06/2007 and 01/2010 and 03/2020 and 09/2020, or the zero-lower bound period (the period during which the 90-day USD LIBOR rate is below 1%). In parentheses are standard errors computed following Driscoll and Kraay (1998). In square brackets are bootstrapped standard errors. Significance at the 1%, 5%, and 10% is denoted by $***$, $**$, and $*$, respectively. The estimation is carried out using monthly data obtained from Barclays Bank International and Reuters. The full sample covers the period between 01/1983 and 09/2020. The panel consists of the exchange rate between the USD and the following 10 currencies: AUD, CAD, DKK, EUR, JPY, NOK, NZD, SEK, CHF, and GBP.

To account for the high uncertainty during periods surrounding recessions, I exclude the periods between 06/2007 and 01/2010 and 03/2020 and 09/2020 from the test sample. The results are reported in columns (5) and (6). The β_{Δ}^{F} coefficient is estimated to be negative. Additionally, $\beta^{F} + \beta_{\Delta}^{F}$ is not economically dissimilar from the full sample UIP coefficient.

In columns (7) and (8), I examine foreign exchange return predictability during the sample during which the 90-day USD LIBOR rate is below 1%.[21] The results in these columns are similar to the results in columns (5) and (6) and reinforce the findings in columns (3) and (4). The results provide additional evidence that interest rates being at the ZLB did not have a major impact on foreign exchange return predictability.

1.6 Conclusion

I demonstrate that one of the most widely studied phenomena in finance, the violations of the UIP condition, is confined to economic expansions. I additionally show that the cyclical behavior of interest rate expectation stickiness is consistent with the foreign exchange return predictability patterns: consensus expectations underreact to interest rate innovations during expansions but react according to the full information rational expectations paradigm during recessions.

I incorporate extrapolation and time-varying expectation stickiness into a simple present-value model and discipline the model using survey evidence. The model is capable of accounting for novel stylized fact documented in this paper. I also study the implications of the model regarding bond return predictability, the profitability of time-series momentum strategies, and foreign exchange return predictability during the post-2007 period.

Extensions of the model that incorporate rich dynamics of subjective risk premia and models featuring counter-cyclical probability of portfolio adjustment appear to be candidates that could improve the quantitative fit of the model.

[21] A zero lower bound cutoff of USD LIBOR $< 0.5\%$ delivers qualitatively similar results.

APPENDICES

1.A Implications of Rational Expectations Models

1.A.1 Preliminaries

Assuming log-normal foreign and domestic stochastic discount factors (SDFs), the domestic and foreign risk-free rates are defined as:

$$i_t = -\log \mathbb{E}_t[M_{t+1}] = -\mathbb{E}_t[m_{t+1}] - \frac{1}{2}\text{var}_t(m_{t+1}) \tag{1.26}$$

and

$$i_t^* = -\log \mathbb{E}_t[M_{t+1}^*] = -\mathbb{E}_t[m_{t+1}^*] - \frac{1}{2}\text{var}_t(m_{t+1}^*) \tag{1.27}$$

In a setting with complete markets, the SDF is unique. Therefore, the real exchange rate appreciation is equal to the ratio of the SDFs at home and abroad:

$$\frac{S_{t+1}}{S_t} = \frac{M_{t+1}^*}{M_{t+1}} \tag{1.28}$$

Therefore, the expected log exchange rate appreciation (Δs_{t+1}) is:

$$\mathbb{E}_t[m_{t+1}^*] - \mathbb{E}_t[m_{t+1}] = i_t^D - \frac{1}{2}\text{var}_t(m_{t+1}^*) + \frac{1}{2}\text{var}_t(m_{t+1}) \tag{1.29}$$

Using the definition of excess returns, we obtain the following expression for $\mathbb{E}_t[r_{t+1}]$:

$$\mathbb{E}_t[r_{t+1}] = \mathbb{E}_t[\Delta s_{t+1}] - i_t^D = \frac{1}{2}\text{var}_t(m_{t+1}) - \frac{1}{2}\text{var}_t(m_{t+1}^*) \tag{1.30}$$

1.A.2 Habits-based model

The model examined below is based on the framework developed by Verdelhan (2010).

Domestic and foreign log consumption growth is i.i.d.

$$\Delta c_{t+1} = g + \sigma w_{t+1} \tag{1.31}$$

$$\Delta c^*_{t+1} = g^* + \sigma^* w_{t+1} \tag{1.32}$$

Habit-based utility implies the following domestic and foreign SDFs:

$$M_{t+1} = \beta e^{-\gamma \Delta c_{t+1}} \left(\frac{H_{t+1}}{H_t} \right)^{-\gamma} \tag{1.33}$$

and

$$M^*_{t+1} = \beta^* e^{-\gamma \Delta c^*_{t+1}} \left(\frac{H^*_{t+1}}{H^*_t} \right)^{-\gamma} \tag{1.34}$$

where H_t is the surplus consumption ratio.

The log of the surplus consumption ratio evolves according to:

$$h_{t+1} = (1 - \phi)\bar{h} + \phi h_t + v(h_t)\sigma w_{t+1} \tag{1.35}$$

$$h^*_{t+1} = (1 - \phi^*)\bar{h}^* + \phi^* h^*_t + v(h^*_t)\sigma^* w^*_{t+1} \tag{1.36}$$

where the sensitivity function, $v(h_t)$, describes how habits are formed from past aggregate consumption:

$$\begin{aligned} v(h_t) &= \frac{1}{\bar{H}} \sqrt{1 - 2\left(h_t - \bar{h}\right)} - 1, \text{ when } h \leq h_{\max}, \text{ 0 elsewhere} \\ \bar{H} &= \sigma \sqrt{\frac{\gamma}{1 - \phi - B/\gamma}} \end{aligned} \tag{1.37}$$

where $h_{\max} = \bar{h} + (1 - \bar{H}^2)/2$ and B is a preference parameter that governs the cyclical behavior of interest rates. The $v\left(h^*_t\right)$ is defined analogously.

As is standard in the literature, I assume that the preferences and endowments of the domestic and foreign investors are characterized by the same parameters: $\beta^* = \beta$, $\gamma^* = \gamma$, $\phi^* = \phi$, $g^* = g$, $\sigma^* = \sigma$, and $\bar{h}^* = \bar{h}$.

Given the set-up above, the domestic interest rate is:

$$i_t = -\log\beta + \gamma\mathbb{E}_t\left[\Delta c_{t+1}\right] + \gamma\mathbb{E}_t\left[\Delta h_{t+1}\right] - \gamma^2\text{var}_t\left(\Delta c_{t+1} + \Delta h_{t+1}\right) \tag{1.38}$$

$$= -\log\beta + \gamma g - \frac{\gamma^2\sigma^2}{2\bar{H}^2} - B\left(h_t - \bar{h}\right) \tag{1.39}$$

and the foreign interest rate is:

$$i_t^* = -\log\beta + \gamma g - \frac{\gamma^2\sigma^2}{2\bar{H}^2} - B\left(h_t^* - \bar{h}\right) \tag{1.40}$$

Under a symmetric calibration the interest rate differential is:

$$i_t^D = B\left(h_t^* - h_t\right) \tag{1.41}$$

The variance of the log SDF is equal to:

$$\text{var}_t\left(m_{t+1}\right) = \text{var}_t\left(\log\beta - \gamma g + \sigma w_{t+1} - \gamma\left((1-\phi)\bar{h} + (\phi-1)h_t + v(h_t)\sigma w_{t+1}\right)\right) \tag{1.42}$$

$$= \frac{\gamma^2\sigma^2}{\bar{H}^2}\left(1 - 2\left(h_t - \bar{h}\right)\right) \tag{1.43}$$

Therefore, the expected return is equal to:

$$\mathbb{E}_t\left[r_{t+1}\right] = \frac{\gamma^2\sigma^2}{\bar{H}}\left(h_t^* - h_t\right) \tag{1.44}$$

Given the $\mathbb{E}_t[r_{t+1}]$ and i_t^D above, the UIP coefficient is equal to:

$$\beta^F = \frac{B\bar{H}}{\gamma^2\sigma^2} \tag{1.45}$$

The sign of B determines the sign of β^F. In order to replicate the evidence presented in Section 1.3.2, B needs to be negative during expansions and non-negative during recessions.

$B \geq 0$ in equation (1.38) implies that (real) interest rates increase in response to negative consumption shocks. In order to test this prediction of the model, I study the correlation between risk-free rates rates and several proxies of economic activity during recessions. To increase the power of the tests, I only consider proxies that are available at the monthly frequency: industrial production growth and unemployment growth. Real rates are constructed using the following formula:

$$r_t = i_t - \mathbb{E}_t^{SPF}[\pi_{t+1}]$$

where i_t is the nominal interest rate (3-month T-Bill rate) and π_{t+1} is inflation realized three months from time t. The SPF superscript indicates that the expectations are based on the SPF.[22]

The expected growth rates of macroeconomic variables are constructed using the following formula:

$$\mathbb{E}^{SPF}\left[\Delta x_{t+1}\right] = \mathbb{E}^{SPF}\left[x_{t+1}\right] - x_t$$

where x is the unemployment rate or the natural logarithm of industrial production.

The correlations between i_t (r_t) and $\mathbb{E}^{SPF}\left[\Delta x_{t+1}\right]$ and are reported in the table below. The table also reports the p-value values of one-tailed tests. Recessions are identified using the CFNAI.

	Real rates	Nominal rates
$\mathbb{E}^{SPF}\left[\Delta\text{Ind. Prod.}\right]$	0.26^{**}	0.27^{**}
p-value	(0.05)	(0.04)
$-\mathbb{E}^{SPF}\left[\Delta\text{UE}\right]$	0.25^{*}	0.26^{**}
p-value	(0.05)	(0.04)

Negative economic shocks during recessions are associated with lower, not higher interest rates. Therefore, the Verdelhan (2010) model needs to make a counter-factual prediction regarding the behavior of interest rates to generate $\beta_\Delta > 0$. Tests based on macroeconomic

[22] I assign the forecasts in the SPF to the second month in a quarter, e.g. February in Q1, and interpolate between forecasts to generate monthly forecasts. Assigning the SPF forecast to each month in a quarter produces qualitatively similar results.

variables reported at the quarterly frequency and tests using realized growth rates produce correlations similar to the ones reported above.

1.A.3 Long-run Risks

The model examined below is based on the long-run risks framework in Bansal and Shaliastovich (2013). For simplicity, I ignore the distinction between nominal and real variables in the original paper. This assumption does not affect the theoretical conclusions of the model presented in this section.

A generic feature of models with recursive preferences is that the conditional volatility of the SDF and, hence, expected excess return is only affected by the variables that drive the conditional volatility of innovations to consumption (Chernov and Creal, 2020). Therefore, I abstract from the predictable consumption growth component in Bansal and Shaliastovich (2013). The domestic and foreign consumption growth processes take on the following form:

$$\Delta c_{t+1} = g + \sigma_t w_{t+1} \tag{1.46}$$

$$\Delta c^*_{t+1} = g^* + \sigma^*_t w^*_{t+1} \tag{1.47}$$

where w_{t+1} is i.i.d. standard normal and the domestic and foreign volatility evolves according to:

$$\sigma^2_{t+1} = \phi\sigma^2_t + (1-\phi)\varsigma + \omega\eta_{t+1} \tag{1.48}$$

$$\sigma^{*2}_{t+1} = \phi^*\sigma^{*2}_t + (1-\phi^*)\varsigma^* + \omega^*\eta^*_{t+1} \tag{1.49}$$

where η_{t+1} and η^*_{t+1} are i.i.d. standard normal shocks uncorrelated with w_{t+1} and w^*_{t+1}.[23]

Recursive utility implies the following domestic SDF:

$$M_{t+1} = \beta^\theta e^{-\frac{\theta}{\psi}\Delta c_{t+1} + (\theta-1) r_{c,t+1}} \tag{1.50}$$

[23] Equations (1.48) and (1.49) imply that variance can be negative. For ease of exposition, I follow Bansal and Shaliastovich (2013) and proceed as if σ^2_t and σ^{*2}_t are always nonnegative.

where γ is the coefficient of relative risk aversion, $\theta = (1-\gamma)/(1-\psi^{-1})$ with ψ denoting the elasticity of intertemporal substitution, and $r_{c,t+1}$ is the return on an asset that delivers aggregate consumption as its dividend each time period (the wealth portfolio).

From the representative agent's first order condition:

$$\begin{aligned} 1 &= \mathbb{E}_t \left[M_{t+1} R_{c,t+1} \right] &(1.51)\\ &= \beta^{\theta} \mathbb{E}_t \left[e^{-\frac{\theta}{\psi}\Delta c_{t+1} + \theta r_{c,t+1}} \right] &(1.52) \end{aligned}$$

Let pc_t be the log price-consumption ratio. Using the standard Campbell and Shiller (1988) approximation $r_{t+1} = \kappa_0 + \kappa pc_{t+1} - pc_t + \Delta c_{t+1}$, Equation 1.52 can be rewritten as:

$$1 = \beta^{\theta} \mathbb{E}_t \left[e^{(1-\gamma)\Delta c_{t+1} + \theta\kappa_0 + \theta\kappa pc_{t+1} - \theta pc_t} \right] \tag{1.53}$$

I conjecture that the wealth-consumption ratio is an affine function of the only state variable in the model, σ_t^2,

$$pc_t = c - A\sigma_t^2 \tag{1.54}$$

Then,

$$\begin{aligned} 1 &= \beta^{\theta} \mathbb{E}_t \left[e^{(1-\gamma)(g+\sigma_t w_{t+1}) + \theta\kappa_0 + \theta\kappa(c - A\sigma_{t+1}^2) - \theta(c - A\sigma_t^2)} \right] &(1.55)\\ &= \beta^{\theta} \mathbb{E}_t \left[e^{(1-\gamma)g + \frac{(1-\gamma)^2\sigma_t^2}{2} + \theta\kappa_0 + \theta\kappa\left(c - A\left((1-\phi)\varsigma + \phi\sigma_t^2\right)\right) - \theta(c - A\sigma_t^2)} \right] &(1.56) \end{aligned}$$

Ignoring any terms that do not multiply σ_t^2 we have:

$$\mathbb{E}_t \left[e^{\frac{(1-\gamma)^2}{2}\sigma_t^2 - \theta\kappa A\phi\sigma_t^2 + \theta A\sigma_t^2} \right] \tag{1.57}$$

Hence,

$$\frac{(1-\gamma)^2}{2} - \theta\kappa A\phi + \theta A = 0 \tag{1.58}$$

Therefore,

$$A = -\frac{1}{2}\frac{(1-\gamma)(1-1/\psi)}{1-\kappa\phi} \tag{1.59}$$

The risk-free rate is given by:

$$\begin{aligned}
i_t &= -\log \mathbb{E}_t\left[M_{t+1}\right] && (1.60)\\
&= -\log \beta^{\theta}\mathbb{E}_t\left[e^{-\gamma\Delta c_{t+1}+(\theta-1)\left(\kappa_0+\kappa c-A\kappa\sigma_{t+1}^2-c+A\sigma_t^2\right)}\right] && (1.61)\\
&= -\theta\log\beta - \log\mathbb{E}_t\left[e^{-\gamma g-\gamma\sigma_t w_{t+1}+(\theta-1)\left(\kappa_0+\kappa c-A\kappa\left((1-\phi)\varsigma+\phi\sigma_t^2+\omega\eta_{t+1}\right)-c+A\sigma_t^2\right)}\right] && (1.62)\\
&= -\theta\log\beta - \left(-\gamma g+\frac{\gamma^2}{2}\sigma_t^2+(\theta-1)\left(\kappa_0+\kappa c-A\kappa((1-\phi)\varsigma+\phi\sigma_t^2)-c+A\sigma_t^2\right)\right) && (1.63)\\
&- \left(\frac{1}{2}\left(\frac{\gamma^2}{2}+(\theta-1)\kappa A\right)^2\omega^2\right) && (1.64)
\end{aligned}$$

The functional form of c has no bearing on my analysis as c and c^* cancel each other out under a symmetric calibration ($g = g^*$, $\sigma = \sigma^*$, $\phi = \phi^*$, $\varsigma = \varsigma^*$, $\omega = \omega^*$, $\beta = \beta^*$, $\psi = \psi^*$, $\gamma = \gamma^*$).

Under a symmetric calibration,

$$i_t^D = i_t - i_t^* = \left((\theta-1)\,A\left(\kappa\phi-1\right)-\frac{\gamma^2}{2}\right)\left(\sigma_t^2-\sigma_t^{*2}\right) \tag{1.65}$$

The variance of the domestic log SDF is:

$$\mathrm{var}_t(m_{t+1}) = \gamma^2\sigma_t^2 + \left((\theta-1)A\kappa\right)^2\omega^2 \tag{1.66}$$

Therefore, the expected excess return is:

$$\mathbb{E}_t\left[r_{t+1}\right] = \frac{\gamma^2}{2}\left(\sigma_t^2-\sigma_t^{*2}\right) \tag{1.67}$$

Given $\mathbb{E}_t[r_{t+1}]$ and i_t^D above, the UIP coefficient takes on the following form:

$$\beta^F = \frac{\frac{\gamma^2}{2}}{\left((\theta - 1) A (\kappa\phi - 1) - \frac{\gamma^2}{2}\right)} \tag{1.68}$$

Plugging in the expressions for A and θ:

$$\beta^F = \frac{\gamma^2}{\frac{1}{\psi} - \frac{\gamma}{\psi} - \gamma} \tag{1.69}$$

As shown in Equation 1.69, the sign of the β^F in the context of the long-run risks model depends on the relation between γ and $1/\psi$. In order to replicated the evidence presented in Section 1.3.2, the long-run risks model needs to generate (real) interest rates that increase in response to an increase in consumption volatility during recessions.

In order to test this prediction of the model, I examine the correlation between the real 3-month T-Bill rate and a proxy for macroeconomic volatility, VXO^2 (the squared S&P 100 implied volatility). The correlation between the two time series is -0.47 (one-tailed p-value < 0.01). The correlation between the nominal rates and the VXO^2 is even more negative: -0.55 (one-tailed p-value < 0.01). Therefore, high macroeconomic volatility is associated with low, not high interest rates. The long-run risks model, similar to the habits-based model, needs to generate counter-factual behavior of interest rates to match the patterns in foreign exchange return predictability.

1.A.4 Heterogeneous Agent Models

In this section, I consider a domestic economy that consists of two agents: Agent A and Agent B. Each of the agents has recursive preferences, as in Appendix 1.A.3. Following Gârleanu and Panageas (2015), I assume that:

(i) $\gamma^A < \gamma^B$

(ii) $\psi^A > \psi^B$

Consumption growth is i.i.d. with mean μ and variance σ^2.

The foreign economy is structured in the same manner.

In this section, I limit my focus on the behavior of interest rates. A discussion related to counter-cyclical risk premia is presented in the main body of the paper. A standard result for an economy with a representative agent is:

$$i = \beta + \frac{\mu}{\psi} - \frac{1}{2}\gamma\left(1 + \frac{1}{\psi}\right)\sigma^2 \tag{1.70}$$

Let $x_t \equiv \frac{w_A}{w_A + w_B}$ represent the wealth share of the less risk-averse agent, A.

Assuming that ψ and γ are a function of the state variable x_t, the interest rate in a set-up with heterogeneous agents becomes:

$$i(x_t) = \beta + \frac{\mu}{\psi(x_t)} - \frac{1}{2}\gamma(x_t)\left(1 + \frac{1}{\psi(x_t)}\right)\sigma^2 \tag{1.71}$$

In the context of this model, recession correspond to low x_t periods. Therefore, in order to replicate the empirical evidence presented in this paper, the following inequalities must hold:[24]

$$\begin{aligned} \gamma(x_t) &< \frac{2\mu}{\sigma^2} \text{ if } x_t < \bar{x} \\ \gamma(x_t) &> \frac{2\mu}{\sigma^2} \text{ if } x_t > \bar{x} \end{aligned} \tag{1.72}$$

where $\bar{x}$ is the recession threshold.

The two inequalities in (1.72) cannot be true at the same time as $\partial\gamma(x_t)/\partial x_t < 0$. Therefore, the model featuring two agents fails to produce results consistent with those presented in this paper.

The heterogeneous agent model is able to reproduce the results presented in this paper through the introduction of additional agent types. However, the ability of the model to do so hinges on (real) interest rates increasing in response to negative macroeconomic shocks.

[24] As shown by Schneider (2022), $\partial i/\partial x_t < 0$ if $\gamma(x_t) < 2\mu/\sigma^2$.

As shown in Appendix 1.A.2, interest rates do not increase but decrease in response to negative macroeconomic shocks during recessions.

1.B Derivations

1.B.1 FE-on-FR Coefficient

First, I derive expressions for several terms that will come up a number of times throughout this section.

$$\begin{aligned}
&\mathrm{cov}_s\left(\mathbb{E}^S_{t-1}[i^D_t], i^D_{t-m-1}\right) = \mathrm{cov}_s\left(\lambda_s\tilde{\phi}\sum_{j=0}^{m-1}(1-\lambda_s)^j\,\tilde{\phi}^j i^D_{t-j-1} + (1-\lambda_s)^m\,\tilde{\phi}^{m+1} i^D_{t-m-1}, i^D_{t-m-1}\right) = \\
&= \left[\frac{\lambda_s\tilde{\phi}\phi^m\left(1-\left((1-\lambda_s)\frac{\tilde{\phi}}{\phi}\right)^{m-1}\right)}{1-(1-\lambda_s)\frac{\tilde{\phi}}{\phi}} + (1-\lambda_s)^m\tilde{\phi}^{m+1}\right]\mathrm{var}_s(i^D_t)
\end{aligned} \tag{1.73}$$

$$\begin{aligned}
&\mathrm{cov}_s\left(\mathbb{E}^S_{t-1}[i^D_t], i^D_t\right) = \mathrm{cov}_s\left(\lambda_s\tilde{\phi}\sum_{j=0}^{m-1}(1-\lambda_s)^j\,\tilde{\phi}^j i^D_{t-j-1} + (1-\lambda_s)^m\,\tilde{\phi}^{m+1} i^D_{t-m-1}, i^D_t\right) = \\
&= \left[\frac{\lambda_s\tilde{\phi}\phi\left(1-\left((1-\lambda_s)\tilde{\phi}\phi\right)^{m-1}\right)}{1-(1-\lambda_s)\tilde{\phi}\phi} + (1-\lambda_s)^m\tilde{\phi}^{m+1}\phi^{m+1}\right]\mathrm{var}_s(i^D_t)
\end{aligned} \tag{1.74}$$

$$\begin{aligned}
&\mathrm{var}_s\left(\mathbb{E}^S_{t-1}[i^D_t]\right) = \mathrm{var}_s\left(\lambda_s\tilde{\phi}\sum_{j=0}^{m-1}(1-\lambda_s)^j\,\tilde{\phi}^j i^D_{t-j-1} + (1-\lambda_s)^m\,\tilde{\phi}^{m+1} i^D_{t-m-1}\right) = \\
&= \lambda_s^2\tilde{\phi}^2\mathrm{var}_s\left(\sum_{j=0}^{m-1}(1-\lambda_s)^j\,\tilde{\phi}^j i^D_{t-j-1}\right) + \\
&+ (1-\lambda_s)^{2m}\,\tilde{\phi}^{2m+2}\mathrm{var}_s\left(i^D_t\right) + 2\lambda_s(1-\lambda_s)^m\tilde{\phi}^{m+2}\left(\frac{\phi^m\left(1-\left((1-\lambda_s)\frac{\tilde{\phi}}{\phi}\right)^{m-1}\right)}{1-(1-\lambda_s)\frac{\tilde{\phi}}{\phi}}\right)\mathrm{var}_s\left(i^D_t\right)
\end{aligned} \tag{1.75}$$

The variance of the finite sum is equal to the sum of the variance terms and twice the sum of the covariance terms.

The sum of the variances terms is:

$$\frac{1-\left((1-\lambda_s)^2\tilde{\phi}^2\right)^{m-1}}{1-(1-\lambda_s)^2\tilde{\phi}^2} \tag{1.76}$$

The sum of the covariances term is:

$$\begin{aligned}&\frac{(1-\lambda_s)\tilde{\phi}\phi}{1-(1-\lambda_s)\tilde{\phi}\phi}\left[1-\left((1-\lambda_s)\tilde{\phi}\phi\right)^{m-2}\right]+\frac{(1-\lambda_s)^3\tilde{\phi}^3\phi}{1-(1-\lambda_s)\tilde{\phi}\phi}\left[1-\left((1-\lambda_s)\tilde{\phi}\phi\right)^{m-3}\right]+\\&+\frac{(1-\lambda_s)^5\tilde{\phi}^5\phi}{1-(1-\lambda_s)\tilde{\phi}\phi}\left[1-\left((1-\lambda_s)\tilde{\phi}\phi\right)^{m-4}\right]+...\end{aligned} \tag{1.77}$$

The sum of the positive terms in Equation (1.77) is:

$$\frac{(1-\lambda_s)\tilde{\phi}\phi}{1-(1-\lambda_s)\tilde{\phi}\phi}\frac{1-\left((1-\lambda_s)^2\tilde{\phi}^2\right)^{m-2}}{1-(1-\lambda_s)^2\tilde{\phi}^2} \tag{1.78}$$

The sum of the negative terms in Equation (1.77) is:

$$\frac{\left((1-\lambda_s)\tilde{\phi}\phi\right)^{m-1}}{1-(1-\lambda_s)\tilde{\phi}\phi}\frac{1-\left((1-\lambda_s)\frac{\tilde{\phi}}{\phi}\right)^{m-2}}{1-(1-\lambda_s)\frac{\tilde{\phi}}{\phi}} \tag{1.79}$$

I plug Equations (1.76), (1.78), and (1.79) into Equation (1.75) and obtain an expression for the variance of $\mathbb{E}^S_{t-1}[i^D_t]$. The FE-on-FR coefficient is

$$\beta^{CG}_s=\frac{\operatorname{cov}\left(i^D_{t+1}-\mathbb{E}^S_t[i^D_{t+1}],\mathbb{E}^S_t[i^D_{t+1}]-\mathbb{E}^S_{t-1}[i^D_{t+1}]\right)}{\operatorname{var}\left(\mathbb{E}^S_t[i^D_{t+1}]-\mathbb{E}^S_{t-1}[i^D_{t+1}]\right)} \tag{1.80}$$

First, I focus on the covariance term:

$$\begin{aligned}&\operatorname{cov}_s(\phi i^D_t+\varepsilon_{t+1}-\tilde{\phi}\lambda_s i^D_t-(1-\lambda_s)\tilde{\phi}\mathbb{E}^S_{t-1}[i^D_t]-(1-\lambda_s)^{m+1}\,\tilde{\phi}^{m+1}i^D_{t-m}+(1-\lambda_s)^{m+1}\,\tilde{\phi}^{m+2}i^D_{t-m-1},\\&\tilde{\phi}\lambda_s i^D_t-\lambda_s\tilde{\phi}\mathbb{E}^S_{t-1}[i^D_t]+(1-\lambda_s)^{m+1}\tilde{\phi}^{m+1}i^D_{t-m}-(1-\lambda_s)^{m+1}\tilde{\phi}^{m+2}i^D_{t-m-1})\end{aligned} \tag{1.81}$$

I evaluate the covariance and plug in the expression derived in Equations (1.73), (1.74) and (1.75) to obtain an expression for the numerator of β^{CG}.

The variance term in Equation (1.80) can be written as:

$$\text{var}_s\left(\tilde{\phi}\lambda_s i_t^D - \lambda_s\tilde{\phi}\mathbb{E}_{t-1}^S[i_t^D] + (1-\lambda_s)^{m+1}\tilde{\phi}^{m+1}i_{t-m}^D - (1-\lambda_s)^{m+1}\tilde{\phi}^{m+2}i_{t-m-1}^D\right) \tag{1.82}$$

I evaluate the variance and plug in the expressions derived in Equations (1.73), (1.74) and (1.75) to obtain an expression for the denominator of β^{CG}.

1.B.2 UIP Coefficient

The deviation from rational expectations in Equation (1.18) is equivalent to:

$$\begin{aligned}&\mathbb{E}_t^P[s_{t+1}] - \mathbb{E}_t^S[s_{t+1}] = -\mathbb{E}_t^P[i_{t+1}^D] - \frac{1}{1-\tilde{\phi}}\mathbb{E}_t^P\left[\mathbb{E}_{t+1}^S[i_{t+2}^D]\right] + \\ &\mathbb{E}_t^S[i_{t+1}^D] + \frac{\tilde{\phi}}{1-\tilde{\phi}}\mathbb{E}_t^S[i_{t+1}^D]\end{aligned} \tag{1.83}$$

Note that the law of iterated expectations does not hold for $\mathbb{E}_t^P\left[\mathbb{E}_{t+1}^S[i_{t+2}^D]\right]$. Instead, $\mathbb{E}_t^P\left[\mathbb{E}_{t+1}^S[i_{t+2}^D]\right] = \lambda_s\tilde{\phi}\mathbb{E}_t^P[i_{t+1}^D] + (1-\lambda_s)\tilde{\phi}\mathbb{E}_t^S[i_{t+1}^D]$. Using the definitions of $\mathbb{E}^S$ and $\mathbb{E}^P$, Equation (1.83) can be rewritten as:

$$\left(1 + \frac{\tilde{\phi}\lambda_s}{1-\tilde{\phi}}\right)\left(\mathbb{E}_t^S[i_{t+1}^D] - \mathbb{E}_t^P[i_{t+1}^D]\right) \tag{1.84}$$

Given the expressions above, the UIP coefficient can be written as:

$$\beta_s^F = \frac{\text{cov}_s\left(\left(1+\frac{\tilde{\phi}\lambda_s}{1-\tilde{\phi}}\right)\left(\mathbb{E}_t^S[i_{t+1}^D] - \mathbb{E}_t^P[i_{t+1}^D]\right), i_t^D\right)}{\text{var}_s\left(i_t^D\right)} \tag{1.85}$$

The covaraince term can be written as:

$$\left(1 + \frac{\tilde{\phi}\lambda}{1-\tilde{\phi}}\right)\left(\text{cov}_s\left(i_t^D, \mathbb{E}_t^S[i_{t+1}^D]\right) - \phi\text{var}_s(i_t^D)\right) \tag{1.86}$$

I plug the definition of $\text{cov}_s\left(i_t^D, \mathbb{E}_t^S[i_{t+1}^D]\right)$ into the equation above and obtain the following expression for the UIP coefficient:

$$\beta_s^F = \left(1 + \frac{\tilde{\phi}\lambda}{1-\tilde{\phi}}\right)\left(\lambda_s\tilde{\phi}\frac{1-\left((1-\lambda_s)\tilde{\phi}\phi\right)^{m-1}}{1-(1-\lambda_s)\tilde{\phi}\phi} + (1-\lambda_s)^m\,\tilde{\phi}^{m+1}\phi^m - \phi\right) \tag{1.87}$$

1.B.3 Bond Return Predictability Coefficient

The expected excess bond return under measure $\mathbb{P}$ is equivalent to:

$$\mathbb{E}_t^P[p_{t+1}^N] - \mathbb{E}_t^S[p_{t+1}^N] = \left(1 + \frac{\bar{\lambda}_s\tilde{\phi}(1-\tilde{\phi}^{N-1})}{1-\tilde{\phi}}\right)\left(\mathbb{E}_t^S[i_{t+1}] - \mathbb{E}_t^P[i_{t+1}]\right) \tag{1.88}$$

Therefore, the covariance term in Equation (1.23) can be written as:

$$\text{cov}_s\left(\left(1 + \frac{\lambda_s\tilde{\phi}(1-\tilde{\phi}^{N-1})}{1-\tilde{\phi}}\right)\left(\mathbb{E}_t^S[i_{t+1}] - \mathbb{E}_t^P[i_{t+1}]\right), \frac{1}{N}\left[\left(\frac{1-\tilde{\phi}^N}{1-\tilde{\phi}}\right)\mathbb{E}_t^S[i_{t+1}] + (1-N)i_t\right]\right) \tag{1.89}$$

This expression is equivalent to:

$$\frac{1}{N}\left(1 + \frac{\lambda_s\tilde{\phi}(1-\tilde{\phi}^{N-1})}{1-\tilde{\phi}}\right)\text{cov}_s\left(\mathbb{E}_t^S[i_{t+1}] - \mathbb{E}_t^P[i_{t+1}], \left(\frac{1-\tilde{\phi}^N}{1-\tilde{\phi}}\right)\mathbb{E}_t^S[i_{t+1}] + (1-N)i_t\right) \tag{1.90}$$

Using the definitions of $\mathbb{E}^S$ and $\mathbb{E}^P$, the covariance term in Equation (1.90) can be written as:

$$\begin{aligned}&\text{cov}_s(\lambda_s\tilde{\phi}\sum_{j=0}^{m-1}(1-\lambda_s)^j\,\tilde{\phi}^j i_{t-j} + (1-\lambda_s)^m\,\tilde{\phi}^{m+1}i_{t-m} - \phi i_t,\\ &\left(\frac{1-\tilde{\phi}^N}{1-\tilde{\phi}}\right)\left(\lambda_s\tilde{\phi}\sum_{j=0}^{m-1}(1-\lambda_s)^j\,\tilde{\phi}^j i_{t-j} + (1-\lambda_s)^m\,\tilde{\phi}^{m+1}i_{t-m}\right) + (1-N)i_t)\end{aligned} \tag{1.91}$$

The covaraince term is equivalent to:

$$\left(\frac{1-\tilde{\phi}^N}{1-\tilde{\phi}}\right)\text{var}_s\left(\mathbb{E}_t^S[i_{t+1}]\right) + \left(1 - N - \phi\left(\frac{1-\tilde{\phi}^N}{1-\tilde{\phi}}\right)\right)\text{cov}_s\left(i_t, \mathbb{E}_t^S[i_{t+1}]\right) - \phi(1-N)\text{var}_s(i_t) \tag{1.92}$$

Plugging in the expressions for the $\text{var}_s\left(\mathbb{E}_t^S[i_{t+1}]\right)$ and the $\text{cov}_s\left(i_t, \mathbb{E}_t^S[i_{t+1}]\right)$ terms, derived analogously to expressions in Equations (1.74) and (1.75), I obtain an expression for the numerator of the bond return predictability coefficient.

The variance term in Equation (1.23) can be written as:

$$\frac{1}{N^2}\left(\left(\frac{1-\tilde{\phi}^N}{1-\tilde{\phi}}\right)^2 \text{var}_s\left(\mathbb{E}_t^S[i_{t+1}]\right) + (1-N)^2\text{var}_s\left(i_t\right) + 2\left(\frac{1-\tilde{\phi}^N}{1-\tilde{\phi}}\right)(1-N)\text{cov}_s\left(\mathbb{E}_t^S[i_{t+1}], i_t\right)\right) \tag{1.93}$$

Plugging in the expressions for $\text{var}_s\left(\mathbb{E}_t^S[i_{t+1}]\right)$ and $\text{cov}_s\left(i_t, \mathbb{E}_t^S[i_{t+1}]\right)$ terms above provides us with an expression for the denominator of β_s^{CS}.

1.B.4 Time-Series Momentum

We are interested in the following covariance term:[25]

$$\text{cov}\left(r_t, \mathbb{E}_t^P[r_{t+1}]\right) \tag{1.94}$$

First, I derive an expression for r_t:

$$\begin{aligned} r_t = s_t - s_{t-1} - i_{t-1}^D = -i_t^D - \frac{1}{1-\tilde{\phi}}\mathbb{E}_t^S\left[i_{t+1}^D\right] + i_{t-1}^D + \frac{1}{1-\tilde{\phi}}\mathbb{E}_{t-1}^S\left[i_t^D\right] - i_{t-1}^D = \\ - i_t^D - \frac{1}{1-\tilde{\phi}}\left(\mathbb{E}_t^S\left[i_{t+1}^D\right] - \mathbb{E}_{t-1}^S\left[i_t^D\right]\right) \end{aligned} \tag{1.95}$$

The expression for $\mathbb{E}_t^P\left[r_{t+1}\right]$ is given in Equation 1.84.

Now, I evaluate the covariance term:

$$\begin{aligned} \left(1+\frac{\lambda_s\tilde{\phi}}{1-\tilde{\phi}}\right)\text{cov}\left(\mathbb{E}_t^S\left[i_{t+1}^D\right] - \mathbb{E}_t^P\left[i_{t+1}^D\right], -i_t^D - \frac{1}{1-\tilde{\phi}}\left(\mathbb{E}_t^S\left[i_{t+1}^D\right] - \mathbb{E}_{t-1}^S\left[i_t^D\right]\right)\right) = \\ = -\beta^F + \left(1+\frac{\lambda_s\tilde{\phi}}{1-\tilde{\phi}}\right)\frac{1}{1-\tilde{\phi}}\text{cov}\left(\mathbb{E}_t^P\left[i_{t+1}^D\right] - \mathbb{E}_t^S\left[i_{t+1}^D\right], \mathbb{E}_t^S\left[i_{t+1}^D\right] - \mathbb{E}_{t-1}^S\left[i_t^D\right]\right) \end{aligned} \tag{1.96}$$

[25]For ease of exposition, I assume that the look-back period is one month.

where β^F is the UIP coefficient.

Let's focus on the $\text{cov}\left(\mathbb{E}_t^P\left[i_{t+1}^D\right] - \mathbb{E}_t^S\left[i_{t+1}^D\right], \mathbb{E}_t^S\left[i_{t+1}^D\right] - \mathbb{E}_{t-1}^S\left[i_t^D\right]\right)$ term:

$$
\begin{aligned}
&\text{cov}\left(\mathbb{E}_t^S\left[i_{t+1}^D\right] - \mathbb{E}_t^P\left[i_{t+1}^D\right], \mathbb{E}_t^S\left[i_{t+1}^D\right] - \mathbb{E}_{t-1}^S\left[i_t^D\right]\right) = \\
&\text{cov}\left(\lambda_s\tilde{\phi} i_t^D + (1-\lambda_s)\,\tilde{\phi}\mathbb{E}_{t-1}^S\left[i_t^D\right] - \phi i_t^D, \lambda_s\tilde{\phi} i_t^D + (1-\lambda_s)\tilde{\phi}\mathbb{E}_{t-1}^S\left[i_t^D\right] - \mathbb{E}_{t-1}^S\left[i_t^D\right]\right) = \\
&\frac{\lambda_s\tilde{\phi}\beta^F}{1+\frac{\lambda_s\tilde{\phi}}{1-\tilde{\phi}}} + \left((1-\lambda_s)\tilde{\phi} - 1\right)\text{cov}\left((\lambda_s\tilde{\phi} - \phi) i_t^D + (1-\lambda_s)\tilde{\phi}\mathbb{E}_{t-1}^S\left[i_t^D\right], \mathbb{E}_{t-1}^S\left[i_t^D\right]\right) = \\
&\frac{\lambda_s\tilde{\phi}\beta^F}{1+\frac{\lambda_s\tilde{\phi}}{1-\tilde{\phi}}} + \left((1-\lambda_s)\tilde{\phi} - 1\right)\left((\lambda_s\tilde{\phi} - \phi)\text{cov}\left(i_t^D, \mathbb{E}_{t-1}^S[i_t^D]\right) + (1-\lambda_s)\tilde{\phi}\text{var}\left(\mathbb{E}_{t-1}^S\left[i_t^D\right]\right)\right)
\end{aligned}
\tag{1.97}
$$

The expressions for $\text{cov}\left(i_t^D, \mathbb{E}_{t-1}^S[i_t^D]\right)$ and $\text{var}\left(\mathbb{E}_{t-1}^S\left[i_t^D\right]\right)$ are given in Equations (1.73) and (1.74), respectively.

1.C Equilibrium Model

1.C.1 Preliminaries

I build a stylized model of exchange rate determination. The model extends the Gabaix and Maggiori (2015) segmented markets model by relaxing the assumption that the expectations of the intermediaries are fully rational. I keep the model as simple as possible to assess how far deviations from rational expectations take us towards accounting for the foreign exchange return predictability patterns documented in Section 1.3.

Agents and Market Structure

The model features two types of agents: representative foreign and domestic households and a continuum of overlapping generations of competitive intermediaries who belong to families indexed by $i \in [0, 1]$. All agents live for two periods. The markets are segmented, as in Gabaix and Maggiori (2015). Households can only access international capital markets through intermediaries. Intermediaries facilitate international capital flows resulting from households' consumption-savings decisions by engaging in carry trade. Consequently, the exchange rate is pinned down by the intermediaries' first-order condition. Haddad and Muir (2021) rank foreign exchange as one of the most highly intermediated asset classes, which lends credence to the segmented markets assumption. The main purpose of the assumption in this paper is to simplify the analysis by separating the portfolio choice problem from the consumption-savings decision.

1.C.2 Households

Households consume three types of goods: nontradables, foreign tradables and domestic tradables. Households trade in international goods markets, but not in international capital markets. Consequently, households invest with the intermediaries in risk-free bonds denominated in their own currency. The interest rates are determined by households' Euler equation.

The two representative households in my economy have the following exogenous, strictly positive endowments:

$$\left\{Y_{NT,\tau}, Y_{D,\tau}, Y^*_{NT,\tau}, Y^*_{F,\tau}\right\}_{\tau=t}^{t+1} \tag{1.98}$$

where $Y_{NT,t}$ and $Y^*_{NT,t}$ are the domestic and foreign households' respective endowments of nontradables, $Y_{D,t}$ is the domestic households' endowment of tradables, and $Y^*_{F,t}$ is the foreign households' endowment of tradables.

The foreign households derive utility from consumption according to:

$$\theta \log C^*_t + \frac{1}{\delta^*_t}\mathbb{E}^S_t\left[\theta \log C^*_{t+1}\right] \tag{1.99}$$

where C^* is the consumption basket defined as:

$$C^*_t = \left((C^*_{NT,t})^{\chi_t}(C^*_{F,t})^{a}(C^*_{D,t})^{\zeta}\right)^{\frac{1}{\varsigma_t}} \tag{1.100}$$

where $C^*_{NT,t}$ is the foreign households' consumption of nontradable goods, $C^*_{F,t}$ is the foreign households' consumption of foreign tradable goods, and $C^*_{D,t}$ is the foreign households' consumption of domestic tradable goods. χ_t, a, and ζ are preference parameters and $\chi_t + a + \zeta = \varsigma_t$.[26] For expositional simplicity, I assume that the preference parameters for the domestic and foreign households are identical. The nontradable good in the numerariere in each country and $p^*_{NT} = 1$.

Domestic households' consumption basket is defined as:

$$C_t = \left((C_{NT,t})^{\chi_t}(C_{D,t})^{a}(C_{F,t})^{\zeta}\right)^{\frac{1}{\varsigma_t}} \tag{1.101}$$

I allow households' expectations to deviate from rational expectations. Hence, households maximize their expected utility under the subjective measure $\mathbb{S}$. However, within the context of the additional assumptions I make, the fact that the households' expectations deviate from rational expectations does not influence the implications of my model regarding exchange rate movements.

I deviate from the usual assumption in the asset pricing literature by assuming that the time preference parameter δ^* is time-varying. There is some empirical support for this assumption. For instance, Meier and Sprenger (2015) find that the within-individual correlation of the time discount factor is about 25%.

Households can trade in both domestic and foreign tradable goods but can only trade in their own nontradable goods market. The households can only borrow and lend through risk-free bonds denominated in their own currency, i.e. domestic households cannot directly borrow from or lend to foreign households. Such trades need to go through the intermediary.

The foreign households' optimization problem can be written as:

$$\max_{C^*_{NT,t},C^*_{D,t},C^*_{F,t}} \left\{\theta \log C^*_t + \frac{1}{\delta^*_t}\mathbb{E}^S_t\left[\theta \log C^*_{t+1}\right]\right\} \tag{1.102}$$

[26]The assumption that ζ and a are constant is made to emphasize the role of deviations of rational expectations. Time-variation in the ζ parameter would create a stochastic bonds supply effect that causes the UIP to fail. By assuming a constant ζ, I abstract from that effect.

subject to Equation (1.100) and the following budget constraint:

$$\sum_{\tau=t}^{t+1} e^{-i_{\tau}^{*}(\tau-t)} \left(Y_{NT,\tau} + p_{F,\tau}Y_{F,\tau}\right) = \sum_{\tau=t}^{t+1} e^{-i_{\tau}^{*}(\tau-t)} \left(C_{NT,\tau}^{*} + p_{D,\tau}^{*}C_{D,\tau}^{*} + p_{F,\tau}^{*}C_{F,\tau}^{*}\right) \tag{1.103}$$

where $p_{D,t}^{*}$ is the foreign price of domestic tradables, $p_{F,t}^{*}$ is the foreign price of foreign tradables.

The foreign households' static optimization problem can be written as:

$$\max_{C_{NT,t}^{*},C_{D,t}^{*},C_{F,t}^{*}} \left\{\chi \log C_{NT,t}^{*} + a \log C_{F,t}^{*} + \zeta \log C_{D,t}^{*} + \mu_{t}^{*}\left(CE_{t}^{*} - C_{NT,t}^{*} - p_{D,t}^{*}C_{D,t}^{*} - p_{F,t}^{*}C_{F,t}^{*}\right)\right\} \tag{1.104}$$

where μ_{t}^{*} is the Lagrange multiplier, CE_{t}^{*} is the aggregate consumption expenditure of foreign households, which I take as given.

The first order condition for tradables gives us the following relation: $\chi_{t} = \mu_{t}^{*}C_{NT,t}^{*}$. Following Gabaix and Maggiori (2015), I assume that the endowment of nontradables is such that $Y_{NT,t} = \chi_{t}$. This assumption together with the market clearing condition for foreign nontradables, $Y_{NT,t}^{*} = \mu_{t}^{*}C_{NT,t}^{*}$, leads to $\mu_{t}^{*} = 1$ is all states. With this assumption, the value of foreign country's imports is:

$$\zeta = p_{D,t}^{*}C_{D,t}^{*} \tag{1.105}$$

Given the symmetric nature of the model, domestic households' imports are: $\zeta = p_{F,t}C_{F,t}$.

Given the two expressions for imports, foreign country's net exports, expressed in units of domestic currency, are:

$$NX_{t}^{*} = \zeta - \zeta S_{t} = \zeta\left(1 - S_{t}\right) \tag{1.106}$$

Rewriting S_{t} as $e^{s_{t}}$ where s_{t} is the log exchange rate and using the approximation $e^{x} \approx 1 + x$, foreign household's net exports can be written as:

$$NX_{t}^{*} = -\zeta s_{t} \tag{1.107}$$

Foreign household's net imports (the negative of net exports) are equal to foreign bond supply. Therefore, taking the negative of Equation (1.107) gives us the expression for foreign bond supply denominated in domestic currency units (ζs_t) that is used in the main body of the paper.

Under a regime of autarky, the exchange rate is constant and equal to 1. This mirrors the constant exchange rate obtained by Pavlova and Rigobon (2007) in the absence of home bias in consumption.

Households' Euler equation based on the optimization problem in Equation (1.102) takes on the following form:

$$1 = \mathbb{E}_t^S \left[\frac{1}{\delta_t^*} \frac{U'\left(C_{NT,t+1}^*\right)}{U'\left(C_{NT,t}^*\right)} \right] = \mathbb{E}_t^S \left[\frac{1}{\delta_t^*} e^{i_t^*} \frac{\frac{\chi_{t+1}}{C_{NT,t+1}^*}}{\frac{\chi_t}{C_{NT,t}^*}} \right] = \frac{e^{i_t^*}}{\delta_t^*} \tag{1.108}$$

The equation above implies that changes in the foreign interest rate is driven by changes in the foreign households' time discount factor. This result is driven by the assumption that $C_{NT,t} = \chi_t$ and serves to simplify the portion of the model related to the households, which is not central to my analysis.

Following the same logic we derive an identical expression for the domestic interest rate:

$$e^{i_t} = \delta_t \tag{1.109}$$

In order to keep my analysis as simple as possible, I assume that the subjective discount parameters for the two countries evolve according to the following stationary processes:

$$\delta_t = \phi \delta_{t-1} + (1-\phi)\,\bar{\delta} + \nu_t^{\delta} \tag{1.110}$$

$$\delta_t^* = \phi \delta_{t-1}^* + (1-\phi)\,\bar{\delta} + \nu_t^{\delta^*} \tag{1.111}$$

where $\nu_t^{\delta} \sim N(0, \sigma_{\delta,t}^2)$ and $\nu_t^{\delta^*} \sim N(0, \sigma_{\delta,t}^2)$.

Based on this assumption, the interest rate evolves according to the AR(1) process shown in the main body of the paper.

1.C.3 Intermediaries

Intermediary i is born at time t with zero initial wealth. The intermediary inherits $\mathbb{E}_{t-1}^{i}[i_t^D]$ from the generation $t-1$ intermediary i. Prior to the realization of i_t^D, the agent chooses the probability of updating her beliefs $\lambda_{i,t}$. With probability $\lambda_{i,t}$ the agent incorporates the realization of i_t^D into her information set and with probability $1-\lambda_{i,t}$ the agent uses $\mathbb{E}_{t-1}^{i}[i_t^D]$ when forming expectations regarding i_{t+h}^D.

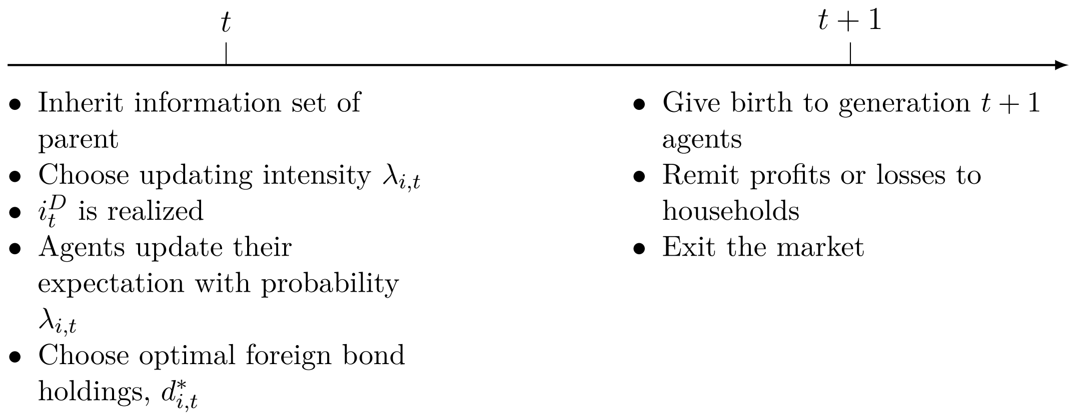
After making the information acquisition choice, the intermediary facilitates international capital flows by engaging in carry trade. In period $t+1$, the profits or losses of generation t intermediaries are realized. Subsequently, the generation t intermediaries remit their profits or losses to the domestic households,[27] and exit the market. The structure of the economy removes the need to keep track of agents' wealth as a state variable. Bacchetta and Van Wincoop (2006) employ the same structure. The life cycle of a generation t intermediary is summarized in Figure 1.C.1.

Carry Trade

Intermediaries in my model pursue a zero-cost investment strategy whereby they borrow in a low interest rate currency and lend in a high interest rate currency. The investors make a single portfolio choice: the amount invested in foreign currency bonds, $d_{i,t}^*$, expressed in units of domestic currency. For a positive $d_{i,t}^*$ the investment position can be illustrated as follows: the investor borrows $d_{i,t}^*$ units in domestic currency and uses the $d_{i,t}^*$ units of domestic currency to purchase $d_{i,t}^*\left(1/S_t\right)$ units of foreign currency. This amount generates $d_{i,t}^*\left(1/S_t\right)\exp\left(i_t^*\right)$ units of foreign currency at time $t+1$. At time $t+1$ the agent also needs to repay $d_{i,t}^*\exp\left(i_t\right)$ units of domestic currency. Thus, the agent uses her foreign currency

[27] Intermediaries that experience negative profits are bailed out by the households.

Figure 1.C.1. Generation t Intermediary's Life Cycle



This figure summarizes the life cycle of an intermediary born in period t, called generation t intermediary.

holdings to purchase $d_{i,t}^* \left(S_{t+1}/S_t\right)$ units of domestic currency. The net end-of-life profit $\pi_{i,t+1}$ is then:

$$\pi_{i,t+1} = d_{i,t}^* \exp\left(i_t\right)\left(\exp\left(s_{t+1} - s_t - i_t^D\right) - 1\right) \tag{1.112}$$

The first-order approximation of portfolio return allows $\pi_{i,t+1}$ to be written as $d_{i,t}^* r_{t+1}$ where r_{t+1} is defined in Equation (1.2). The intermediaries are risk-neutral but incur quadratic trading costs $\frac{\psi}{2}\left(d_{i,t}^*\right)^2$ when submitting their orders for foreign bonds. The $\psi > 0$ parameter governs how costly trading is.[28]

1.C.3.1 Foreign Exchange Market Equilibrium

I solve the model by backward induction. I first solve for agent i's optimal portfolio at time t and substitute in that solution into the optimal updating probability choice problem.

Intermediary i's optimization problem at time t is given by:

$$\max_{d_{i,t}^*} \mathbb{E}_t^i \left[d_{i,t}^* r_{t+1} - \frac{\psi}{2}\left(d_{i,t}^*\right)^2\right] \tag{1.113}$$

[28] This specification can be regarded as a reduced-form way of capturing the risk aversion of agents with CARA utility. If $\psi = \gamma \mathrm{var}_t\left(s_{t+1}\right)$, we obtain the standard mean-variance utility. The choice of a constant ψ allows me to abstract from the effects of changes in the conditional payoff variance.

where $\mathbb{E}_t^i$ reflects the fact that the expectations are taken under intermediary i's subjective measure.

Using the first-order condition of intermediary i's optimization problem, we obtain demand for foreign bonds:

$$d_{i,t}^* = \frac{\mathbb{E}_t^i\left[r_{t+1}\right]}{\psi} = \frac{\mathbb{E}_t^i\left[s_{t+1}\right] - s_t - i_t^D}{\psi} \tag{1.114}$$

Bond demand, aggregated across all agents, is:

$$\int_0^1 d_{i,t}^* di \equiv d_t^* = \frac{\mathbb{E}_t^S\left[r_{t+1}\right]}{\psi} = \frac{\mathbb{E}_t^S\left[s_{t+1}\right] - s_t - i_t^D}{\psi} \tag{1.115}$$

where the $\mathbb{E}^S$ operator indicates consensus expectations under subjective measure $\mathbb{S}$.

Foreign bond supply in units of domestic currency is ζs_t, as derived in Section 1.C.2.

Given the expressions for foreign bond supply and foreign bond demand in Equation (1.115), the market-clearing condition becomes:

$$\mathbb{E}_t^S\left[s_{t+1}\right] - s_t - i_t^D = \psi\zeta s_t \tag{1.116}$$

The market for domestic bonds clears by Walras' law.

I iterate the market clearing in Equation (1.116) forward and obtain the following expression for the log exchange rate:

$$s_t = -\sum_{j=0}^{T} \Theta^{j+1}\mathbb{E}_t^S[i_{t+j}^D] - \Theta^{T+1}\mathbb{E}_t^S[s_{t+T}] \tag{1.117}$$

where $\Theta \equiv (1+\psi\zeta)^{-1} < 1$ is a constant discount factor and $\mathbb{E}_t^S[s_{t+T}]$ is the expected long-run exchange rate.

Letting $T \to \infty$ and assuming that the PPP hods in the long-run (as in Dahlquist and Pénasse, 2022), the expression in Equation (1.117) is equivalent to:

$$s_t = -\Theta i_t^D - \frac{\Theta}{1-\tilde{\phi}\Theta}\mathbb{E}_t^S[i_{t+1}^D] \tag{1.118}$$

Having solved for optimal portfolios and the equilibrium exchange rate, now I can solve for agent i's optimal updating probability choice.

A generation t intermediary's expected utility at the beginning of period time t takes on the following form:

$$U_{t-1}^i = \mathbb{E}_{t-1}^i\left[\mathbb{E}_t^i\left[r_{t+1}\right]d_{i,t}^* - \frac{\psi}{2}\left(d_{i,t}^*\right)^2\right] - C(\lambda_{i,t}) \tag{1.119}$$

The U_{t-1}^i notation reflects the fact that intermediary i does not have access to any time t information when choosing the optimal probability of updating her beliefs.

As in Ball, Mankiw, and Reis (2005), expectation stickiness can be motivated by costly information gathering. Along these lines, I assume that agents bear an information processing cost. $C(\lambda_{i,t})$ is the cost associated with updating probability $\lambda_{i,t}$. I assume that $C\left(\lambda_{i,t}\right)$ is strictly increasing and convex.

For the solution of the updating probability choice problem I consider the following functional form for $C\left(\lambda_t^i\right)$:

$$C\left(\lambda_{i,t}\right) = \frac{\xi}{\kappa+1}\lambda_{i,t}^{\kappa+1} \tag{1.120}$$

where $\xi > 0$ and $\kappa \geq 0$. The ξ parameter shifts the marginal cost of information acquisition and the κ parameter influences the local curvature of the cost function.

As shown in Appendix 1.D.1, Intermediary i's optimal choice of λ solves the following equation:

$$\frac{1}{\psi}(\Theta-1)\frac{\tilde{\phi}^2\Theta(\Theta-1)}{1-\Theta\tilde{\phi}}\sigma_{i,t|t-1}^2 = C'\left(\lambda_{i,t}\right) - \frac{1}{\psi}\left(\frac{\tilde{\phi}^2\Theta(\Theta-1)}{1-\Theta\tilde{\phi}}\right)^2\lambda_{i,t}\sigma_{i,t|t-1}^2 \tag{1.121}$$

where $C^{'}(\lambda_{i,t})$ is the first derivative of the cost function and $\sigma^2_{i,t|t-1}$ the value of the time t variance given intermediary i's information set.

The variable of interest in the expression above is $\sigma^2_{i,t|t-1}$. For the purposes of building intuition, the solution of Equation (1.121) for $\kappa = 1$ is presented below.[29]

$$\lambda_{i,t} = \frac{\frac{1}{\psi}(\Theta - 1)\frac{\tilde{\phi}^2\Theta(\Theta-1)}{1-\Theta\tilde{\phi}}\sigma^2_{i,t|t-1}}{\xi - \frac{1}{\psi}\left(\frac{\tilde{\phi}^2\Theta(\Theta-1)}{1-\Theta\tilde{\phi}}\right)^2\sigma^2_{i,t|t-1}} \tag{1.122}$$

Assuming that intermediary i's family last updating their beliefs at time $t-k$, $\sigma^2_{i,t|t-1} \equiv \sigma^2_{i,t|t-k}$ can be written as:[30]

$$\sigma^2_{i,t|t-k} = \sum_{j=0}^{k-1}\tilde{\phi}^{2j}\text{var}^i_{t-k}(\varepsilon_{t-j}) \tag{1.123}$$

Putting Equations (1.122) and (1.123) together, I obtain the following expression for optimal updating probability:

$$\lambda_{i,t} = \frac{\frac{1}{\psi}(\Theta - 1)\frac{\tilde{\phi}^2\Theta(\Theta-1)}{1-\Theta\tilde{\phi}}\sum_{j=0}^{k-1}\tilde{\phi}^{2j}\text{var}^i_{t-k}(\varepsilon_{t-j})}{\xi - \frac{1}{\psi}\left(\frac{\tilde{\phi}^2\Theta(\Theta-1)}{1-\Theta\tilde{\phi}}\right)^2\sum_{j=0}^{k-1}\tilde{\phi}^{2j}\text{var}^i_{t-k}(\varepsilon_{t-j})} \tag{1.124}$$

The comparative statics of Equation (1.124) that are of particular interest to my analysis are:

- $\frac{\partial\lambda_{i,t}}{\partial k} > 0$: agents have an incentive to account for the progressive widening of their uncertainty during nonupdating periods
- $\frac{\partial\lambda_{i,t}}{\partial\text{var}^i_{t-k}(\varepsilon_{t-j})} > 0 \; \forall j < k$: updating beliefs becomes more valuable when the intermediary perceives the variance during the periods when expectations are not updated to be higher

In order to relate the comparative statics to the analysis in rest of the paper, consider the case of an intermediary whose family update their expectations at time $t-1$. If period $t-1$ is a recession period, $\text{var}^i_{t-1}(\varepsilon_t) = \pi_{hh}\sigma^2_h + (1-\pi_{hh})\sigma^2_l$ and if period $t-1$ is an expansion

[29]Numerical solution for $\kappa = 4$ (the parameterization in Opp, 2015) is presented in Appendix 1.D.2.

[30]I assume that agents do not observe realized volatility during the periods in which they fail to update their expectations.

period, $\text{var}^i_{t-1}(\varepsilon_t) = (1-\pi_{ll})\sigma_h^2 + \pi_{ll}\sigma_l^2$. Given the fact that $\pi_{hh} > 1 - \pi_{ll}$, $\lambda_{i,t}$ is higher conditional on the agent's family having updated their beliefs during a recession.

As the mass of agents who update their expectations during a recession increases, the average updating probability, $\bar{\lambda}_t$, increases.

1.C.4 Closing the Economy

In order to close the economy, I need the market clearing conditions for the domestic and foreign tradables markets. From the law of one price, we obtain the following:

$$p_{D,t} = p^*_{D,t} S_t \tag{1.125}$$

$$p_{F,t} = p^*_{F,t} S_t \tag{1.126}$$

Domestic households' demand for domestic tradables takes on the following form:

$$C_{D,t} = \frac{a}{p_{D,t}} \tag{1.127}$$

and foreign households' demand for domestic tradables is:

$$C^*_{D,t} = \frac{\zeta S_t}{p_{D,t}} \tag{1.128}$$

The world demand for domestic tradables is:

$$D_{D,t} = C_{D,t} + C^*_{D,t} = \frac{a + \zeta S_t}{p_{D,t}} \tag{1.129}$$

From the market clearing condition $D_{D,t} = Y_{D,t}$,

$$p_{D,t} = \frac{a + \zeta S_t}{Y_{D,t}} \tag{1.130}$$

Using the exact same logic,

$$p^*_{F,t} = \frac{a + \frac{\zeta}{S_t}}{Y_{F,t}} \tag{1.131}$$

1.D Additional Derivations and Results

1.D.1 Derivations

Ex-ante Expected Utility and Choice of Updating Probability

The problem solved by a generation t agent at the beginning of her life is:

$$\max_{\lambda_{i,t}} U^i_{t-1} \tag{1.D.132}$$

Given foreign bond demand in Equation (1.114), U^i_{t-1} can be rewritten as:

$$U^i_{t-1} = \mathbb{E}^i_{t-1}\left[\frac{1}{2\psi}\left(\mathbb{E}^i_t[r_{t+1}]\right)^2\right] - C(\lambda_{i,t}) \tag{1.D.133}$$

Using the definition of variance, the expression above is equivalent to:

$$\frac{1}{2\psi}\text{var}^i_{t-1}\left(\mathbb{E}^i_t[r_{t+1}]\right) + \frac{1}{2\psi}\left(\mathbb{E}^i_{t-1}\left[\mathbb{E}^i_t[r_{t+1}]\right]\right)^2 - C(\lambda_{i,t}) \tag{1.D.134}$$

The choice of $\lambda_{i,t}$ only affects expected utility through the variance and cost terms. The squared expectation term, which is not affected by the choice of $\lambda_{i,t}$ can be ignored for the purposes of the analysis in this section. In order to evaluate the variance term, I plug the definition of s_t and s_{t+1} into $\mathbb{E}^i_t\left[r_{t+1}\right]$:

$$\mathbb{E}^i_t[r_{t+1}] = \Theta i^D_t + \frac{\Theta}{1-\tilde{\phi}\Theta}\mathbb{E}^i_t[i^D_{t+1}] - \Theta\mathbb{E}^i_t[i^D_{t+1}] - \frac{\Theta\tilde{\phi}}{1-\Theta\tilde{\phi}}\mathbb{E}^i_t[i^D_{t+1}] - i^D_t \tag{1.D.135}$$

where I use the fact that $\mathbb{E}^i_t[i^D_{t+2}] = \tilde{\phi}\mathbb{E}^i_t[i^D_{t+1}]$.

Equation (1.D.135) can be rewritten as:

$$\mathbb{E}_t^i[r_{t+1}] = (\Theta - 1)\, i_t^D + \frac{\tilde{\phi}\Theta(\Theta-1)}{1-\Theta\tilde{\phi}}\mathbb{E}_t^i[i_{t+1}^D] \tag{1.D.136}$$

Given the chosen updating probability $\lambda_{i,t}$, the expression above is equivalent to:

$$(\Theta-1)i_t^D + \frac{\tilde{\phi}\Theta(\Theta-1)}{1-\Theta\tilde{\phi}}\left[\lambda_{i,t}\tilde{\phi}i_t^D + (1-\lambda_{i,t})\,\tilde{\phi}\mathbb{E}_{t-1}^i[i_t^D]\right] \tag{1.D.137}$$

$(1-\lambda_{i,t})\,\tilde{\phi}\mathbb{E}_{t-1}^i[i_t^D]$ does not contain variables that are not known to intermediary i at time $t-1$. Therefore, the variance of that term is zero.

Therefore, the agents' maximization problem takes on the following form:

$$\max_{\lambda_{i,t}}\left\{\frac{1}{2\psi}\mathrm{var}_{t-1}^i\left(\left((\Theta-1)+\frac{\tilde{\phi}^2\Theta(\Theta-1)}{1-\Theta\tilde{\phi}}\lambda_t^i\right)i_t^D\right) - C\left(\lambda_{i,t}\right)\right\} \tag{1.D.138}$$

I take the first-order condition with respect to $\lambda_{i,t}$, which leads to the expression in Equation (1.121).

UIP Coefficient, Intermediary Model

The subjective risk premium in the intermediary model is equal to:

$$\mathbb{E}_t^S[r_{t+1}] = (\Theta-1)i_t^D + \frac{\tilde{\phi}\Theta(\Theta-1)}{1-\Theta\tilde{\phi}}\mathbb{E}_t^S[i_{t+1}^D] \tag{1.D.139}$$

The deviation from rational expectations in Equation (1.18) is equivalent to:

$$\begin{aligned}&\mathbb{E}_t^P[s_{t+1}] - \mathbb{E}_t^S[s_{t+1}] = -\Theta\mathbb{E}_t^P[i_{t+1}^D] - \frac{\Theta}{1-\tilde{\phi}\Theta}\mathbb{E}_t^P\left[\mathbb{E}_{t+1}^S[i_{t+2}^D]\right] + \\ &\Theta\mathbb{E}_t^S[i_{t+1}^D] + \frac{\Theta\tilde{\phi}}{1-\tilde{\phi}\Theta}\mathbb{E}_t^S[i_{t+1}^D]\end{aligned} \tag{1.D.140}$$

Using the definitions of $\mathbb{E}^S$ and $\mathbb{E}^P$, Equation (1.D.140) can be rewritten as:

$$\left(\Theta + \frac{\Theta\tilde{\phi}\bar{\lambda}_s}{1-\Theta\tilde{\phi}}\right)\left(\mathbb{E}_t^S[i_{t+1}^D] - \mathbb{E}_t^P[i_{t+1}^D]\right) \tag{1.D.141}$$

Given the expressions above, the UIP coefficient can be written as:

$$\beta_s^F = \frac{\text{cov}_s\left(\frac{\Theta\tilde{\phi}}{1-\Theta\tilde{\phi}}(\Theta-1)\mathbb{E}_t^S[i_{t+1}^D] + (\Theta-1)i_t^D + \left(\Theta + \frac{\Theta\tilde{\phi}\bar{\lambda}_s}{1-\Theta\tilde{\phi}}\right)\left(\mathbb{E}_t^S[i_{t+1}^D] - \mathbb{E}_t^P[i_{t+1}^D]\right), i_t^D\right)}{\text{var}_s\left(i_t^D\right)} \tag{1.D.142}$$

The covaraince term can be written as:

$$\begin{aligned}&\left((\Theta-1) - \left(\Theta + \frac{\Theta\tilde{\phi}\bar{\lambda}_s}{1-\Theta\tilde{\phi}}\right)\phi\right)\text{var}_s(i_t^D) + \\ &+\left(\frac{\tilde{\phi}\Theta}{1-\Theta\tilde{\phi}}(\Theta-1) + \Theta + \frac{\Theta\tilde{\phi}\bar{\lambda}}{1-\Theta\tilde{\phi}}\right)\text{cov}_s\left(i_t^D, \mathbb{E}_t^S[i_{t+1}^D]\right)\end{aligned} \tag{1.D.143}$$

I plug the definition of $\text{cov}_s\left(i_t^D, \mathbb{E}_t^S[i_{t+1}^D]\right)$ into the equation above and obtain the following expression for the UIP coefficient:

$$\begin{aligned}\beta_s^F &= (\Theta-1) - \left(\Theta + \frac{\Theta\tilde{\phi}\bar{\lambda}_s}{1-\Theta\tilde{\phi}}\right)\phi + \\ &+\left(\frac{\tilde{\phi}\Theta}{1-\Theta\tilde{\phi}}(\Theta-1) + \Theta + \frac{\Theta\tilde{\phi}\bar{\lambda}}{1-\Theta\tilde{\phi}}\right)\left(\bar{\lambda}_s\tilde{\phi}\frac{1-\left((1-\bar{\lambda}_s)\tilde{\phi}\phi\right)^{m-1}}{1-(1-\bar{\lambda}_s)\tilde{\phi}\phi} + \left(1-\bar{\lambda}_s\right)^m\tilde{\phi}^{m+1}\phi^m\right)\end{aligned} \tag{1.D.144}$$

We obtain the expression in Equation (1.20) by setting $\psi = 0$ $(\Theta = 1)$.

1.D.2 Optimal Updating Probability, Curvature of Cost Function of 4

In order to solve for the optimal probability of updating, $\lambda_{i,t}$, I use the first-order condition from Equation (1.121)

$$\frac{1}{\psi}(\Theta-1)\frac{\tilde{\phi}^2\Theta(\Theta-1)}{1-\Theta\tilde{\phi}}\sigma_{i,t|t-1}^2 = \chi(\lambda_{i,t})^4 - \frac{1}{\psi}\left(\frac{\tilde{\phi}^2\Theta(\Theta-1)}{1-\Theta\tilde{\phi}}\right)^2\lambda_{i,t}\sigma_{i,t|t-1}^2 \tag{1.D.145}$$

and use a non-linear equation solver to solve the resultant equation.

The values I assign to the different parameters are reported in Table D1 below.

Table D1
Parameter Setup

ψ	0.16	χ	1×10^{-5}
ζ	0.4	κ	4
σ_l^2	1.35×10^{-5}	π_{ll}	0.97
σ_h^2	8.10×10^{-5}	π_{hh}	10/14

This table lists the base parameters that are used in the numerical exercise presented in this section.

The $\tilde{\phi}$ parameter is based on the calibration in the main body of the paper. The ψ and ζ parameters are chosen to allow the equilibrium model developed in this section to match the conditional dynamics of foreign exchange return predictability. The calibration of the κ parameter follows Opp (2015) and χ is chosen so that the unconditional mean of subjective variance, estimated using signal-to-noise ratios from Gourinchas and Tornell (2004), corresponds to the unconditional average updating intensity implied by the SPF. The π_{ll} and π_{hh} are calibrated based on the NBER recession dates. The volatility parameter are set so that the model matches the average values of realized variance and the average value of subjective variance implied by the signal-to-noise ratios estimated by Gourinchas and Tornell (2004).

Table D2
Variance and λ

k	var^i	λ_i	k	var^i	λ_i
0	1.55×10^{-5}	0.69	0	6.17×10^{-5}	0.98
1	3.02×10^{-5}	0.82	1	1.01×10^{-4}	1
2	4.37×10^{-5}	0.90	2	–	–
3	5.60×10^{-5}	0.95	3	–	–
4	6.70×10^{-5}	1	4	–	–

This table lists the perceived variance and the corresponding λ for k periods of inattention during expansions/recessions.

1.D.3 Speed of Convergence of the Probability of Updating Expectations

As discussed in Section 1.4, in the main body of this paper I use an expression of the following form to derive closed-form solutions:

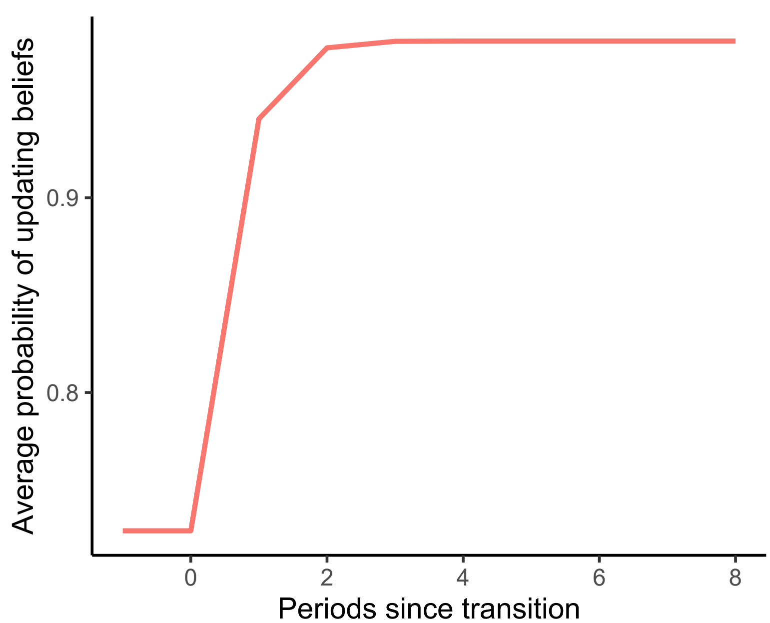
$$\mathbb{E}_t^S[i_{t+1}^D] = \bar{\lambda}_s \tilde{\phi} \sum_{j=0}^{m-1} \tilde{\phi}^{j+1}(1-\bar{\lambda}_s)^j i_{t-j}^D + \tilde{\phi}\left(1-\bar{\lambda}_s\right)^m i_{t-m}^D \tag{1.D.146}$$

where $\bar{\lambda}_s$ is the steady-state average updating probability is state s.

If we were to interpret this expression as an approximation, the accuracy of the approximation partly depends on how quickly $\bar{\lambda}$ converges to its steady-state value following a regime change.

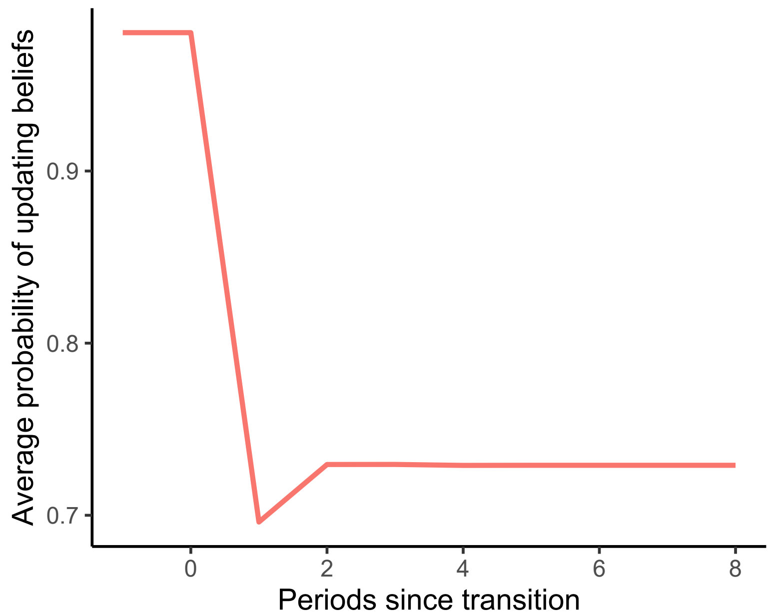
As shown in Table D2, the large values of $\lambda_{i,t}$ implied by the FX4Casts survey ensure that $\bar{\lambda}$ converges to its steady state value within a few periods. Figure D1 below depicts the evolution of $\bar{\lambda}$ during a transition from an expansion to a recession and Figure D2 depicts the evolution of $\bar{\lambda}$ during a transition from a recession to an expansion.

Figure D1. Average Probability of Updating Expectations, Transition into a Recession



This figure illustrates the relation between the number of periods since the transition into a recession state and the average probability of updating beliefs. The probabilities of updating beliefs are based on the figures in Table D2.

Figure D2. Average Probability of Updating Expectations, Transition into an Expansion



This figure illustrates the relation between the number of periods since the transition into an expansion state and the average probability of updating beliefs. The probabilities of updating beliefs are based on the figures in Table D2.

Given a transition at time $t = 0$, the value of $\bar{\lambda}$ becomes indistinguishable from its steady state value after two periods for recessions and after three periods for expansions. Additionally, the value of $\bar{\lambda}$ during the transition periods is reasonably close to its steady state value. The numerical errors associated with using the approximation appear to be minimal.

1.E Additional Robustness

1.E.1 Using Individual Currencies as Test Assets

A possible concern related to the results reported in Section 1.3 is the degree to which the results are driven by the fact that I pool ten currencies together and assume that the slope coefficient is the same for all of them. In order to verify the robustness of my results, I reestimate the regression in Equations (1.1) and (1.3) using nine individual currencies as test assets.[31]

First, I run the UIP regression for using the individual currencies in my sample. The results are reported in Table E1. Consistent with the results in the main body of the paper, the β^F coefficient is negative for all currencies and significantly different from zero for seven of the nine currencies in my sample.

The results related to the regression in Equation (1.3) are reported in Table E2. For parsimony, I only report results that use the CFNAI as a recession proxy. The results are qualitatively similar if I use detrended industrial production as a recession proxy.

The results in Table E2 show that the UIP coefficient for individual currencies mirrors the pattern observed in the pooled sample. The β^F_Δ coefficient is positive for all currencies and significant at conventional significance levels for five out of the nine currencies. $\beta^F_{\text{rec}} = \beta^F_{\text{exp}} + \beta^F_\Delta$ is positive for all nine currencies.

The currency-level evidence is consistent with the presence of significant differences between the UIP coefficients observed during expansions and recessions.

1.E.2 Selectively Omitting Recessions

It is informative to explore if the results reported in Table 1.1 are robust to omitting some of the recessions from my sample. In order to explore this question, I reestimate the regression in Equation (1.3) when omitting recessions from 1983-1989, 1990-1999, 2000-2009, 2010-

[31]The DKK is dropped from the analysis due to the fact that it has been pegged to the EUR since 1999.

Table E1

UIP Tests Using Individual Currencies

	Dependent variable:				
	r_{t+1}				
	(1)	(2)	(3)	(4)	(5)
	AUD	CAD	EUR	JPY	NOK
i_t^D	-1.046^{***}	-1.180^{**}	-0.864	-0.690^{**}	-1.200
	(0.277)	(0.558)	(1.046)	(0.273)	(0.658)
	[0.299]	[0.560]	[0.990]	[0.283]	[0.863]
Constant	-0.000	-0.000	0.002	0.001	-0.001
	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
	[0.002]	[0.001]	[0.002]	[0.001]	[0.002]
Observations	429	452	452	452	452
Adjusted R^2	0.024	0.002	0.000	0.017	0.006
		NZD	SEK	CHF	GBP
i_t^D		-1.397^{***}	-0.309	-2.735^{***}	-2.052^{*}
		(0.287)	(1.029)	(0.841)	(1.134)
		[0.276]	[0.813]	[0.903]	[1.015]
Constant		-0.001	0.000	0.005^{**}	-0.002
		(0.002)	(0.002)	(0.002)	(0.001)
		[0.003]	[0.002]	[0.002]	[0.002]
Observations		429	452	410	452
Adjusted R^2		0.060	0.000	0.025	0.013

This table reports the results for the UIP regression $r_{t+1} = \alpha + \beta^F i_t^D + \varepsilon_{t+1}$. In parentheses are standard errors computed following Newey and West (1987). In square brackets are bootsrapped standard errors. Significance at the 1%, 5%, and 10% is denoted by $*$, $**$, and $***$, respectively. The estimation is carried out using monthly data obtained from Barclays Bank International and Reuters and covers the period between 01/1983 and 09/2020.

Table E2

Conditional UIP Tests Using Individual Currencies

	Dependent variable:				
	r_{t+1}				
	(1)	(2)	(3)	(4)	(5)
	AUD	CAD	EUR	JPY	NOK
i_t^D	-1.109^{***}	-1.465^{**}	-1.505	-0.757^{***}	-1.681^{*}
	(0.290)	(0.593)	(1.060)	(0.263)	(0.671)
	[0.312]	[0.587]	[0.973]	[0.298]	[0.922]
$i_t^D \cdot \mathbf{1}_{\text{rec}\|\text{t}}$	1.276	2.495	10.480^{**}	0.958	14.616^{***}
	(1.293)	(3.058)	(4.766)	(0.333)	(5.427)
	[1.258]	[3.191]	[5.051]	[0.781]	[5.259]
Constant	-0.001	-0.000	0.003^{*}	0.001	-0.001
	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
	[0.002]	[0.001]	[0.002]	[0.001]	[0.002]
$\mathbf{1}_{\text{rec}\|\text{t}}$	0.007	0.001	-0.001	-0.001	0.036^{***}
	(0.012)	(0.009)	(0.007)	(0.003)	(0.013)
	[0.010]	[0.007]	[0.009]	[0.005]	[0.013]
Observations	429	452	452	452	452
Adjusted R^2	0.022	0.000	0.004	0.014	0.036
		NZD	SEK	CHF	GBP
i_t^D		-1.523^{***}	-0.580	-3.534^{***}	-3.111^{***}
		(0.295)	(1.191)	(0.873)	(1.221)
		[0.288]	[0.944]	[0.906]	[1.097]
$i_t^D \cdot \mathbf{1}_{\text{rec}\|\text{t}}$		1.581^{**}	4.837	11.001^{**}	7.023^{**}
		(0.766)	(3.261)	(3.820)	(3.351)
		[0.656]	[2.854]	[4.544]	[3.189]
Constant		-0.001	0.000	0.008^{***}	-0.002
		(0.002)	(0.002)	(0.002)	(0.001)
		[0.002]	[0.002]	[0.003]	[0.001]
$\mathbf{1}_{\text{rec}\|\text{t}}$		0.003	0.006	-0.013^{*}	0.007
		(0.009)	(0.011)	(0.005)	(0.010)
		[0.007]	[0.009]	[0.007]	[0.008]
Observations		429	452	410	452
Adjusted R^2		0.062	0.004	0.043	0.032

This table reports the results for the modified UIP regression $r_{t+1} = \alpha_{\exp} + \alpha_{\Delta}\mathbf{1}_{\text{rec}|\text{t}} + \left(\beta^F_{\exp} + \beta^F_{\Delta}\mathbf{1}_{\text{rec}|\text{t}}\right) i_t^D + \varepsilon_{t+1}$. The recession indicator is based on the Chicago Fed Activity Index (CFNAI). Recessions are associated with index values below -0.60. In parentheses are standard errors computed following Newey and West (1987). In square brackets are bootstrapped standard errors. Significance at the 1%, 5%, and 10% is denoted by $***$, $**$, and $*$, respectively. The estimation is carried out using monthly data obtained from Barclays Bank International and Reuters and covers the period between 01/1983 and 09/2020.

2020. The results are reported in Table E3. The results of tests that use the recession proxy constructed using detrended industrial production are reported.[32]

[32] I choose the industrial production proxy instead of the CFNAI as 24 out of the 39 months identified as recessions by the CFNAI fall between 2000 and 2009. This significantly decreases the power of the tests that omit the recessions between 2000 and 2009.

Table E3

Selectively Excluded Recessions

	Dependent variable:							
	r_{t+1}							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
i_t^D	-1.191^{***}	-1.260^{***}	-1.191^{***}	-1.296^{***}	-1.191^{***}	-1.295^{***}	-1.191^{***}	-1.298^{***}
	(0.252)	(0.287)	(0.250)	(0.290)	(0.250)	(0.290)	(0.250)	(0.290)
	[0.246]	[0.291]	[0.246]	[0.297]	[0.246]	[0.296]	[0.246]	[0.297]
$i_t^D \cdot \mathbf{1}_{\text{rec}\vert\text{t}}$	1.682^{*}	1.674^{*}	1.309^{**}	1.336^{**}	1.627^{**}	1.646^{**}	1.312^{**}	1.330^{**}
	(0.857)	(0.847)	(0.660)	(0.658)	(0.689)	(0.683)	(0.573)	(0.570)
	[0.891]	[0.879]	[0.481]	[0.471]	[0.458]	[0.453]	[0.476]	[0.463]
Constant	0.000		0.000		0.000		0.000	
	(0.001)		(0.001)		(0.001)		(0.001)	
	[0.001]		[0.001]		[0.001]		[0.001]	
$\mathbf{1}_{\text{rec}\vert\text{t}}$	0.001	0.001	0.005	0.005	0.005	0.005	0.001	0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)
	[0.004]	[0.005]	[0.004]	[0.004]	[0.005]	[0.005]	[0.003]	[0.004]
p-value ($H_0 : \beta_{\text{rec}}^F = 0$)	> 0.10	> 0.10	> 0.10	> 0.10	> 0.10	> 0.10	> 0.10	> 0.10
Observations	4,061	4,061	4,091	4,091	3,910	3,910	4,090	4,090
Omitted rec.	1983-1989		1990-1999		2000-2009		2010-2020	
Currency FE	No	Yes	No	Yes	No	Yes	No	Yes
Adjusted R^2	0.020	0.019	0.021	0.021	0.022	0.021	0.019	0.018

This table reports the results for the modified UIP regression $r_{j,t+1} = \alpha_{\exp} + \alpha_{\Delta}\mathbf{1}_{\text{rec}|\text{t}} + \left(\beta_{\exp}^F + \beta_{\Delta}^F \mathbf{1}_{\text{rec}|\text{t}}\right) i_{j,t}^D + \varepsilon_{j,t+1}$, where the recession indicator is an indicator constructed using detrended industrial production. The regression is estimated when omitting recession observations from 1983-1989, 1990-1999, 2000-2009, or 2010-2020. In parentheses are standard errors computed following Driscoll and Kraay (1998). In square brackets are bootstrapped standard errors. Significance at the 1%, 5%, and 10% is denoted by $***$, $**$, and $*$, respectively. The estimation is carried out using monthly data obtained from Barclays Bank International and Reuters. The full sample covers the period between 01/1983 and 09/2020. The panel consists of the exchange rate between the USD and the following 10 currencies: AUD, CAD, DKK, EUR, JPY, NOK, NZD, SEK, CHF, and GBP.

In general, the results in Table E3 seem to indicate that the results reported in Table 1.1 are not driven by a single recession: all β^F_Δ and β^F_{rec} coefficients are estimated to be positive and of similar magnitude to those in columns (3) and (4) of Table 1.1. Interestingly, the results become stronger if the recessions between 2000 and 2009 are dropped from the sample. Additionally, all β^F_Δ coefficients are estimated to be statistically significant at conventional levels (the p-values columns (1) and (2), which omit the recession between 1980 and 1989, are 5.9% and 5.6%, respectively).

1.E.3 Placebo Tests

As an additional robustness exercise, I provide the results of tests based on recession proxies that are not *ex-ante* expected to produce results similar to those reported in Table 1.1. In particular, I use the OECD recession dates for several small open economies (SOEs).[33] I opt for these proxies as local recessions within SOEs are unlikely to significantly impact the foreign exchange market at large. The existence of recessions specific to each country can be verified based on the relatively low correlations between the SOE recession proxies and the recession proxies I consider in Table 1.1. The results of the tests are reported in Table E4. None of the β^F_Δ coefficients reported in the table are significant. Additionally, all β^F_Δ coefficients are economically smaller than the coefficients in Table 1.1.

I also conduct placebo tests using three simulated random variables: placebo CFNAI, placebo industrial production, and placebo OECD recessions. These are binary variables that take on the value of 1 with a probability of 8.6%, 12.6%, and 50.1%, respectively (these correspond to the unconditional recession probabilities based on the CFNAI, detrended industrial production, and OECD recession proxies). These placebo tests yield β^F_{rec} coefficients greater than their empirical counterparts in Table 1.1 only 1.46%, 1.58%, and 1.2% of the time. The distributions associated with the simulated return predictability coefficients are shown in Figure D1.

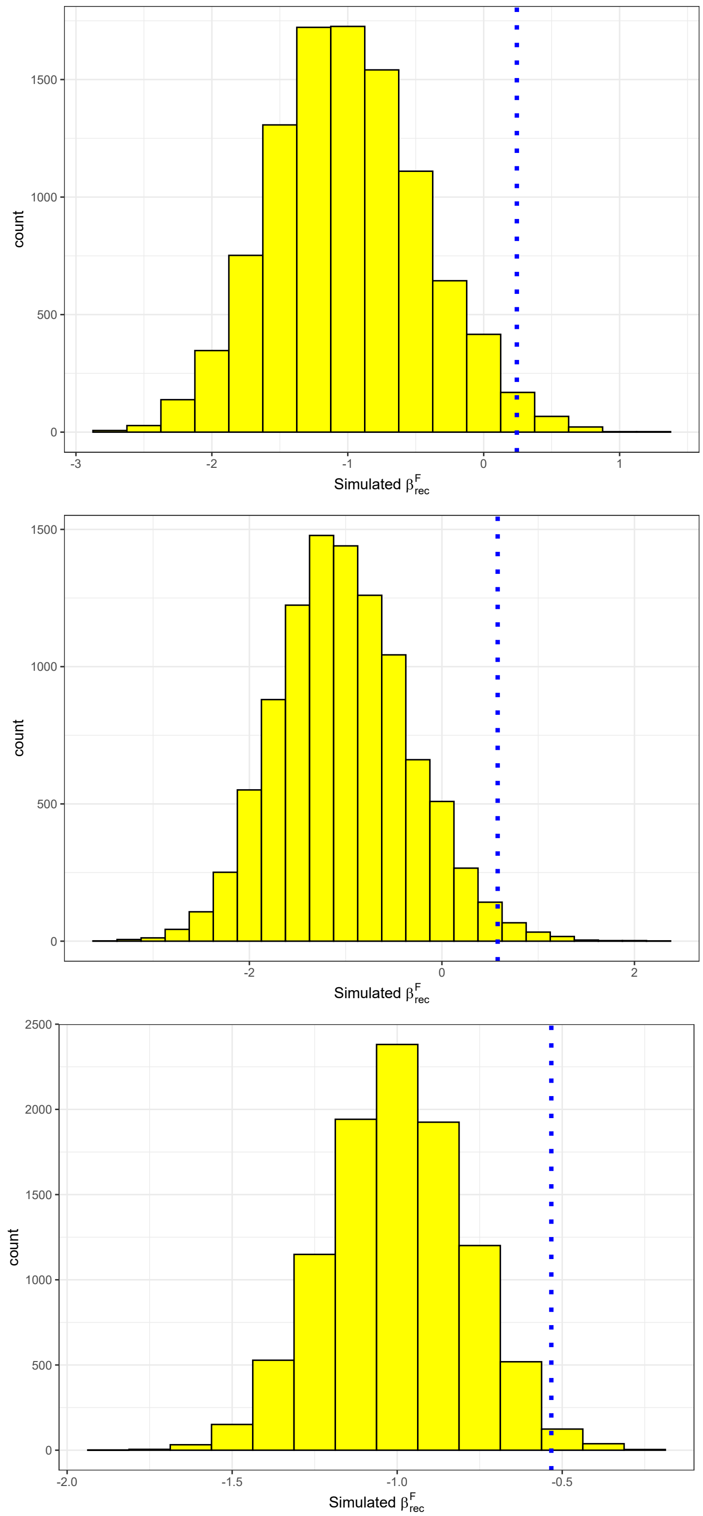
[33] The following SOEs are considered in this exercise: Denmark, Ireland, Korea, Turkey, and South Africa.

Table E4
Return Predictability Regressions, SOE Recessions

	Dependent variable:				
	r_{t+1}				
	(1)	(2)	(3)	(4)	(5)
i_t^D	-1.134^{***}	-1.043^{***}	-0.951^{***}	-1.232^{***}	-1.142^{***}
	(0.310)	(0.228)	(0.241)	(0.210)	(0.232)
	[0.351]	[0.217]	[0.214]	[0.165]	[0.269]
$i_t^D \cdot \mathbf{1}_{\text{rec}\mid\text{t}}$	0.317	0.046	-0.063	0.487	0.322
	(0.409)	(0.430)	(0.489)	(0.413)	(0.419)
	[0.435]	[0.380]	[0.414]	[0.407]	[0.469]
Constant	0.000	0.002	0.002	0.001	0.000
	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
$\mathbf{1}_{\text{rec}\mid\text{t}}$	0.000	-0.003	-0.004^{*}	0.000	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
	[0.002]	[0.002]	[0.002]	[0.003]	[0.002]
Observations	4,171	4,171	4,171	4,171	4,171
Rec. ind.	OECD DNK	OECD KOR	OECD TUR	OECD ZAF	OECD IRL
Currency FE	No	No	No	No	No
Cor with CFNAI	0.26	0.12	0.14	0.26	0.25
Cor with Ind. Prod.	0.20	0.02	0.08	0.33	0.32
Adj. R^2	0.016	0.017	0.019	0.016	0.016

This table reports the results for the modified UIP regression $r_{j,t+1} = \alpha_{\exp} + \alpha_{\Delta}\mathbf{1}_{\text{rec}\mid\text{t}} + \left(\beta^F_{\exp} + \beta^F_{\Delta}\mathbf{1}_{\text{rec}\mid\text{t}}\right) i^D_{j,t} + \varepsilon_{j,t+1}$, where the recession indicator is the OECD indicator for Denmark, Korea, Turkey, South Africa, and Ireland. In parentheses are standard errors computed following Driscoll and Kraay (1998). In square brackets are bootstrapped standard errors. Significance at the 1%, 5%, and 10% is denoted by $***$, $**$, and $*$, respectively. The estimation is carried out using monthly data obtained from Barclays Bank International and Reuters and covers the period between 01/1983 and 09/2020. The panel consists of the exchange rate between the USD and the following 10 currencies: AUD, CAD, DKK, EUR, JPY, NOK, NZD, SEK, CHF, and GBP.

Figure D1. Placebo Test: Simulated Return Predictability Coefficients



This figure illustrates the β^{F}_{rec} coefficients obtained using simulated recessions based on industrial production, the CFNAI, and the OECD global recession proxy. The dotted lines represent the empirical value of β^{F}_{rec}.

Taken together, the two sets of placebo tests provide further evidence that the results in Table 1.1 are unlikely to be obtained by chance.

1.E.4 Additional Return Predictors

A number of variables beyond interest rate differentials have been shown to predict foreign exchange returns. In this subsection, I consider the degree to which the results reported in the main body of the paper are robust to the inclusion of additional return predictors in the foreign exchange return predictability regressions. In order to do so, I estimate the following regressions:

$$r_{t+1} = \alpha + \beta^F i_t^D + \beta^q q_t + \varepsilon_{t+1} \tag{1.E.147}$$

and

$$r_{t+1} = \alpha_{\text{exp}} + \alpha_\Delta 1_{\text{rec}} + \left(\beta^F_{\text{exp}} + \beta^F_\Delta 1_{\text{rec}}\right) i_t^D + \left(\beta^q_{\text{exp}} + \beta^q_\Delta 1_{\text{rec}}\right) q_t + \varepsilon_{t+1} \tag{1.E.148}$$

where q_t is the additional return predictor.

I consider the following predictors: real exchange rates (Dahlquist and Pénasse, 2022), year-on-year and month-on-month inflation differentials (Engel et al., 2021), the US equity variance risk premium (Londono and Zhou, 2017), the VIX (Kalemli-Özcan and Varela, 2021), and the He et al. (2017) intermediary factor.

Dahlquist and Pénasse (2022) propose a present value model which implies that real exchange rates should predict currency returns. The real exchange rate is defined as: $s_t + p_t^* - p_t$ where p_t^* and p_t are log consumer price indeces obtained from the OECD. The price index data is available at the monthly frequency for all currencies except for the AUD and the NZD, for which data is available at the quarterly frequency. I forward fill the price indeces from the AUD and the NZD in the months until the next quarter. This approach avoids the avoids the use of information that is not available to the investors in real time at the cost of using stale information.

Engel et al. (2021) posit that UIP violations are driven by delayed reaction to monetary policy changes (the mechanism explored in this paper) and that year-on-year inflation captures central banks' monetary policy stance. Year-on-year inflation for the domestic and foreign countries is computed as $\pi_t = p_t - p_{t-12}$ and $\pi_t^* = p_t^* - p_{t-12}^*$, respectively. Here p_t and p_t^* are the consumer price indeces obtained from the OECD. The inflation differential is defined as $\pi_t - \pi_t^*$. As an additional robustness exercise, the results of tests using month-on-month inflation differentials as addtional predictors are also reported.

Londono and Zhou (2017) and Kalemli-Özcan and Varela (2021) argue that high values of the VRP (Londono and Zhou, 2017) or VIX (Kalemli-Özcan and Varela, 2021) indicate greater economic uncertainty and that investors demand higher returns as a compensation for bearing foreign exchange risk during periods of high uncertainty. I obtain the VRP data from Hao Zhou's website and the VIX data from the FRED.

Haddad and Muir (2021) rank foreign exchange as one of the most intermediated asset classes. Therefore, it would stand to reason that intermediaries are likely to act as the marginal investors in the foreign exchange market. This implies that the intermediaries' marginal value of wealth should be related to foreign exchange returns. He et al. (2017) argue theoretically that the intermediary capital ratio captures the intermediaries' marginal value of wealth. I download the intermediary capital ratio data and the HKM intermediary factor from Zhiguo He's website.[34]

The results of the predictability regressions are reported in Tables E5 and E6. For parsimony, only results based on the CFNAI recession proxy are shown. The results based on alternative recession proxies are qualitatively similar. Only the results of regressions with currency fixed effects are reported in order to ensure greater comparability with the original papers.

The results reported in Tables E5 and E6 indicate that the results reported in the main body of the paper are robust to the inclusion of additional variables in the predictability

[34]Unlike He et al. (2017), I do not find evidence that the inverse of the squared intermediary capital ratio predicts foreign exchange returns. However, I find that the HKM intermediary factor predicts returns. Therefore, I only report the results of test that include the HKM factor as a predictor.

Table E5

Return Predictability Regressions: Controlling for Existing Predictors

	Dependent variable:					
	r_{t+1}					
	RER		Y-O-Y inflation		M-O-M inflation	
	(1)	(2)	(3)	(4)	(5)	(6)
i_t^D	-1.173^{***}	-1.300^{***}	-0.882^{***}	-1.082^{***}	-1.093^{***}	-1.250^{***}
	(0.251)	(0.265)	(0.270)	(0.283)	(0.265)	(0.278)
	[0.250]	[0.268]	[0.262]	[0.252]	[0.249]	[0.266]
$i_t^D \cdot \mathbf{1}_{\text{rec}\vert\text{t}}$		1.623^{**}		2.212^{**}		1.977^{***}
		(0.717)		(1.019)		(0.761)
		[0.575]		[0.862]		[0.695]
q_t	-0.014^{***}	-0.014^{**}	-0.130^{***}	-0.098^{**}	-0.009	-0.011
	(0.006)	(0.007)	(0.048)	(0.046)	(0.007)	(0.007)
	[0.004]	[0.004]	[0.035]	[0.039]	[0.005]	[0.005]
$\mathbf{1}_{\text{rec}\vert\text{t}} \cdot q_t$		-0.000		-0.284^{*}		0.029
		(0.002)		(0.169)		(0.033)
		[0.000]		[0.158]		[0.024]
$\mathbf{1}_{\text{rec}\vert\text{t}}$		-0.002		-0.001		0.001
		(0.007)		(0.005)		(0.007)
		[0.006]		[0.006]		[0.006]
Observations	4,171	4,171	4,171	4,171	4,171	4,171
Sample	1983-2020	1983-2020	1983-2020	1983-2020	1983-2020	1983-2020
Currency FE	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.020	0.024	0.018	0.024	0.014	0.019

This table reports the results for the modified UIP regressions: $r_{j,t+1} = \alpha + \beta^F i_{j,t}^D + \beta^q q_{j,t} + \varepsilon_{j,t+1}$ and $r_{j,t+1} = \alpha_{\text{exp}} + \alpha_\Delta \mathbf{1}_{\text{rec}|\text{t}} + \left(\beta_{\text{exp}}^F + \beta_\Delta^F \mathbf{1}_{\text{rec}|\text{t}}\right) i_{j,t}^D + \left(\beta_{\text{exp}}^q + \beta_\Delta^q \mathbf{1}_{\text{rec}|\text{t}}\right) q_{j,t} + \varepsilon_{j,t+1}$, where the recession indicator is an indicator constructed using the Chicago Fed National Activity Index (CFNAI) and $q_{j,t}$ is an existing return predictor: real exchange rate (RER), the year-on-year inflation differential (Y-O-Y inflation), or the month-on-month inflation differential (M-O-M inflation). In parentheses are standard errors computed following Driscoll and Kraay (1998). Bootstrapped standard errors are in square brackets. Significance at the 1%, 5%, and 10% is denoted by $***$, $**$, and $*$, respectively. The estimation is carried out using monthly data obtained from Barclays Bank International and Reuters and covers the period indicated in the "Sample" row. The panel consists of the exchange rate between the USD and the following 10 currencies: AUD, CAD, DKK, EUR, JPY, NOK, NZD, SEK, CHF, and GBP.

Table E6

Return Predictability Regressions: Controlling for Existing Predictors, Continued

	Dependent variable:					
	r_{t+1}					
	VRP		VIX		HKM factor	
	(1)	(2)	(3)	(4)	(5)	(6)
i_t^D	-0.756^{**}	-0.941^{***}	-0.756^{**}	-0.967^{***}	-1.066^{***}	-1.194^{***}
	(0.311)	(0.337)	(0.312)	(0.332)	(0.258)	(0.268)
	[0.259]	[0.278]	[0.258]	[0.276]	[0.244]	[0.258]
$i_t^D \cdot \mathbf{1}_{\text{rec\|t}}$		2.042^{*}		2.000^{*}		1.652^{**}
		(1.045)		(1.039)		(0.804)
		[0.838]		[0.808]		[0.666]
q_t	0.404	1.141	0.019	0.029	-0.055^{***}	-0.066^{***}
	(0.678)	(1.114)	(0.021)	(0.021)	(0.020)	(0.018)
	[0.552]	[0.931]	[0.020]	[0.020]	[0.019]	[0.017]
$\mathbf{1}_{\text{rec\|t}} \cdot q_t$		−0.978		0.004		0.050
		(1.144)		(0.044)		(0.052)
		[1.076]		[0.051]		[0.054]
$\mathbf{1}_{\text{rec\|t}}$		0.002		−0.004		−0.002
		(0.006)		(0.011)		(0.006)
		[0.006]		[0.011]		[0.005]
Observations	3,377	3,377	3,377	3,377	4,171	4,171
Sample	1990-2020	1990-2020	1990-2020	1990-2020	1983-2020	1983-2020
Currency FE	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.005	0.010	0.005	0.012	0.026	0.031

This table reports the results for the modified UIP regressions: $r_{j,t+1} = \alpha + \beta^F i_{j,t}^D + \beta^q q_t + \varepsilon_{j,t+1}$ and $r_{j,t+1} = \alpha_{\text{exp}} + \alpha_\Delta \mathbf{1}_{\text{rec|t}} + \left(\beta_{\text{exp}}^F + \beta_\Delta^F \mathbf{1}_{\text{rec|t}}\right) i_{j,t}^D + \left(\beta_{\text{exp}}^q + \beta_\Delta^q \mathbf{1}_{\text{rec|t}}\right) q_t + \varepsilon_{j,t+1}$, where the recession indicator is an indicator constructed using the Chicago Fed National Activity Index (CFNAI) and q_t is an existing return predictor: variance risk premium (VRP), the VIX, or the He et al. (2017) intermediary factor (HKM). In parentheses are standard errors computed following Driscoll and Kraay (1998). Bootstrapped standard errors are in square brackets. Significance at the 1%, 5%, and 10% is denoted by $***$, $**$, and $*$, respectively. The estimation is carried out using monthly data obtained from Barclays Bank International and Reuters and covers the period indicated in the "Sample" row. The panel consists of the exchange rate between the USD and the following 10 currencies: AUD, CAD, DKK, EUR, JPY, NOK, NZD, SEK, CHF, and GBP.

regressions. In columns (1), (3), and (5) of the two tables the coefficient of the interest rate differential is negative and significant under all specifications considered in the two tables. This result is consistent with the finding that the additional predictors considered in the literature are unable to resolve the UIP puzzle.

The inclusion of the additional predictors has no marked impact on the main stylized fact documented in the first section of this paper: the β^F_Δ coefficient is positive and statistically different from zero (at the 6% level in columns (2) and (4) of Table E6) and β^F_{rec} is positive after the inclusion of the additional return predictors. In fact, the magnitude of the β^F_Δ coefficient is just as large (or larger in certain specifications) as the one reported in column (6) of Table 1.1. These results indicate that the return predictability patterns reported in the main body of the paper are not captured by any well-established return predictors.

1.E.5 Additional Tests of Forecast Rationality

1.E.5.1 SPF, Individual Forecast Horizons

In Table 1.5, I pool three forecast horizons together to increase the power of conditional tests. In this section, I estimate the forecast error predictability regressions for the three forecasting horizons separately. The results are reported in Table E7.

1.E.5.2 Mincer and Zarnowitz (1969) and Kohlhas and Walther (2021) Regressions

Kohlhas and Walther (2021) propose an alternative test of rational expectations: their methodology involves regressing time $t+h$ forecast errors on time t realizations:

$$i_{t+h} - i^D_{t+h|t} = \alpha^{KW} + \beta^{KW} i_t + \varepsilon_{t+h} \tag{1.E.149}$$

The β^{KW} coefficient is a strictly decreasing function of λ: more negative values of the coefficient are associated with less inattention.

Table E7

Forecast Error Predictability Regressions, Individual Forecast Horizons

	Dependent variable:		
	$FE_{t+h\|t}$		
	(1)	(2)	(3)
$FR_{t,h}$	0.408^{***}	0.506^{***}	0.689^{***}
	(0.085)	(0.088)	(0.119)
$FR_{t,1} \cdot \mathbf{1}_{\text{rec\|t}}$	-0.912^{***}	-0.715^{**}	-0.506
	(0.239)	(0.333)	(0.545)
Constant	-0.095^{***}	-0.212^{***}	-0.322^{***}
	(0.027)	(0.047)	(0.072)
$\mathbf{1}_{\text{rec\|t}}$	-0.895^{***}	-1.039^{***}	-1.049^{***}
	(0.219)	(0.131)	(0.273)
h	1	2	3
Observations	147	147	147
Rec. ind.	CFNAI	CFNAI	CFNAI
Adjusted R^2	0.244	0.162	0.141

This table reports the results for the forecast error-on-forecast revision regression $i_{t+h} - i_{t+h|t} = \alpha^{CG}_{\text{exp}} + \alpha^{CG}_{\Delta}\mathbf{1}_{\text{rec|t}} + \left(\beta^{CG}_{\text{exp}} + \beta^{CG}_{\Delta}\mathbf{1}_{\text{rec|t}}\right)\left(i_{t+h|t} - i_{t+h|t-1}\right) + \varepsilon_{t,t+h}$ for forecast horizons (h) of one, two, and three quarters, where the recession indicator is an indicator constructed using the Chicago Fed National Activity Index (CFNAI). In parentheses are standard errors computed following Newey and West (1987). Significance at the 1%, 5%, and 10% is denoted by $***$, $**$, and $*$, respectively. The estimation is carried out using quarterly SPF data obtained from the Philadelphia Fed and covers the period between 1983:Q1 and 2020:Q3.

I modify the regression in Equation (1.E.149) in a way that allows the regression coefficients to switch between recessions and expansions:

$$i^D_{t+h} - i^D_{t+h|t} = \alpha^{KW}_{\exp} + \alpha^{KW}_{\Delta} 1_{\mathrm{rec}|t} + \left(\beta^{KW}_{\exp} + \beta^{KW}_{\Delta} 1_{\mathrm{rec}|t}\right) i_t + \varepsilon_{t+h} \tag{1.E.150}$$

The Mincer and Zarnowitz (1969) tests of rational expectations involve examining the correlation between forecasts and interest rate realizations:

$$i_{t+h} = \alpha^{MZ} + \beta^{MZ} i_{t+h|t} + \varepsilon_t \tag{1.E.151}$$

In this framework, rational expectations correspond to a β^{MZ} of 1. A β^{MZ} bigger than 1 implies underreaction and β^{MZ} smaller than 1 implied overreaction.

I modify the regression in Equation (1.E.151) in a way that allows the regression coefficients to differ between recessions and expansions:

$$i_{t+h} = \alpha^{MZ}_{\exp} + \alpha^{MZ}_{\Delta} \mathbf{1}_{\mathrm{rec}|t} + \left(\beta^{MZ}_{\exp} + \beta^{MZ}_{\Delta}\right) i_{t+h|t} + \varepsilon_t \tag{1.E.152}$$

The results of the two sets of tests conducted using the SPF are reported in Table E8. The results of the Kohlhas and Walther (2021) tests are in the first four columns and the results of the Mincer and Zarnowitz (1969) are in columns (5) through (8). The results of the conditional tests reported in the table are qualitatively consistent with the results in Table 1.5. The point estimates of the coefficients of the interaction term are negative in all six specifications. Negative β_Δ indicates that the stickiness of the interest rate expectations declines during recessions. However, the results are somewhat weaker quantitatively: the coefficients are only significant when we utilize the industrial production indicator as a recession proxy.

Table E8

Alternative Tests of Rational Expectations, SPF

	Dependent variable:							
	$FE_{t+h\|t}$				i_t			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
i_t	-0.030	-0.035	-0.034	-0.021				
	(0.033)	(0.035)	(0.033)	(0.056)				
$i_t \cdot \mathbf{1}_{\text{rec}\|\text{t}}$		-0.144^{*}	-0.059	-0.030				
		(0.081)	(0.090)	(0.064)				
$i_{t\|t-h}$					0.964^{***}	0.959^{***}	0.960^{***}	0.972^{***}
					(0.034)	(0.035)	(0.033)	(0.056)
$i_{t\|t-h} \cdot \mathbf{1}_{\text{rec}\|\text{t}}$						-0.167^{**}	-0.079	-0.032
						(0.081)	(0.099)	(0.063)
Constant	-0.227^{**}	-0.104	-0.152	0.015	-0.198	-0.073	-0.120	0.049
	(0.108)	(0.113)	(0.101)	(0.158)	(0.112)	(0.114)	(0.101)	(0.167)
$\mathbf{1}_{\text{rec}\|\text{t}}$		-0.344^{**}	-0.670^{**}	-0.446^{**}		-0.262	-0.612^{**}	-0.439^{**}
		(0.162)	(0.258)	(0.193)		(0.172)	(0.287)	(0.200)
Observations	441	441	441	441	441	441	441	441
Rec. ind.	None	Ind. prod.	CFNAI	OECD+NM	None	Ind. prod.	CFNAI	OECD+NM
Adjusted R^2	0.009	0.126	0.086	0.134	0.918	0.929	0.925	0.929

This table reports the results for the forecast error-on-past realization regression $i_{t+h} - i_{t+h|t} = \alpha_{\text{exp}}^{KW} + \alpha_{\Delta}^{KW} \mathbf{1}_{\text{rec}|\text{t}} + \left(\beta_{\text{exp}}^{KW} + \beta_{\Delta}^{KW} \mathbf{1}_{\text{rec}|\text{t}}\right) i_t + \varepsilon_{t,t+h}$ and the realization-on-forecast regression $i_{t+h} = \alpha_{\text{exp}}^{MZ} + \alpha_{\Delta}^{MZ} \mathbf{1}_{\text{rec}|\text{t}} + \left(\beta_{\text{exp}}^{MZ} + \beta_{\Delta}^{MZ} \mathbf{1}_{\text{rec}|t}\right) i_{t+h|t} + \varepsilon_t$ for a pooled panel of $h = \{1, 2, 3\}$, where the recession indicator is an indicator constructed using detrended industrial production data, the Chicago Fed National Activity Index (CFNAI), or the OECD and Non-member recession indicator (OECD+NM). In parentheses are standard errors computed following Driscoll and Kraay (1998). Significance at the 1%, 5%, and 10% is denoted by $***$, $**$, and $*$, respectively. The estimation is carried out using quarterly data from Philadelphia Fed and covers the period between 1983:Q1 and 2020:Q3.

CHAPTER 2

Earnings Expectations and Asset Prices

With Gabriel Cuevas Rodriguez *(UCLA Anderson)* **and Danyu Zhang** *(UCLA Anderson)*

2.1 Introduction

Bouchaud, Kruger, Landier, and Thesmar (2019) use the methodology developed by Coibion and Gorodnichenko (2015) to examine the rationality of financial analysts' earnings forecasts and find that analysts' forecasts underreact to news unconditionally. A related finding in the literature, most notably in Coibion and Gorodnichenko (2015), is that there is state dependence in the expectation formation process. In this paper, we examine the degree to which the state dependence of information rigidity extends to equity analysts' earnings expectations. We use financial analysts' earnings per share (EPS) forecasts from the Thomson Reuters Institutional Brokers Estimate System (I/B/E/S) and document the following stylized facts: (1) unconditionally, earnings expectations underreact to news, in line with the findings of Abarbanell and Bernard (1992) and Bouchaud et al. (2019); (2) the stickiness of earnings expectations declines significantly during periods of high market volatility; and (3) the stickiness of earnings expectations declines significantly over our sample period (01/1986 to 12/2021).[1]

[1]Throughout this paper, we use the terms information rigidity and expectation stickiness interchangeably as the two terms are equivalent within our theoretical framework.

We develop a simple model featuring rational inattention, as in Mankiw and Reis (2002), to explain the state dependence in equity analysts' expectation formation process. In our model, a fraction of the market participants optimally remain inattentive to news. The presence of inattentive agents causes the average (consensus) earnings expectations to underreact to earnings shocks. The time-varying costs and benefits of paying attention to earnings news drive the state-dependence of expectation stickiness. The benefits of acquiring information increase during high-volatility periods, which translates into lower expectation stickiness during high-volatility periods. Our model is also able to account for expectation stickiness declining over our sample, conditional on technological advancements reducing the cost of information acquisition over time.

Given the ability of our model to replicate the conditional behavior of equity analysts' earnings expectations, we focus on its asset pricing implications. The momentum anomaly (Jegadeesh and Titman, 1993) provides a natural starting point for our analysis. Momentum is one of the most robust empirical results in financial economics. Consequently, it has garnered significant attention in the asset pricing literature. Our framework fits into the class of models that relate the profitability of momentum to the slow diffusion of information (Hong and Stein, 1999): the presence of inattentive agents causes prices to incorporate publicly available information gradually. That is, stock prices underreact news, so that prices are, on average, too low following positive news and too high following negative news. As a consequence, positive lagged returns predict high subsequent returns and negative lagged returns predict low subsequent returns, thus generating return patterns consistent with the unconditional profitability of momentum.

More recently, Daniel and Moskowitz (2016) find that momentum experiences negative returns (crashes) following high-volatility periods and Chordia, Subrahmanyam, and Tong (2014) find that the profitability of momentum has diminished significantly over the period between 1976 and 2011. First, we verify the robustness of these results in our extended sample. Then, we show that our model is capable of accounting for these results. According to our model, the profitability of momentum is related to the stickiness of market participants' expectations. The increasing relative cost of being inattentive during high-volatility periods accounts for momentum crashes and the declining cost of information acquisition accounts for the attenuating profitability of momentum over our sample.

A prediction of our model is that the relative profitability of momentum strategies with different lookback periods differs based on the level of market volatility. During low-volatility periods, momentum strategies with longer lookback periods (e.g. the Jegadeesh and Titman, 1993 $t-12$ to $t-2$ strategy) tend to outperform momentum strategies with shorter lookback periods (e.g. a $t-3$ to $t-2$ momentum strategy). However, our model predicts that short-run momentum strategies may deliver higher returns than long-run momentum strategies during high-volatility periods. Based on this model prediction, we propose a trading strategy that mixes long-run and short-run momentum signals (similar to Goulding, Harvey, and Mazzoleni, 2022), with greater weight placed on the short-run signal during high-volatility periods. The resultant mixed momentum strategy lessens the impact of momentum crashes and earns a significant α with respect to the baseline 12-2 momentum strategy.

A prediction that our model shares with any model featuring deviations from full-information rational expectations is that the wedge between the objective earnings expectations and analysts' forecasts predicts stock returns. In order to test this model prediction, we use the Extreme Gradient Boosting machine learning algorithm to construct a proxy for objective expectations (similar to Van Binsbergen, Han, and Lopez-Lira, 2022). Our approach serves as an extension of papers in the literature that use linear regression frameworks to extract the predictable component of analysts' forecast errors (e.g. So, 2013 and Frankel and Lee, 1998).

Armed with a measure of objective expectations, we sort stocks into portfolios based on the value of the predictable component of forecast errors. We call the long-short portfolio that is long on the stocks with the most pessimistic earnings expectations and short on the stocks with the most optimistic earnings expectations pessimistic-minus-optimistic (PMO). The PMO strategy generates an annualized Sharpe ratio of 1.16 and its returns cannot be fully explained by standard multifactor models.

Related literature

Our paper contributes to a number of different strands of the literature. In terms of documenting state-dependent expectation stickiness, our paper is most closely related to Coibion and Gorodnichenko (2015). In other related work, Loungani et al. (2013) show that professional forecasters increase the rate at which they incorporate news into their forecasts as the economy enters a recession and Andrade and Le Bihan (2013) document that the Great Recession is associated with increased attentiveness to unemployment, real GDP, and inflation news among professional forecasters surveyed by the European Central Bank.

Our work is also related to the literature that documents systematic errors in equity analysts' earnings expectations and relates the systematic errors to the profitability of various trading strategies. For instance, Bouchaud et al. (2019) show that analysts' short-run earnings expectations underreact to news and propose an explanation for the profitability anomaly (Novy-Marx, 2013) based on the underreaction to earnings surprises. On the other hand, Bordalo, Gennaioli, La Porta, and Shleifer (2022) show that analysts' long-term earnings growth expectations overreact to news and develop a model in which the profitability of the Fama and French (2015) factors is driven by the overreaction of long-term expectations. Engelberg, McLean, and Pontiff (2018) use *ex-post* forecast errors and show that analysts tend to have overly optimistic (pessimistic) expectations for stocks in the short (long) leg of various anomalies.

Our framework is based on the idea that costly information acquisition causes earnings shocks to be incorporated into consensus expectations slowly. The slow diffusion of information, in turn, causes momentum. It is well-established in the asset pricing literature that limited investor attention is associated with slow diffusion of information and underreaction to news. For instance, Ben-Rephael, Da, and Israelsen (2017) provide a measure of abnormal institutional attention and show that the post-earnings announcement drift is driven by announcements which do not receive sufficient attention from institutional investors. In related work, Hirshleifer, Lim, and Teoh (2009) show that the underreaction to earnings announcements is stronger on days with more earnings announcements and Dellavigna and

Pollet (2009) show that the underreaction is more pronounced for earnings announcements that take place on Friday. Two papers that relate the speed with which information is incorporated into aggregate expectations to the profitability of momentum are Hong, Lim, and Stein (2000) and Da, Gurun, and Warachka (2014). The two papers propose different proxies for the speed at which news is incorporated into consensus expectations, residual analyst coverage and information discreteness, respectively and find that stocks for which information is incorporated into expectations more slowly deliver higher momentum returns.

In terms of the use of non-linear methods to construct *ex-ante* forecast errors, our paper is most closely related to Van Binsbergen et al. (2022), de Silva and Thesmar (2023), and Cao and You (2021). Van Binsbergen et al. (2022) find that stocks with upward- (downward-) biased earnings forecasts tend to earn lower (higher) returns going forward. de Silva and Thesmar (2023) decompose analysts' forecast errors at different horizons into soft information, forecast bias, and forecast noise. Cao and You (2021) document that earnings information uncovered by machine learning algorithms (over extant models) is significantly associated with future stock returns and earnings forecast errors.

2.2 Data

2.2.1 Analysts' forecasts

We obtain consensus (median) earnings-per-share (EPS) forecasts from the I/B/E/S Unadjusted Summary file. Following Bouchaud et al. (2019), we focus on the one-year and two-year earnings forecasts.[2] I/B/E/S updates earnings forecasts monthly. Our tests of forecast error predictability are based on the forecasts immediately following the announcement of the previous fiscal year's earnings.

We match earnings forecasts with earnings realizations from the I/B/E/S actual file using ticker and fiscal end date.[3] Before merging the two datasets, we adjust the realized EPS

[2]The forecasting horizon is identified using the I/B/E/S Forecast Period Indicator variable FPI.

[3]Fiscal end dates are denoted by PENDS in the actual file and by FPEDATS in the summary file.

values for stock splits using the CRSP cumulative adjustment factor CFACSHR (Diether, Malloy, and Scherbina, 2002):

$$AdjustedEPS_{f,t} = \frac{CFACSHR_{f,t-1}}{CFACSHR_{f,t}} \times EPS_{f,t}$$

The forecast error predictability tests are based on firms with fiscal year ends between 01/1986 and 12/2021. Our final dataset contains 78,287 firm-year observations.

2.2.2 Stock and trading strategy returns

We obtain monthly return and stock price data from CRSP. We start with all firms in the monthly CRSP database between 1986 and 2021 and apply the following filters: we only keep the common stocks (share codes 10 and 11) of firms listed on the NYSE, Amex, or Nasdaq (exchange code 1, 2, and 3). We also exclude firms whose stock price is below $1. We then match the CRSP data with the analyst forecast data described in the previous section.[4]

In this paper, we also utilize a number of off-the-shelf trading strategies and factor returns, which serve as control variables or building blocks for the strategies proposed in this paper. We make the decision to use data used in previous research to ensure greater comparability with existing work. We obtain the returns of the trading strategies from one of two sources: Kenneth French's data library or the Global Factor Data repository (Jensen, Kelly, and Pedersen, 2022).

2.3 Forecast Error Predictability

We start our analysis by examining the ability of earnings forecast revisions to predict earnings forecast errors. The resultant regression coefficient allows us to draw conclusions regarding the degree of information rigidity in equity analysts' expectations (Coibion and Gorodnichenko, 2015).

[4] We merge I/B/E/S data with CRSP data using the link table provided by Wharton Research Data Services.

To examine the ability of forecast revisions to predict forecast errors, we first construct forecast revisions. Forecast revisions are defined as the difference between the time t consensus forecast for firm f's fiscal year τ earnings (one-year forecast) and the time $t-1$ consensus forecast for firm f's fiscal year τ earnings (two-year forecast). Throughout our analysis, we use τ to denote fiscal years and t to denote calendar time.

Following Bouchaud et al. (2019), we normalize the revision by firm f's stock price in year $t-1$, $P_{f,t-1}$.[5] Therefore,

$$FR_{f,t} = \frac{\mathbb{F}_t\left[e_{f,\tau}\right] - \mathbb{F}_{t-1}\left[e_{f,\tau}\right]}{P_{f,t-1}}$$

where e denotes earnings per share.

Forecast errors are defined as the difference between the actual fiscal year τ earnings and the year t earnings forecast, normalized by the time $t-1$ stock price:

$$FE_{f,\tau} = \frac{e_{f,\tau} - \mathbb{F}_t\left[e_{f,\tau}\right]}{P_{f,t-1}}$$

We begin our analysis by estimating the following regression, following Bouchaud et al. (2019):

$$FE_{f,\tau} = \alpha^{CG} + \beta^{CG} FR_{f,t} + \varepsilon_{t,\tau} \tag{2.1}$$

where we winsorize the forecast errors and forecast revisions at the 1% and 99% levels.

If the analysts' information set includes all information available at time t, we would not be able to predict forecast errors using any time t information, including forecast revisions, i.e. the regression coefficient, β^{CG}, is equal to zero under the null of rational expectations. If analysts' expectations underreact to earnings shocks, we would observe $\hat{\beta}^{CG} > 0$. The mechanism underlying this result is the following: let us assume that agents receive a piece of positive news at time t. This implies that $FR_{f,t} > 0$. However, if agents underreact to

[5] $P_{f,t-1}$ is the stock price at the end of the month used to determine $\mathbb{F}_{t-1}\left[e_{f,\tau}\right]$.

the news, the forecast errors will also be positive, on average, i.e. forecast revisions will be positively correlated with forecast errors.

Conversely, overreaction to news implies $\hat{\beta}^{CG} < 0$. The mechanism underlying this result is the same as the one outlined in the underreaction case. The difference is that the forecast errors will be negative, on average, following positive news. That is, forecast errors are negatively correlated with forecast revisions in the case of overreaction.

The results in the main body of this paper are based on a panel regression of all firm-level observations.The results of the baseline regression are reported in the first column of Table 1. We estimate the β^{CG} coefficient to be positive ($\hat{\beta}^{CG} = 0.167$) and highly statistically significant (t-statistic $= 5.195$). The positive regression coefficient indicates that equity analysts' expectations underreact to earnings shocks. This result is consistent with the findings in Bouchaud et al. (2019) who estimate β^{CG} to be 0.165.

Volatility and information rigidity

Coibion and Gorodnichenko (2015) show that the rigidity of SPF survey respondents' expectations declines during NBER recessions. They hypothesize that the decline in information rigidity is driven by the high macroeconomic volatility prevailing during recessions. In order to test the validity of this hypothesis in the context of earnings expectations we examine the relation between information rigidity and stock market volatility. In particular, we estimate the following regression:

$$FE_{f,\tau} = \alpha^{CG} + \alpha_{\Delta}^{CG} z_t + \left(\beta^{CG} + \beta_{\Delta}^{CG} z_t\right) FR_{f,t} + \varepsilon_{t,\tau} \tag{2.2}$$

where z_t is either time t stock market volatility ($\hat{\sigma}_{\text{mkt}}$) or an indicator, which takes on the value of 1 if stock market volatility at time t is above a certain threshold ($\mathbf{1}_{\text{HV}}$). In our analysis, z_t is computed by taking the average of beginning-of-month and end-of-month volatility for month t. Alternative methods of computing volatility produce qualitatively similar results.

Table 1
Forecast Error Predictability Regressions

	Dependent variable:						
	$FE_{f,\tau}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		RV			IV		
$FR_{f,t}$	0.167***	0.243***	0.201***	0.197***	0.285***	0.195***	0.199***
	(5.195)	(5.497)	(7.230)	(7.546)	(4.817)	(7.280)	(7.722)
$FR_{f,t} \cdot \hat{\sigma}_{\text{mkt}}$		−0.065			−0.056*		
		(−1.541)			(−1.848)		
$FR_{f,t} \cdot \mathbf{1}_{\text{HV}}$			−0.126*	−0.160*		−0.128*	−0.190**
			(−1.900)	(−1.848)		(−1.984)	(−2.288)
Constant	−0.012***	−0.013***	−0.011***	−0.012***	−0.007**	−0.011***	−0.011***
	(−8.190)	(−5.942)	(−8.438)	(−8.782)	(−2.603)	(−7.739)	(−8.436)
$\hat{\sigma}_{\text{mkt}}$		0.001			−0.002		
		(0.502)			(−1.404)		
$\mathbf{1}_{\text{HV}}$			−0.001	0.001		−0.005	−0.002
			(−0.236)	(0.340)		(−1.634)	(−0.420)
Observations	78,287	78,287	78,287	78,287	77,969	77,969	77,969
$P(\mathbf{1}_{\text{HV}} = 1)$	–	–	0.20	0.10	–	0.20	0.10
Adjusted R^2	0.014	0.015	0.015	0.016	0.015	0.015	0.016

This table reports the results for the forecast error predictability regression $FE_{f,\tau} = \alpha^{CG} + \beta^{CG} FR_{f,t} + \varepsilon_{t,\tau}$ and the modified regression $FE_{f,\tau} = \alpha^{CG} + \alpha^{CG}_{\Delta} z_t + \left(\beta^{CG} + \beta^{CG}_{\Delta} z_t\right) FR_{f,t} + \varepsilon_{t,\tau}$, where z_t in columns (2) and (5) is realized and implied volatility, respectively, and a high volatility indicator in columns (3), (4), (6), and (7). Realized volatility is based on daily value-weighted CRSP returns for a period of 126 days. The CBOE S&P100 Volatility Index is used as a measure of implied volatility. $FE_{f,\tau}$ is defined as $\left(e_{f,\tau} - \mathbb{F}_t\left[e_{f,\tau}\right]\right)/P_{f,t-1}$ and $FR_{f,t}$ is defined as $\left(\mathbb{F}_t\left[e_{f,\tau}\right] - \mathbb{F}_{t-1}\left[e_{f,\tau}\right]\right)/P_{f,t-1}$ Standard errors are double-clustered by firm and year. The corresponding t-statistics are in square brackets. Significance at the 1%, 5%, and 10% is denoted by $***$, $**$, and $*$, respectively. Earnings forecasts are obtained from I/B/E/S and cover firms with fiscal year ends between 1/1986 and 11/2021. The volatility measures are in percent per month.

The results from the estimation of the regression in Equation (2.2) are reported in columns (2) through (7) of Table 1. In this table, we consider two measures of volatility: realized volatility (RV) and implied volatility (IV). RV is the standard deviation of the daily market returns over the 126 days prior to time t (Daniel and Moskowitz, 2016):

$$\hat{\sigma}_{mkt,t} = \sqrt{\sum_{d=1/126}^{1} \left(R_{mkt,t+d} - \frac{\sum_{d=1/126}^{1} R_{mkt,t+d}}{126} \right)^2} \tag{2.3}$$

where $R_{mkt,t}$ represents the day t return of value-weighted CRSP.

We use the VXO (implied volatility of S&P 100 index options) as a measure of IV. The "high volatility" indicators ($\mathbf{1}_{\text{HV}}$) considered in the table take on the value of 1 if volatility is above the 80^{th} or 90^{th} percentile of the volatility estimates within our full sample.

The results in Table 1, show that information rigidity declines significantly during high-volatility periods: we estimate β_{Δ}^{CG} coefficients to be negative and statistically significant. The two volatility measures considered in the table generate similar results. We estimate $\hat{\beta}_{\Delta}^{CG}$ to be -0.056 using implied volatility. The effect of using realized volatility is slightly weaker compared to implied volatility.

Information rigidity over time

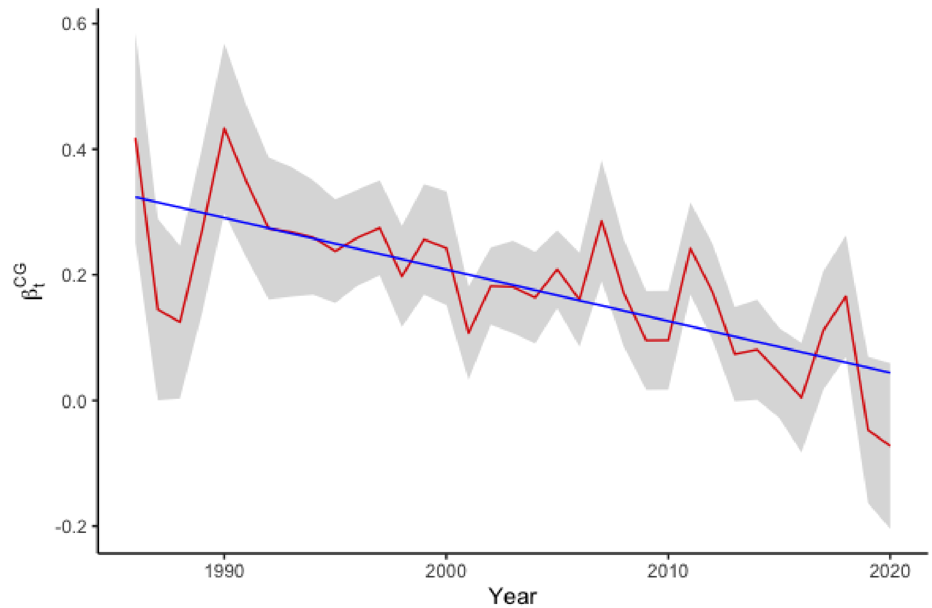
Coibion and Gorodnichenko (2015) document a low-frequency variation in the stickiness of macroeconomic expectations. In particular, they find that the rigidity of SPF respondents' expectations increased significantly during the period characterized by low macroeconomic volatility known as the Great Moderation.

In this section, we examine if equity analysts' earnings expectations display similar low-frequency patterns. To do so, we estimate the regression in (2.1) using overlapping two-year windows.[6] Figure 1 plots the dynamics of the β_t^{CG} coefficients obtained using this methodology as well as the associated 95% confidence intervals and a fitted trend line.

[6] For instance, the $\hat{\beta}^{CG}$ coefficient associated with 1986 is based on a sample that includes 1986 and 1987.

Figure 1 indicates that the rigidity of earnings expectations declines significantly throughout the sample. In fact, we cannot reject rational expectations during the post-2010 sample ($\hat{\beta}^{CG} = 0.077$, t-statistic $= 1.616$).

Figure 1. Information Rigidity Over Time



The figure above depicts the coefficient β_t^{CG} on forecast revisions in specification (2.1) estimated using overlapping two-year windows. The shaded region represents the 95% confidence interval for $\hat{\beta}_t^{CG}$ and the blue line is the fitted trend line. The estimation is carried out using data from I/B/E/S and covers firms with fiscal year ends between 01/1986 and 12/2021.

The lack of stickiness in analysts' earnings expectations during the latter parts of our sample is consistent with the findings of Martineau (2023) who shows that the post-earnings announcement drift has disappeared in 2006 for large stocks and in the 2010s for microcaps.

2.4 Model

In Section 2.3, we identify the following patterns in equity analysts' consensus earnings expectations:

1. Equity analysts' earnings expectations underreact to shocks unconditionally.

2. The degree of underreaction to earnings shocks declines significantly during high volatility periods.

3. The degree of underreaction to earnings shocks declines over our sample.

In this section, we develop a simple model capable of accounting for the three stylized facts presented above.

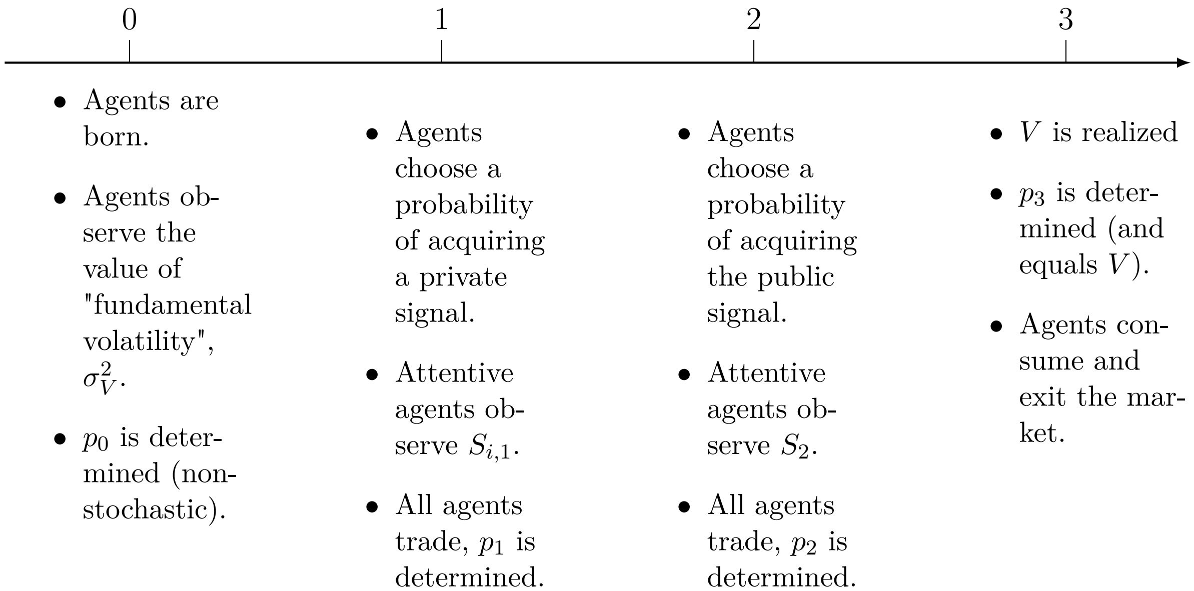
The economy we study is based on the one in Pouget, Sauvagnat, and Villeneuve (2016). It consists of two assets: a riskless asset in perfectly elastic supply and a risky asset in fixed supply, x. The gross return of the riskless asset is normalized to 1. There are $T-1$ dates of trading, indexed by $t \in \{0, 1, ..., T-1\}$. Consumption takes place at date T during which the payoff of the risky asset is drawn from a normal distribution with a mean of 0 and variance of σ_V^2. There is no consumption prior to the final period and there is no time discounting.

We set $T = 3$, which allows us to derive closed-form solutions for moments of interest. The signal structure we consider follows Daniel, Hirshleifer, and Subrahmanyam (1998). At time $t = 1$, overconfident agent i (in measure one) obtains a private signal $S_{i,1}$ with probability λ_1 that is determined endogenously. Fraction $1-\lambda_1$ of the market participants fail to obtain the private signal. The inattentive agents use their prior beliefs to form expectations about the future, as in Mankiw and Reis (2002). Following the standard approach in the literature, we assume that $S_{i,1} = V + \varepsilon_{i,1}$. Where the noise term, $\varepsilon_{i,1}$, is independently distributed across agents and normally distributed with a mean of 0 and a variance of σ_S^2. Following Daniel et al. (1998), we model overconfidence as agents overassessing the quality of their private signals, i.e. agents perceive the variance of their private signal to be $\sigma_C^2 < \sigma_S^2$.[7]

At time $t = 2$, a public signal S_2 is realized. Agent i observes the signal and updates her expectations with probability $\lambda_{i,2}$, which depends on the agent's information set at the beginning of period 2. The public signal is equal to $V + \varepsilon_2$ where $\varepsilon_2 \sim N(0, \sigma_S^2)$. We assume that agents perceive the precision of the public signal correctly.

[7] Overconfidence has limited bearing on the ability of our model to replicate the patterns in equity analysts' earnings expectations. Overconfidence plays an important role in Section 2.5, in which we study the asset pricing implication of our model.

Figure 2. Model Representation



This figure summarizes the life cycle of an agent born in period 0.

Figure 2 provides a graphical illustration of our model. To relate our model to the empirical results in Section 2.3, we interpret the private signals as a representation of the information collected by equity analysts prior to an earnings release. The public signal and the final asset payoff represent earnings announcements.

In order to focus on the expectation formation process presented in this paper, we make several assumptions that follow Pouget et al. (2016). First, we assume that the market participants are risk-neutral and incur an exogenous trading cost that is quadratic in their portfolio positions, i.e. the total cost for trader i is $\frac{\psi}{2}q_{i,t}^2$. We also assume that agents cannot costlessly extract signals from market prices.[8] Finally, we assume that market participants agree to disagree and trade at the prevailing market price during periods 1 and 2.

[8]The framework we have in mind involves inattentive agents submitting demand schedules to a Walrasian auctioneer. These demand schedules are only revised during periods during which agents acquire a new signal.

2.4.1 Portfolio choice problem

Investors in our model derive utility from end-of-life consumption, C_3. The end-of-life consumption of agent i can be represented as $\sum_{t=0}^{3} q_{i,t}\left(V - p_t\right)$.

Therefore, agent i's portfolio choice problem can be written as:

$$\max_{q_{i,t}} \left\{ \sum_{t=1}^{3} U_{i,t} \right\} \equiv \max_{q_{i,t}} \left\{ \sum_{t=1}^{3} \left(q_{i,t} \mathbb{E}_t^i \left[V - p_t \right] - \frac{\psi}{2} q_{i,t}^2 \right) \right\} \tag{2.4}$$

As shown by Pouget et al. (2016), this maximization problem is equivalent to maximizing utility for each period separately.

2.4.2 Information acquisition problem

At the beginning of period 1, investor i chooses the probability λ_1 with which she acquires signal $S_{i,1}$ and updates her expectations. At the beginning of period 2, agent i chooses the probability with which she acquires the public signal S_2.

The cost of information acquisition is specified by the function $C(\lambda_{i,t})$. We assume that $C(\cdot)$ is a strictly increasing convex function of $\lambda_{i,t}$. In particular, we consider the following functional form:

$$C(\lambda_{i,t}) = \frac{\varphi}{\kappa + 1} \lambda_{i,t}^{\kappa+1}, \text{ with } \kappa > 1 \text{ and } \varphi > 0 \tag{2.5}$$

In this specification, the φ parameter shifts the marginal cost of information acquisition and κ influences the local curvature of the $C(\cdot)$ function.

Given the setup presented in this section, agents i's choice of λ at the beginning of period t solves the following problem:

$$\max_{\lambda_{i,t}} \left\{ \mathbb{E}_{t-1}^i \left[U_{i,t} \right] - C(\lambda_{i,t}) \right\} \text{ for } t \in \{1, 2\} \tag{2.6}$$

subject to the constraint that $\lambda_{i,t} \in [0, 1]$.

2.4.3 Equilibrium

In order to solve the model, we follow the approach in Kacperczyk et al. (2016) that involves three steps.

Step 1: Solve for optimal portfolios, given information sets.

Agent i's time t demand takes on the following form:

$$\begin{aligned} q_{i,0} &= 0 \\ q_{i,t} &= \frac{\mathbb{E}_t^i\left[V\right] - p_t}{\psi} \text{ for } t = \{1, 2\} \end{aligned} \tag{2.7}$$

Step 2: Clear the asset market.

The market-clearing condition for the risky asset is:

$$\int_0^1 q_{i,t} di = x \tag{2.8}$$

Plugging in the expression for agent i's demand into the market clearing condition, we obtain the following expression for the price of the risky asset:

$$p_t = \int_0^1 \mathbb{E}_t^i[V] di - \psi x = \bar{\mathbb{E}}_t[V] - \psi x \tag{2.9}$$

where we use the $\bar{\mathbb{E}}_t[\cdot]$ notation to denote average (consensus) expectations.

The market participants are homogeneous at the beginning of period 1. Then the period 1 price of the risky asset is:

$$p_1 = \lambda_1 \mathbb{E}_1\left[V\right] + \left(1 - \lambda_1\right) \mathbb{E}_0[V] - \psi x = \lambda_1 \mathbb{E}_1[V] - \psi x \tag{2.10}$$

where $\mathbb{E}_1[V] = \frac{\sigma_V^2}{\sigma_V^2 + \sigma_C^2} S_1$ represents the expectation of the attentive agents.

At the beginning of period 2, attentive (A) and inattentive (I) agents have different information sets. Then the time 2 price of the risky asset is::

$$p_2 = \lambda_1 \lambda_{2|A} \mathbb{E}_2^C[V] + (1-\lambda_1)\lambda_{2|I}\mathbb{E}_2^P[V] + (1-\lambda_{2|A})\lambda_1 \mathbb{E}_1[V] + (1-\lambda_1)(1-\lambda_{2|I})\mathbb{E}_0[V] - \psi x \quad (2.11)$$

where $\lambda_{2|A}$ and $\lambda_{2|I}$ represent the optimal updating probabilities for agents who are attentive and inattentive to the private signals, respectively. Here $\mathbb{E}_2^C[V] = \frac{\sigma_V^2 \sigma_S^2}{\sigma_V^2 \sigma_S^2 + \sigma_V^2 \sigma_C^2 + \sigma_C^2 \sigma_S^2} S_1 + \frac{\sigma_V^2 \sigma_C^2}{\sigma_V^2 \sigma_S^2 + \sigma_V^2 \sigma_C^2 + \sigma_C^2 \sigma_S^2} S_2$ and $\mathbb{E}_2^P[V] = \frac{\sigma_V^2}{\sigma_V^2 + \sigma_S^2} S_2$.[9]

Step 3: Solve for information choices.

In Appendix 2.A.1, we show that agent i's optimal $\lambda_{i,t}$ for $t \in \{1, 2\}$ is the solution to the following equation:

$$\frac{\lambda_{i,t}}{\psi} \mathrm{var}_{i,t-1}\left(\mathbb{E}_t[V]\right) - \varphi \lambda_{i,t}^{\kappa} = 0 \qquad (2.12)$$

Based on the expression in (2.12), the optimal values of λ_1, $\lambda_{2|A}$, and $\lambda_{2|I}$ solve the following equations:

$$\begin{aligned}
\lambda_1&: \ \frac{\sigma_V^4 \left(\sigma_V^2 + \sigma_S^2\right)}{\psi \left(\sigma_V^2 + \sigma_C^2\right)^2} \lambda - \varphi \lambda^{\kappa} = 0 \\
\lambda_{2|I}&: \ \frac{\sigma_V^4}{\psi \left(\sigma_V^2 + \sigma_S^2\right)} \lambda - \varphi \lambda^{\kappa} = 0 \\
\lambda_{2|A}&: \ \frac{\sigma_V^4 \left(\sigma_S^4 + \sigma_C^4\right)\left(\sigma_V^2 + \sigma_S^2\right)}{\psi \left(\sigma_V^2 \left(\sigma_C^2 + \sigma_S^2\right) + \sigma_C^2 \sigma_S^2\right)^2} \lambda - \varphi \lambda^{\kappa} = 0
\end{aligned} \qquad (2.13)$$

2.4.4 Model implications for forecast error predictability

In the context of our model, the tests presented in Section 2.3 examine the reaction of consensus expectations to the public signal, S_2. Therefore, the model-implied version of the forecast error predictability coefficients is:

[9] An assumption underlying this expression is that agents who are inattentive in period 1 do not gain access to a private signal when the update their expectations in period 2.

$$\beta^{CG} = \frac{\operatorname{cov}\left(V - \bar{\mathbb{E}}_2\left[V\right], \bar{\mathbb{E}}_2\left[V\right] - \bar{\mathbb{E}}_1\left[V\right]\right)}{\operatorname{var}\left(\bar{\mathbb{E}}_2\left[V\right] - \bar{\mathbb{E}}_1\left[V\right]\right)} \tag{2.14}$$

where $\bar{\mathbb{E}}_1$ and $\bar{\mathbb{E}}_2$ are the consensus expectations of V at time $t = 1$ and $t = 2$. The theoretical expectations correspond $\mathbb{F}_t\left[e_{f,\tau}\right]$ and $\mathbb{F}_{t-1}\left[e_{f,\tau}\right]$ from Section 2.3.

An analytical expression for the error predictability coefficient and its derivation are outlined in Appendix 2.A.2. The actual expression for $\beta^C G$ is not particularly intuitive. Therefore, in this section, we focus on a model with a single trading period. In a one-period setting with a single public signal,

$$\beta^{CG} = \frac{1-\lambda}{\lambda} \tag{2.15}$$

Consequently, the partial derivative of the forecast error predictability coefficient with respect to fundamental volatility takes on the following form:

$$\frac{\partial \beta^{CG}}{\partial \sigma_V^2} = -\frac{\frac{\partial \lambda}{\partial \sigma_V^2}}{\lambda^2} < 0 \tag{2.16}$$

The expression in Equation (2.16) shows that our model generates information rigidity patterns consistent with those documented in the previous section: our model predicts that high-volatility periods are associated with lower information rigidity. The logic underlying this relation extends to the multi-period version of our model.

Volatility and information rigidity

In the context of the stylized model developed in this paper, we study the effects of volatility on the information acquisition decision by examining the relation between the probability of signal acquisition in periods 1 and 2 and fundamental variance σ_V^2.

In order to provide analytical expressions for λ, we focus on the non-zero solutions of the equations in (2.12) for the special case of $\kappa = 2$.[10] The optimal probabilities of information acquisition take on the following form:

$$\begin{aligned}
\lambda_1 &= \frac{\sigma_V^4\left(\sigma_V^2+\sigma_S^2\right)}{\psi\varphi\left(\sigma_V^2+\sigma_C^2\right)^2} \\
\lambda_{2|I} &= \frac{\sigma_V^4}{\psi\varphi\left(\sigma_V^2+\sigma_S^2\right)} \\
\lambda_{2|A} &= \frac{\sigma_V^4\left(\sigma_S^4+\sigma_C^4\right)\left(\sigma_V^2+\sigma_S^2\right)}{\psi\varphi\left(\sigma_V^2\left(\sigma_C^2+\sigma_S^2\right)+\sigma_C^2\sigma_S^2\right)^2}
\end{aligned} \tag{2.17}$$

The comparative statics with respect to σ_V^2 are detailed in Appendix 2.A.3. All of the derivatives with respect to σ_V^2 are positive, i.e. agents are willing to expend more resources to acquire signals during high-volatility periods. Assuming that the state of the technology determines the cost of acquiring information and is, therefore, fixed in the short-run (i.e. it is unaffected by the state of the business cycle), our model produces results that are consistent with the patterns of information rigidity documented in Section 2.3.

Information rigidity over time

The assumption that the cost of information acquisition at time t is determined by the state of the technology at time t also allows the model to account for the decline in information rigidity between 1986 and 2021.

In the context of our framework, the φ parameter governs the cost of information acquisition. We model technological innovations as lower value of φ. Lower values of φ correspond to higher values of λ_1, $\lambda_{2|A}$, and $\lambda_{2|I}$, as we indicated by the expressions in (2.17).

[10]Through numerical solutions we find that the logic of this special case is generalizable to the $\kappa > 2$ case.

2.5 Asset Pricing Implications

Our model generates patterns of information rigidity that are consistent with the stylized facts documented in Section 2.3. In this section, we explore the asset pricing implications of the model.

2.5.1 Momentum

The model developed in the previous section provides us with three momentum signals, which we term long-run momentum signal ($p_2 - p_0$) and short-run momentum signals ($p_1 - p_0$ and $p_2 - p_1$).

Following Luo, Subrahmanyam, and Titman (2020), we define the unconditional short-run momentum parameter as:

$$\mathrm{MOM}_S = \frac{\mathrm{cov}\left(p_2 - p_1, p_1 - p_0\right) + \mathrm{cov}\left(V - p_2, p_2 - p_1\right)}{2} \tag{2.18}$$

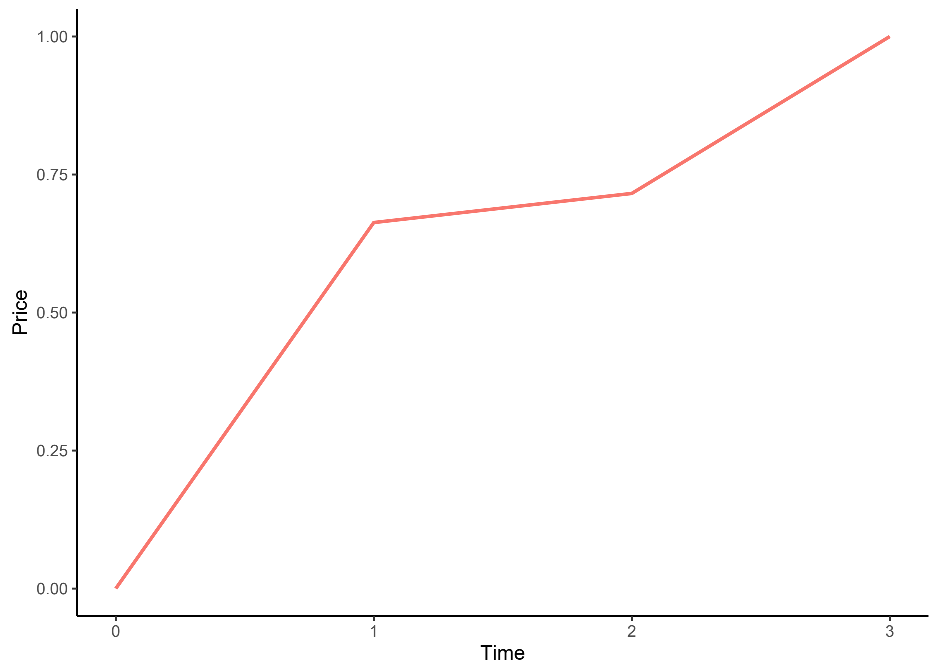
and the long-run momentum parameter as:

$$\mathrm{MOM}_L = \mathrm{cov}\left(V - p_2, p_2 - p_0\right) \tag{2.19}$$

We derive closed-form expressions for the momentum parameters in Appendix 2.A.4. In order to build intuition regarding the predictions of the model, we rely on simulations involving representative price paths. The average price path based on 10,000 simulated paths is shown in Figure 3. In this example, we focus on the case of a positive innovation, $V > 0$. The case with a negative innovation is entirely symmetric.

The figure is based on the following parameters: $\sigma^2_{V,low} = 0.35$, $\sigma^2_{V,high} = 0.6$, $\sigma^2_S = 0.7$, $\sigma^2_C = 0.3$. The transaction cost parameter ψ is set to 1 and the total supply of the risky asset, x, is set to 10^{-5}. We set the probability of a low volatility state is set to 0.90. The parameters associated with the cost function are $\varphi = 0.7$ and $\kappa = 4$. This combination of

Figure 3. Average Price Path



The figure above depicts the price paths for the model in Section 2.4. The x-axis represents time and the y-axis is the price scaled by fundamental value V . The figure is based on the following parameters:$\pi_{hv} = 0.10$, $\sigma^2_{V,lv} = 0.35$, $\sigma^2_{V,hv} = 0.60$, $\sigma^2_S = 0.70$, $\sigma^2_C = 0.30$, $\psi = 1$, $x = 10^{-5}$, $\varphi = 0.70$, $\kappa = 4$.

parameters is chosen to generate probabilities of information acquisition roughly in line with those documented in Section 2.3.

As indicated in Figure 3, the model developed in Section 2.4 generates both short-run and long-run momentum. In the context of our model, the consensus reaction to news is determined by the interaction of two effects, which push expectations in opposite directions: the rational inattention effect generates underreaction to both private and public signals as information is incorporated into aggregate expectations with a delay (Hong and Stein, 1999). On the other hand, the overconfidence effect generates overreaction to private signals, as in Daniel et al. (1998). Given our parameter choices, the rational inattention effect dominates the overconfidence effect unconditionally and we observe underreaction in both the short run and the long run, i.e. both MOM_S and MOM_L are positive. The result is consistent with the fact that both slow (e.g. 12-2) and fast (e.g. 7-2) momentum strategies are profitable unconditionally.

2.5.2 Momentum and volatility

Daniel and Moskowitz (2016) show that long-run (12-month) momentum tends to have low returns following high volatility periods. In Table 2, we verify that these results continue to hold in our extended sample. In particular, we estimate the following regression to examine the effects of volatility on the profitability of momentum:

$$r_{\mathrm{WML},t} = \alpha + \alpha_{\Delta} z_{t-1} + \varepsilon_t \tag{2.20}$$

where $r_{\mathrm{WML},t}$ is the return of the 12-2 momentum strategy, obtained from Kenneth French's data library and z_{t-1} is one of the volatility-related variables considered in Section 2.3 (realized/implied volatility or a high volatility dummy).

The results in Table 2 are consistent with the findings of Daniel and Moskowitz (2016): momentum earns significant positive returns in the order of 10.6% per year over our sample, unconditionally. The momentum strategy delivers significantly lower returns following high-volatility periods. Following periods which are in the top 20% in terms of volatility, momentum returns tend to be negative. This result is also consistent with the findings of Barroso and Wang (2022) who show that the momentum profits occur after periods of low volatility. In Appendix 2.B.1, we repeat the tests presented in this table using the $t-12$ to $t-1$ momentum factor from the Jensen-Kelly-Pedersen data repository. The results obtained using the Jensen-Kelly-Pedersen momentum factor mirror those in Table 2.

To understand the implications of our model for the impact of volatility on the profitability of momentum, we simulate our model separately for periods of high and low volatility. The results are presented in Figure 4. The parameters used to generate the figure are discussed in the previous subsection.

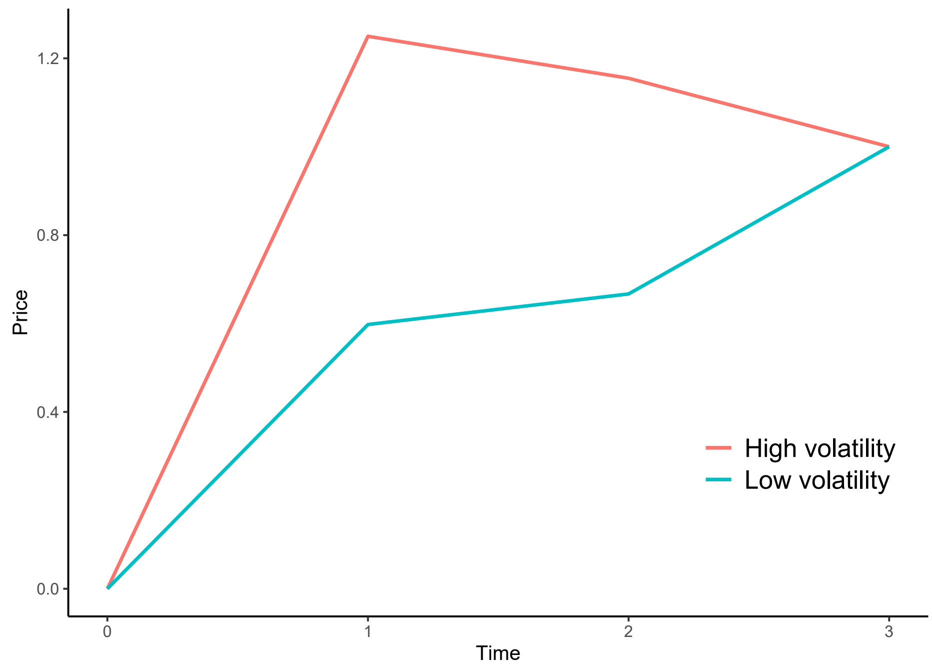
The mechanism that allows our model to generate diminishing momentum profitability during high-volatility episodes goes through the information acquisition channel. Higher values of σ_V^2 translate into higher probabilities of information acquisition, as shown in the

Table 2
Momentum and Volatility

	Dependent variable:						
	$r_{\text{WML},t}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		RV			IV		
$\hat{\sigma}_{\text{mkt}}$		-2.839^{**}			-1.750^{**}		
		[−2.038]			[−1.977]		
$\mathbf{1}_{\text{HV}}$			-0.026^{*}	-0.036		-0.028^{**}	-0.066^{***}
			[−1.805]	[−1.459]		[−2.077]	[−3.092]
Constant	0.009^{**}	0.037^{***}	0.014^{***}	0.012^{***}	0.038^{***}	0.014^{***}	0.015^{***}
	[2.207]	[2.930]	[3.983]	[3.459]	[2.911]	[4.586]	[4.564]
Observations	432	432	432	432	422	422	422
$P(\mathbf{1}_{\text{HV}} = 1)$	–	–	0.20	0.10	–	0.20	0.10

This table reports the results for the following regression: $r_{\text{WML},t} = \alpha + \alpha_{\Delta} z_{t-1} + \varepsilon_t$. In columns (2) and (5) z_{t-1} is realized and implied volatility, respectively. In columns (3), (4), (6), and (7) z_{t-1} is a high volatility indicator. Realized volatility is based on daily value-weighted CRSP returns for a period of 126 days. The CBOE $S\&P100$ Volatility Index is used as a measure of implied volatility and the volatility of value-weighted CRSP is used as a measure of realized volatility. The standard errors are computed using the Newey and West (1987) methodology with six lags. The corresponding t-statistics are in square brackets. Significance at the 1%, 5%, and 10% is denoted by $***$, $**$, and $*$, respectively. The momentum return data is obtained from Kenneth French's website and covers the period between 01/1986 and 12/2021.

Figure 4. Average Price Path



The figure above depicts the price paths based on the model in Section 2.4 for high-volatility and low-volatility states. The x-axis represents time and the y-axis is the price scaled by fundamental value V . The figure is based on the following parameters:$\sigma^2_{V,lv} = 0.35$, $\sigma^2_{V,hv} = 0.60$, $\sigma^2_S = 0.70$, $\sigma^2_C = 0.30$, $\psi = 1$, $x = 10^{-5}$, $\varphi = 0.70$, $\kappa = 4$.

equations in (2.17). As information is incorporated into the aggregate expectations faster, the opportunity for momentum profits diminishes.

If volatility is high enough, the overconfidence effect comes to dominate the rational inattention effect and the consensus expectations overreact to the private signals. The initial overreaction to the private signals is partially corrected during the subsequent trading period but p_2 does not revert all the way to its rational level, i.e. the initial overreaction is further corrected during period 3.

Our model also provides a rationale for volatility management enhancing the profitability of momentum (Barroso and Santa-Clara, 2015; Moreira and Muir, 2017): the volatility-managed momentum strategy avoids the large losses of the baseline momentum strategy by limiting investors' exposure to momentum during high-volatility episodes and increasing investors' exposure to momentum during low-volatility episodes when momentum profits tend to be high.

2.5.2.1 Mixed Momentum Strategy

According to our model, the relative profitability of the long-run and short-run momentum strategies varies with fundamental volatility. During low-volatility periods, the long-run momentum strategy dominates the short-run momentum strategy due to the fact that the short-run momentum signal predicts profits with a low signal-to-noise ratio. The long-run momentum signal has a higher signal-to-noise ratio as the noise components of the $t = 1$ signal and the $t = 2$ signal cancel each other out on average.

On the other hand, the short-run momentum signal at time $t = 2$ is positively correlated with the price appreciation at time $t = 3$ $(V - p_2)$.[11] This implies that the conditional short-run momentum strategy outperforms the long-run momentum strategy during high-volatility periods.

Based on the predictions of our model regarding the relative profitability of short-run and long-run momentum strategies, we propose a mixed momentum strategy similar to the one in Goulding et al. (2022). Our proposed strategy involves mixing short-run and long-run momentum signals in a way that accounts for the impact of market volatility on the relative profitability of the short-run and long-run momentum strategies.

In particular, we propose the following strategy:

$$\begin{aligned} r_{MM,t} &= w_{t-1} r_{LM,t} + (1 - w_{t-1}) r_{SM,t} \\ w_t &= \frac{K}{\sigma_{mkt,t-1}} \end{aligned} \tag{2.21}$$

Here, the weight of the long-run strategy, w_t is inversely related to market volatility and K is constant. We use $r_{LM,t}$ to denote the return of the long-run momentum strategy and $r_{SM,t}$ to denote the return of the short-run momentum strategy.

[11] See Proposition 1 in Daniel et al. (1998).

The $w_t r_{LM,t+1} \left(\frac{K}{\sigma_{mkt,t-1}} r_{LM,t}\right)$ term represents a volatility-managed momentum strategy, which has been shown to significantly enhance the profitability of conventional 12-2 momentum. In order to limit our focus on the implications of mixing short- and long-run momentum, we also consider a restricted version of the strategy in (2.21):

$$\begin{aligned} r_{MMR,t} &= w^{'}_{t-1} r_{LM,t} + (1 - w^{'}_{t-1}) r_{SM,t} \\ w^{'}_{t-1} &= \min\left(\frac{K^{'}}{\sigma_{mkt,t-1}}, 0.999\right) \end{aligned} \tag{2.22}$$

The restricted strategy rules out the possibility of levering up and investing heavily in the long-run momentum strategy during low-volatility periods. It also forces us to invest non-zero amounts in the short-run momentum strategy.

Implementation

In order to examine the empirical performance of the mixed momentum strategy, we start by downloading the long-run momentum signal (12-2) from Kenneth French's data library. Then we construct a short-run momentum signal following the methodology described on Kenneth French's website. We choose to use the returns of the stocks in months $t-2$ and $t-3$ as our short-run momentum signal (3-2). At the end of month t, we sort stocks into deciles based on the short-run momentum signal using NYSE breakpoints. Using the month $t-2$ return as a momentum signal produces similar results. In Appendix 2.B.2, we reimplement the mixed momentum strategies using the $t = 12$ to $t = 1$ and $t = 3$ to $t = 1$ momentum factors from the Jensen-Kelly-Pedersen data repository.

In order to ensure consistency with the rest of our paper, we use the realized variance of daily market returns during the 126 days preceding the portfolio formation date as a proxy for $\sigma_{mkt,t}$. Using the realized volatility of the market for month $t-1$ produces similar results.

To implement our strategy, we start with the sample between 01/1950 and 12/2021. We choose 1950 as a start date to avoid the effects of the Great Depression and World War II.

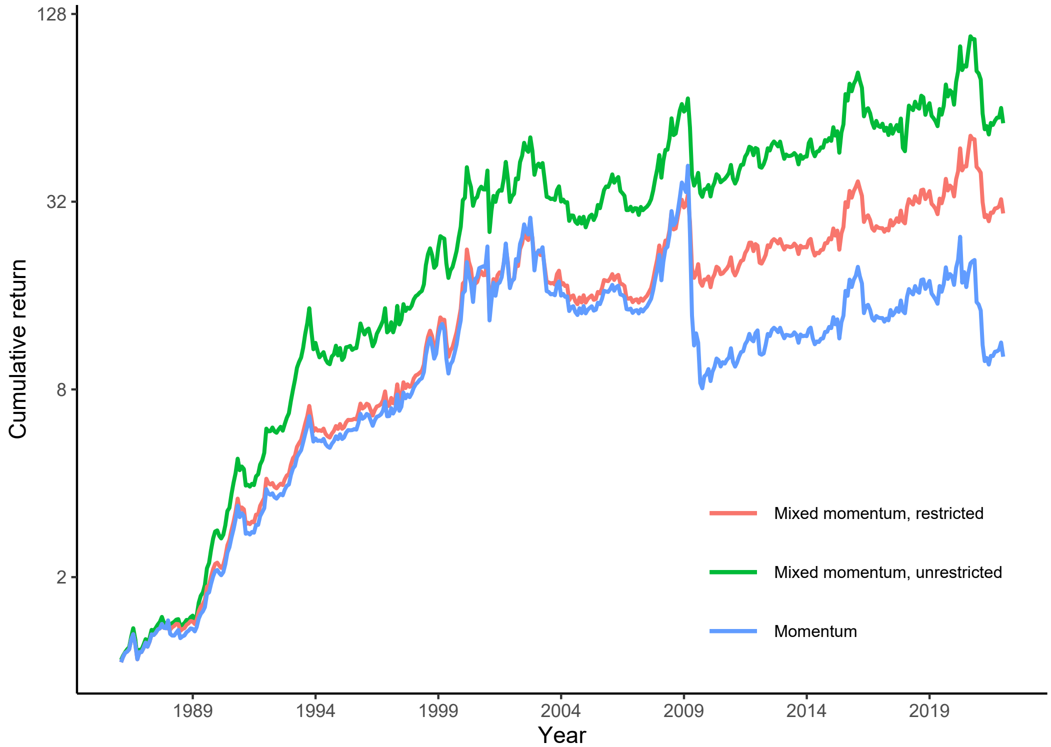
We use the period between 01/1950 and 12/1985 as our training sample and evaluate the performance of the mixed momentum strategy over the period between 01/1986 and 12/2021. The test sample is chosen to match the sample used to test the rationality of equity analysts' earnings expectations. We use the training sample to pin down the value of the K $(K^{'})$ parameter. We choose the value of K $(K^{'})$ that maximizes the Sharpe ratio of the unrestricted (restricted) mixed momentum strategy during the training sample. In order to keep our analysis as simple as possible, we use the same K $(K^{'})$ parameter during our entire test sample.

Performance

Figure 5 plots the cumulative nominal returns to the unrestricted and restricted mixed momentum strategies compared to the 12-2 momentum strategy over our test sample. We invest $1 in 1986 and plot the cumulative returns on a log scale for each strategy. The unrestricted and restricted mixed momentum strategies generate about $57 and $29, respectively at the end of our sample, compared with about $10 for the 12-2 momentum strategy.

The restricted momentum strategy follows the returns of the 12-2 strategy very closely for the majority of our sample: the strategy puts more than 90% weight on the long-run momentum signal about 64% of the time. The divergence between the two strategies takes place during the 6% of our sample when the mixed momentum strategy places less than 50% weight on the short-run momentum signal. These high-volatility episodes coincide with the crashes of the 12-2 strategy and the mixed momentum strategy placing less weight on the long-run momentum signal lessens the impact of those crashes, i.e. the returns of the mixed momentum strategy behave in a way that is consistent with the predictions of our model. The unrestricted strategy achieves returns superior to the restricted strategy by taking on relatively more risk when volatility is low, consistent with previous findings on volatility-managed momentum.

Figure 5. Mixed Momentum Strategy



The figure above depicts the cumulative returns of the restricted and unrestricted versions of the mixed momentum trading strategy developed in Section 2.5.2.1, along with the returns of the baseline momentum strategy. In this figure, we consider the time period between 01/1986 and 12/2021.

Spanning Regressions

To examine the ability of existing trading strategies to account for the profitability of mixed momentum, we estimate time series regressions of mixed momentum on baseline 12-2 momentum, as well as the Fama and French (2015) and Hou, Xue, and Zhang (2014) models, following Moreira and Muir (2017):

$$r_{MM,t} = \alpha + \sum_{j=1}^{N} \beta_j F_j + \varepsilon_t \tag{2.23}$$

A positive α implies that investors who are already trading the explanatory strategies could realize significant gains by also trading the strategy on the left-hand side of the regression.

Table 3

Mixed Momentum Spanning Regressions

	Momentum	FF6	q-Factor
Panel A: Unrestricted			
$\hat{\alpha}$	0.45	0.68	0.89
t-statistic	3.04	3.16	2.25
Panel B: Restricted			
$\hat{\alpha}$	0.31	0.48	0.70
t-statistic	2.87	2.82	2.02

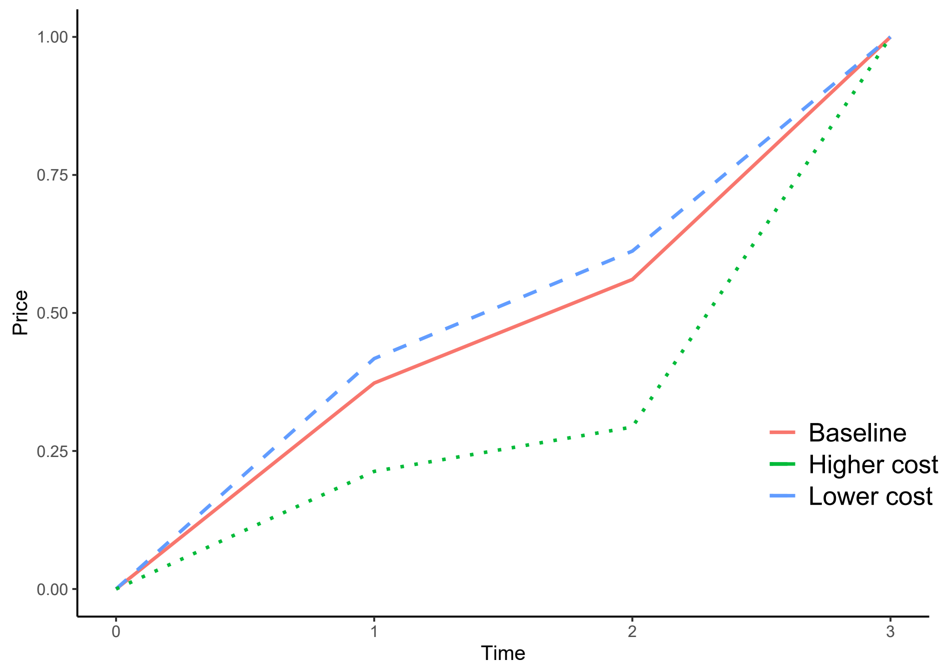
This table reports the $\hat{\alpha}$ (in %) obtained by estimating the following regression: $r_{MM,t} = \alpha + \sum_{j=1}^{N} \beta_j F_{j,t} + \varepsilon_t$ where the factors are: the return on the 12-2 momentum strategy, the five Fama-French factors + WML, or the four Hou-Xue-Zhang factors. The full sample covers the period between 01/1986 and 12/2021. The t-statistics are based on standard errors computed following White (1980).

The intercepts of these spanning regressions for both the restricted and unrestricted versions of the mixed momentum strategy are reported in Table 3. The intercepts are positive and statistically significant in all specifications considered in the table. The two mixed momentum strategies have annualized α's of about 5.4% (3.7%) relative to the 12-2 momentum strategy. The results reported in columns (2) and (3) show that the state-of-the-art factor models cannot fully account for the returns of the mixed momentum strategy.

2.5.3 Attenuation of momentum

A prediction of the model is that high λ periods are associated with low momentum returns. Additionally, we show that information rigidity displays a secular decline during the period between 1986 and 2021. Therefore, a prediction of our model is that the profitability of momentum declines over our sample. In the context of our model, the declining cost of information acquisition leads to declining information rigidity and lower profitability of momentum. Figure 6 depicts the model-implied relation between the marginal cost of information acquisition (φ) and the profitability of momentum.

Figure 6. Cost of Information and Momentum



The figure above depicts the price paths based on the model in Section 2.4 for different marginal costs of information acquisition. The x-axis represents time and the y-axis is the price scaled by fundamental value V. The figure is based on the following parameters: $\sigma_V^2 = 0.35$, $\sigma_S^2 = 0.70$, $\sigma_C^2 = 0.30$, $\psi = 1$, $x = 10^{-5}$, $\varphi_{\text{baseline}} = 0.70$, $\varphi_{\text{low cost}} = 0.50$, $\varphi_{\text{high cost}} = 0.90$ and $\kappa = 4$.

In order to test the prediction of the model regarding the declining profitability of momentum, we estimate the following regression, based on Chordia et al. (2014):

$$Y_t = ae^{bt+u} \tag{2.24}$$

where Y_t is one plus the return of the 12-2 momentum strategy and t is a time index. We scale the time index to be between -1 and 1 so that the mean of the time variable is zero, as in the original paper. We estimate the regression using momentum return data for the period between 01/1986 and 12/2021 to match the data used to generate Figure 1.

The estimate of the b coefficient is -0.014. It is significant at the 5% level (p-value of one-tailed test 0.025).[12] Since the return of the momentum strategy is positive, a negative coefficient signifies a decline in the profitability of momentum over time, thus providing evidence consistent with the prediction of our model.

2.5.4 Return Predictability

2.5.4.1 Return predictability in the model

According to the model, the period 2 expectation of the period 3 price appreciation under the physical measure is:

$$\mathbb{E}_2^P\left[R_3\right] = \mathbb{E}_2^P\left[V - p_2\right] = \bar{\lambda}_2\left(\mathbb{E}_2^P\left[V\right] - \bar{\mathbb{E}}_2\left[V\right]\right) + \psi x \tag{2.25}$$

where $\bar{\lambda}_2 \equiv \int_0^1 \lambda_{i,2} di$ is the average probability of acquiring the signal in period 2.

The risk premium component of returns, ψx is constant in our model. Therefore, return predictability is driven by the gap between the objective payoff expectations and the average payoff expectations of the market. The magnitude of the gap is determined by the interplay of the two effects, which distort market participants' expectations: rational inattention and overconfidence.

In the context of a multi-period economy, Equation (2.25) states that returns are forecastable using the predictable component of earnings forecast errors. This model prediction carries significant intuitive appeal. If a stock's earnings expectations are overly optimistic, its actual earnings will, on average, fail to meet consensus expectations, which will translate into low returns, i.e. stocks with overly optimistic earnings expectations will deliver low returns. A similar logic applies to stocks with overly pessimistic earnings expectations delivering high returns.

[12] Following Chordia et al. (2014), we test the null hypothesis of no decline in the profitability of momentum.

2.5.4.2 Rational earnings expectations

In order to empirically test the model implications regarding return predictability, we need to measure the earnings expectations under the objective measure $\mathbb{P}$.

The traditional approach in the literature has been to use a linear regression framework to estimate the predictable component of equity analysts' forecast errors (So, 2013) or to use the cross-sectional median earnings forecasts as a proxy for objective expectations (La Porta, 1996). However, recent work (e.g. Cao and You, 2021 and Van Binsbergen et al., 2022) documents that models that allow for non-linear relations between realized earnings and earnings predictors significantly improve our ability to forecast earnings.

In this paper, we follow Van Binsbergen et al. (2022) and utilize a tree-based algorithm that can accommodate non-linearities and interactions among the predictors. In particular, we opt for an Extreme Gradient Boosting (XGBoost) algorithm (Chen and Guestrin, 2016). XGBoost is chosen due to its speed and superior performance in a large number of settings. Appendix 2.C contains a brief discussion regarding the technical aspects related to XGBoost. In the main body of the paper, we focus our discussion around issues related to the implementation of the algorithm.

Earnings expectations

Let the earnings of firm f during fiscal year $\tau + h$ be

$$e_{f,\tau+h} = \mathbb{E}_t^P \left[e_{f,\tau+h} \right] + \varepsilon_{f,t,\tau+h} \tag{2.26}$$

where $\mathbb{E}_t^P \left[e_{f,\tau+h} \right] = g(z)$ and z is the $P \times F \times T$-dimensional matrix of predictors. We assume that $g(\cdot)$ is a flexible function of these predictors. We only impose the restriction that $g(\cdot)$ does not depend on f or t, i.e. the function is the same for all firms and over time.

Our goal in this section is to use XGBoost to approximate the function $g(\cdot)$.

XGBoost is based on decision trees in which the data is recursively split into non-intersecting partitions. The algorithm approximates $g(\cdot)$ with the average value of the outcome variable in each partition. At each step, the algorithm groups observations that behave similarly by minimizing the mean squared error when the average value of the dependent variable in each partition is used to form forecasts. Due to the large number of potential splits, tree-based methods rely on "greedy" optimization, which involves myopically minimizing forecast errors during each split.

The $g(\cdot)$ for a tree with K terminal nodes (leaves) can be formally written as:

$$\mathbb{E}^P\left[y\right] = g(z) = \sum_{\kappa=1}^{K} x_\kappa \mathbf{1}_{\{z \in C_\kappa\}} \tag{2.27}$$

where x_κ is the sample average of the dependent variable in partition κ and is given by:

$$x_\kappa = \frac{1}{N_\kappa} \sum_{y: z_p \in C_\kappa} y \tag{2.28}$$

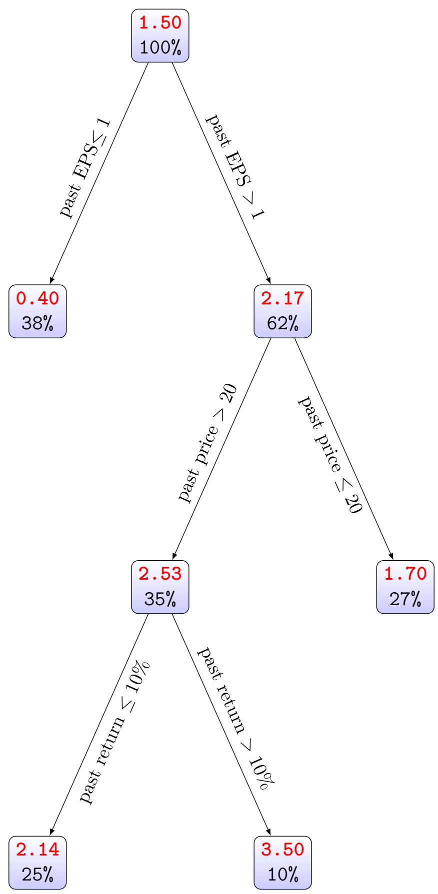
and region C_κ is chosen by forming hyper-regions in the space of predictors:

$$C_\kappa = \left\{ z_p \in \times_{p \in P} Z_p : \underline{z}_p^\kappa < z_p \leq \bar{z}_p^\kappa \right\} \tag{2.29}$$

where $\times$ denotes Cartesian product, P is the number of predictors, and each predictor z_p can take values in set Z_p.

The decision tree in Figure 7 illustrates the contents of Equations (2.27), (2.28), and (2.29) using a simple example. The outcome variable in this example is EPS and the predictors we consider are lagged EPS realizations, lagged prices, and lagged returns. Given the structure in the figure, $g(\cdot)$ takes on the following form:

Figure 7. Decision Tree Example



This figure graphically illustrates the structure of a decision tree. Our goal is to predict EPS and we use past EPS, past price, and past return as predictors. The percentages represent the proportion of our observations that end up in each node and the numbers in red represent the average EPS of the stocks within each node.

$$
\begin{aligned}
g(z;y) = {} & (0.40)\mathbf{1}_{\{past\ EPS \leq 1\}} + (1.70)\mathbf{1}_{\{past\ EPS > 1\}}\mathbf{1}_{\{past\ price \leq 20\}} + \\
& + (2.14)\mathbf{1}_{\{past\ EPS > 1\}}\mathbf{1}_{\{past\ price > 20\}}\mathbf{1}_{\{past\ return \leq 10\%\}} + \\
& + (3.50)\mathbf{1}_{\{past\ EPS > 1\}}\mathbf{1}_{\{past\ price > 20\}}\mathbf{1}_{\{past\ return > 10\%\}}
\end{aligned} \tag{2.30}
$$

Extreme Gradient Boosting

XGBoost is an algorithm that is based on recursively combining forecasts from a large number of weak learners to form a strong learner. During the first step of implementation, the algorithm fits a weak learner to the training sample. At each subsequent step s, the algorithm fits a weak learner to the residuals of a model with $s-1$ trees. The residual forecast is then added to the total with a shrinkage weight $\eta \in (0,1)$. The additional forecasts are shrunken to avoid overfitting the residuals.

In order to implement the XGBoost algorithm, four hyper-parameters need to chosen: γ, `maximum depth`, and `subsample` in addition to the already-described shrinkage parameter η. Hyper-parameter are a characteristic of a model whose value cannot be estimated from data. Therefore, we tune (choose the values of) the hyper-parameters using a cross-validation procedure that we describe later in the paper. The parameter values we use throughout our analysis are presented in Table 4.

Table 4
XGBoost Hyper-parameters

η	0.01
γ	0.15
`maximum depth`	7
`subsample`	0.15
`nrounds`	10000

This table reports the hyper-parameters chosen for the XGBoost algorithm.

The γ parameter determines the minimum loss reduction required to make a further partition on a leaf node of the tree. This parameter controls the total number of trees in the ensemble.

`Maximum depth` determines the maximum depth (complexity) of the tree. More complex models are more likely to overfit the training sample.

`Subsample` determines the ratio of training instance. A value of 0.15 means that XG-Boost randomly collects 15% of the observations to use a training sample for each decision tree.

Earnings forecasts

In our analysis, we focus on two-year earnings forecasts ($FPI = 2$). The long forecasting horizon maximizes the scope for expectation errors and increases the chances of uncovering interesting asset pricing dynamics. Van Binsbergen et al. (2022) find that the bias of equity analysts' expectations increases with the forecasting horizon and de Silva and Thesmar (2023) find that analysts' forecasts outperform statistical forecasts for forecasting horizons of less than one year. To match the frequency of I/B/E/S analyst forecasts, we construct objective expectations for each month.

We deviate from existing papers that forecast earnings by only using variables that are available through CRSP (prices and returns) or I/B/E/S (earnings forecasts and past EPS realizations). We choose this limited set of predictors to avoid basing expectations on information not available to the equity analysts in real time. Variables extracted from financial statements, which have been shown to predict future earnings, may be restated after initial publication and accounting restatements affect stock prices (Hribar and Jenking, 2004 and Palmrose, Richardson, and Scholz, 2004).[13] Therefore, using variables extracted from financial statements may contaminate the out-of-sample tests presented in this paper.

[13] Realized earnings in I/B/E/S are not restated.

An additional advantage of restricting our attention to a limited number of earnings predictors is that by doing so we sidestep the missing data problem outlined in Bryzgalova, Lerner, Lettau, and Pelger (2022). Using the readily-available predictors allows to avoid having to take a stand regarding the appropriate approach to filling in missing firm characteristics.

Using the notation established in this section, the objective earnings expectations are represented as:

$$\mathbb{E}_t^P\left[e_{f,\tau+1}\right] = g\left(e_{f,\tau-1}, \mathbb{F}_t\left[e_{f,\tau+1}\right], P_{f,t}, r_{f,t}\right) \tag{2.31}$$

To implement the XGBoost algorithm, we split our sample into three non-overlapping time periods: training subsample, validation subsample, and testing subsample. The training sample covers the period between 1976 and 1983 and is used to estimate the model using a set of hyper-parameter values.

Our validation sample encompasses the period between 01/1984 and 11/1985 and is used to conduct quasi-out-of-sample tests: we construct predicted earnings for 1986 and 1987 based on our model and compute the mean squared error corresponding to the set of hyper-parameters used to train the model. Next, we reestimate the model using a different set of hyper-parameters. We grid-search over various hyper-parameter combinations and select the combination that minimize the out-of-sample mean squared error of our model. Once the hyper-parameters are chosen, we use the same parameters for all of our out-of-sample tests.

We start our out-of-sample forecasts in 11/1985 and use ten-year rolling windows to train the model. The algorithm provides us with two-year earnings forecasts for the period between 12/1985 and 11/2019. The objective forecasts do not rely on information that is not available to the market participants by the end of month t. Therefore, the trading strategy proposed in the next subsection is implementable in real time.

2.5.4.3 Return predictability in the data

The XGBoost algorithm outlined in the previous subsection provides us with objective two-year earnings forecasts for each month t between 12/1985 and 11/2019. Given the objective forecasts, we can compute the predictable component of analysts' expectation errors using the following formula:

$$\widehat{\mathrm{EE}}_{f,t}^{\tau+1} = \frac{\mathbb{E}_t^P \left[e_{f,\tau+1}\right] - \mathbb{F}_t \left[e_{f,\tau+1}\right]}{P_{f,t-1}} \tag{2.32}$$

Negative values of $\widehat{\mathrm{EE}}$ indicate that analysts' expectations are excessively optimistic and positive values of the measure indicate excessive pessimism. Following Engelberg et al. (2018), the difference between the expectations under the objective and subjective forecasts is normalized by lagged stock prices to ensure greater comparability across firms.

In order to test the predictions of the model regarding return predictability, we sort stocks into quintiles based on the value of $\widehat{\mathrm{EE}}$ at the end of month t. Panel A of Table 5 reports the one-month holding period returns of the five portfolios. The results reported in the table are for the period between 01/1986 and 12/2019.

The results in Panel A of Table 5 support the predictions of the model: the value-weighted portfolio returns increase monotonically in $\widehat{\mathrm{EE}}$. In particular, a portfolio that is long on stocks in the fifth quintile and short on stocks in the first quintile earns an average return of 1.59% per month. For the rest of this paper, we refer to this long-short trading strategy as pessimistic-minus-optimistic or PMO. The t-statistic testing whether the PMO premium is zero is 6.82. Thus, PMO clears the hurdle of a t-statistic ≥ 3.0 proposed by Harvey, Liu, and Zhu (2015). Monthly returns of 1.59% seem large in comparison to existing trading strategies (for comparison, the average value-weighted market return between 01/1986 and 12/2019 is about 1%). However, the profitability of our strategy is in line with the profitability of the strategies considered by Van Binsbergen et al. (2022).

In order to assess the degree to which PMO's profitability is driven by the smallest and most illiquid stocks within our sample, we construct an alternative long-short trading strategy by restricting our investment opportunity set to only include stocks whose market

Table 5
Portfolios Sorted on Expectation Errors

Quintile	1	2	3	4	5	$5-1$
Panel A: All Stocks						
Mean	-0.11	0.59	0.88	1.03	1.48	1.59
t-statistic	-0.32	2.19	3.77	5.10	7.08	6.82
Panel B: Market cap above 90^{th} percentile						
Mean	0.56	0.90	0.85	1.07	1.33	0.77
t-statistic	2.04	3.95	4.00	5.47	6.15	4.74
Panel C: Post-2002 Sample						
Mean	0.15	0.54	0.77	0.94	1.68	1.53
t-statistic	0.30	1.59	2.59	3.83	6.31	5.23

This table reports the time-series average returns on value-weighted portfolios formed based on the predictable component of equity analysts' forecast errors, $\widehat{\text{EE}}$. The full sample in Panel A includes all stocks and covers the period between 01/1986 and 12/2019. In Panel B, the sample is restricted to stocks with market caps above the 90^{th} percentile of market caps within a given year. The sample in Panel C covers the period between 01/2003 and 01/2020.

capitalization is above the 90$^{\text{th}}$ percentile of market capitalizations in a given year.[14] The results of the trading strategy that invests only in large-cap stocks are reported in Panel B of Table 5. Unsurprisingly, the results in this instance are somewhat weaker than those reported in Panel A but the return of the long-short portfolio is still about 0.77% (t-statistic = 4.74) per month. Implementing the PMO trading strategy using only the most liquid stocks in our sample generates significant positive returns.

In Panel C, we examine if the profitability of the PMO strategy is driven by the earlier years within our sample. In this test, we exclude the years between 1986 and 2002 from our sample. The return of the long-short portfolio for the period between 01/2003 and 12/2019 is about 1.53% per month (t-statistic = 5.23) and we are unable to reject the hypothesis that the returns during the second half of our sample are equal to the returns during the first half of our sample (t-statistic = −0.34).

Spanning regressions

We further use spanning regressions to assess the ability of traditional factor models to explain the profitability of the PMO strategy. Specifically, we estimate the following regressions and examine the significance of the regression intercepts (alphas):

$$PMO_t = \alpha + \sum_{j=1}^{N} \beta_j F_{j,t} + \varepsilon_t \tag{2.33}$$

where $F_{j,t}$, $j = \{1, 2, 3, ..., N\}$ are the excess market return, the five Fama and French (2015) factors augmented with the WML factor, or the four factors of the q-factor model (Hou et al., 2014). The results of the spanning regressions are reported in Table 6.

The full-sample tests in Panel A of Table 6 show that PMO earns a significant alpha with t-statistics greater than 5.0 relative to the three models considered in the table. This suggests that the factors considered in the table are unable to fully account for the profitability of the

[14] The minimum market capitalization of a firm included in the megacap sample is about 2 billion in the early parts of the sample and over 21 billion near the end of the sample.

Table 6
PMO Spanning Regressions

	CAPM	FF6	q-Factor
Panel A: Full Sample			
$\hat{\alpha}$	1.90	1.51	1.63
t-statistic	8.97	8.78	6.30
Panel B: 01/1986-12/2002			
$\hat{\alpha}$	1.81	1.22	1.41
t-statistic	5.27	5.55	2.92
Panel C: 01/2003-01/2020			
$\hat{\alpha}$	2.08	1.68	1.87
t-statistic	8.91	8.88	8.00

This table reports the $\hat{\alpha}$ (in %) obtained by estimating the following regression: $PMO_t = \alpha + \sum_{j=1}^{N} \beta_j F_{j,t} + \varepsilon_t$ where the factors are: excess market return, the five Fama-French factors + WML, or the four Hou-Xue-Zhang factors. The full sample covers the period between 01/1986 and 12/2019. The t-statistics are based on standard errors computed following White (1980).

PMO strategy. In Panels B and C we verify that our conclusions regarding the inability of factor models to account for the profitability of PMO hold in both the first half (1986-2002) and the second half (2003-2019) of our sample. It is noteworthy that the performance of PMO improves significantly relative to the explanatory trading strategies in the second half of our sample.

In unreported tests, we show that the profitability of the PMO strategy cannot be explained using a combination of popular trading strategies proposed in the literature. While an explanation of the profitability of the PMO strategy is beyond the scope of this paper, we believe that the PMO strategy should receive a rigorous treatment in the literature going forward.

2.6 Conclusion

In this paper, we document the existence of time variation in the stickiness of financial analysts' expectations. The stickiness of analysts' expectations declines during high-volatility periods. Additionally, expectation stickiness experiences a sustained decline over our sample.

To account for these stylized facts, we build a simple featuring time-varying inattention. We explore the asset pricing implications of our model and show that it is consistent with positive unconditional momentum returns, momentum crashes, the profitability of volatility-managed momentum, and the diminishing profitability of momentum over our sample.

In order to test our model's prediction regarding return predictability, we extract the predictable component of analysts' forecast errors and propose a trading strategy that is long (short) on stocks with excessively pessimistic (optimistic) earnings forecasts. Existing prominent factor models cannot fully explain the profitability of our trading strategy, especially during the second half of our sample.

2.A Derivations

2.A.1 Information Acquisition Problem

Our goal is to obtain an expression for $\mathbb{E}^i_{t-1}\left[U_{i,t}\right]$.

First, we plug in the expression for optimal $q_{i,t}$ and obtain the following maximization problem:

$$\mathbb{E}^i_{t-1}\left[\frac{\mathbb{E}^i_t[V]-p_t}{\psi}\left(\mathbb{E}^i_t[V]-p_t\right)-\frac{\psi}{2}\frac{\left(\mathbb{E}^i_t[V]-p_t\right)^2}{\psi^2}\right]=\frac{1}{2\psi}\mathbb{E}^i_{t-1}\left[\left(\mathbb{E}^i_t[V]-p_t\right)^2\right] \tag{2.34}$$

We expand the square:

$$\frac{1}{2\psi}\mathbb{E}^i_{t-1}\left[\left(\mathbb{E}^i_t[V]\right)^2-2p_t\mathbb{E}^i_t[V]+p_t^2\right] \tag{2.35}$$

The price p_t is only affected by aggregate information choices, therefore each agent i takes the price as given (as in Kacperczyk et al., 2016). The agents taking prices as given, combined with the fact that the law of iterated expectations holds at the individual level, allows us to rewrite the expression above as:

$$\frac{1}{2\psi}\left(\mathbb{E}^i_{t-1}\left[\left(\mathbb{E}^i_t[V]\right)^2\right]-2p_t\mathbb{E}^i_{t-1}[V]+p_t^2\right) \tag{2.36}$$

In this expression, $\mathbb{E}^i_{t-1}\left[\left(\mathbb{E}^i_t[V]\right)^2\right]$is the only term that depends on $\lambda_{i,t}$. Therefore, we need to evaluate the following expression:

$$\frac{1}{2\psi}\mathbb{E}^i_{t-1}\left[\left(\mathbb{E}^i_t[V]\right)^2\right] \tag{2.37}$$

We use the definition of variance to rewrite the expression as:

$$\frac{1}{2\psi}\left(\mathrm{var}^i_{t-1}\left(\mathbb{E}^i_t[V]\right)+\left(\mathbb{E}^i_{t-1}\left[\mathbb{E}^i_t[V]\right]\right)^2\right) \tag{2.38}$$

The second term does not depend on $\lambda_{i,t}$, so we focus on the variance term:

$$\frac{1}{2\psi}\left(\mathrm{var}_{t-1}^{i}\left(\lambda_{i,t}\mathbb{E}_t[V]+(1-\lambda_{i,t})\mathbb{E}_{t-1}[V]\right)\right)=\frac{\lambda_{i,t}^2}{2\psi}\mathrm{var}_{t-1}^{i}\left(\mathbb{E}_t^i[V]\right) \tag{2.39}$$

2.A.2 Information Rigidity Coefficient

The model-implied version of the forecast error predictability coefficients is:

$$\beta^{CG}=\frac{\mathrm{cov}\left(V-p_2,p_2-p_1\right)}{\mathrm{var}\left(p_2-p_1\right)}. \tag{2.40}$$

The numerator and denominator in equation (2.40) are given by

$$\begin{aligned}\mathrm{cov}\left(V-p_2,p_2-p_1\right)=&\tfrac{\bar{K}}{D}\Bigg[\lambda_1^2\lambda_{2|A}\left(\sigma_V^2+\sigma_S^2\right)\left(D\sigma_V^2\sigma_C^2\left(\sigma_S^2-\sigma_C^2\right)-\lambda_{2|A}\sigma_V^2\sigma_C^4\left(\sigma_C^4+2\sigma_V^2\sigma_S^2+\sigma_S^2\sigma_C^2\right)\right)\\&+(1-\lambda_1)\lambda_{2|I}D^2\left(\sigma_V^2+\sigma_C^2\right)\Bigg(\left(\sigma_V^2+\sigma_C^2\right)-2\lambda_1\lambda_{2|A}\sigma_C^2-\lambda_1\sigma_V^2\\&-(1-\lambda_1)\lambda_{2|I}\left(\sigma_V^2+\sigma_C^2\right)\Bigg)+\lambda_1\lambda_{2|A}D\left(\sigma_V^2+\sigma_S^2\right)\left(\sigma_V^2+\sigma_C^2\right)\sigma_C^4\Bigg] \qquad (2.41)\\\mathrm{var}\left(p_2-p_1\right)=&\tfrac{\bar{K}}{D}\Bigg[\lambda_1^2\lambda_{2|A}^2\left(2\sigma_V^4\sigma_S^2+2\sigma_V^2\sigma_S^2\sigma_C^2+\sigma_V^2\sigma_C^4+\sigma_S^2\sigma_C^4\right)\left(\sigma_V^2+\sigma_S^2\right)\sigma_C^4\\&+2\lambda_1(1-\lambda_1)\lambda_{2|A}\lambda_{2|I}D^2\left(\sigma_V^2+\sigma_C^2\right)\sigma_C^2+(1-\lambda_1)^2\lambda_{2|I}^2D^2\left(\sigma_V^2+\sigma_C^2\right)^2\Bigg] \qquad (2.42)\end{aligned}$$

where $D=\sigma_V^2\sigma_S^2+\sigma_V^2\sigma_C^2+\sigma_S^2\sigma_C^2$, and $\bar{K}=\frac{\sigma_V^4}{D\left(\sigma_V^2+\sigma_S^2\right)\left(\sigma_V^2+\sigma_C^2\right)^2}$.

2.A.3 Comparative Statics

The comparative statics of the probability of information acquisition with respect to σ_V^2 are shown below:

$$\frac{\partial\lambda_1}{\partial\sigma_V^2}=\frac{\sigma_V^2\left(2\sigma_C^2\sigma_S^2+3\sigma_C^2\sigma_V^2+\sigma_V^4\right)}{\psi\varphi\left(\sigma_V^2+\sigma_C^2\right)^3}>0 \tag{2.43}$$

$$\frac{\partial\lambda_{2|I}}{\partial\sigma_V^2}=\frac{\sigma_V^2\left(2\sigma_S^2+\sigma_V^2\right)}{\psi\varphi\left(\sigma_V^2+\sigma_S^2\right)^2}>0 \tag{2.44}$$

$$
\begin{aligned}
\frac{\partial \lambda_{2|I}}{\partial \sigma_V^2} &= \frac{\sigma_V^2 \left(\sigma_C^6(\sigma_S^2+\sigma_V^2)(2\sigma_S^2+\sigma_V^2)+\sigma_C^4\sigma_S^2\sigma_V^4\right)}{\psi\varphi(\sigma_C^2(\sigma_S^2+\sigma_V^2)+\sigma_S^2\sigma_V^2)^3}+ \\
&+\frac{\sigma_V^2 \left(\sigma_C^2\sigma_V^4(\sigma_S^2+\sigma_V^2)(4\sigma_S^2+3\sigma_V^2)+\sigma_S^2\sigma_V^6(2\sigma_S^2+3\sigma_V^2)\right)}{\psi\varphi(\sigma_C^2(\sigma_S^2+\sigma_V^2)+\sigma_S^2\sigma_V^2)^3} > 0
\end{aligned} \tag{2.45}
$$

2.A.4 Momentum

The short-run momentum parameter is defined as

$$
\mathrm{MOM}_S = \frac{\operatorname{cov}(p_2-p_1, p_1-p_0)+\operatorname{cov}(V-p_2, p_2-p_1)}{2} \tag{2.46}
$$

The covariances from equation (2.46) are:

$$
\operatorname{cov}(p_2-p_1, p_1-p_0) = \lambda_1\sigma_V^2\bar{K}\Big[(1-\lambda_1)\lambda_{2|I}D\left(\sigma_V^2+\sigma_C^2\right)-\lambda_1\lambda_{2|A}\sigma_C^2\left(\sigma_V^2+\sigma_S^2\right)\left(\sigma_S^2-\sigma_C^2\right)\Big] \tag{2.47}
$$

$$
\begin{aligned}
\operatorname{cov}(V-p_2, p_2-p_1) =& \tfrac{\bar{K}}{D}\Big[\lambda_1^2\lambda_{2|A}\left(\sigma_V^2+\sigma_S^2\right)\Big(D\sigma_V^2\sigma_C^2\left(\sigma_S^2-\sigma_C^2\right)-\lambda_{2|A}\sigma_V^2\sigma_C^4\left(\sigma_C^4+2\sigma_V^2\sigma_S^2+\sigma_S^2\sigma_C^2\right)\Big) \\
&+(1-\lambda_1)\lambda_{2|I}D^2\left(\sigma_V^2+\sigma_C^2\right)\Big(\left(\sigma_V^2+\sigma_C^2\right)-2\lambda_1\lambda_{2|A}\sigma_C^2-\lambda_1\sigma_V^2 \\
&-(1-\lambda_1)\lambda_{2|I}\left(\sigma_V^2+\sigma_C^2\right)\Big)+\lambda_1\lambda_{2|A}D\left(\sigma_V^2+\sigma_S^2\right)\left(\sigma_V^2+\sigma_C^2\right)\sigma_C^4\Big]
\end{aligned} \tag{2.48}
$$

where $D=\sigma_V^2\sigma_S^2+\sigma_V^2\sigma_C^2+\sigma_S^2\sigma_C^2$, and $\bar{K}=\frac{\sigma_V^4}{D\left(\sigma_V^2+\sigma_S^2\right)\left(\sigma_V^2+\sigma_C^2\right)^2}$. Therefore, the unconditional short-run momentum parameter MOM_S from equation (2.46) is

$$
\begin{aligned}
\mathrm{MOM}_S =& \frac{\bar{K}}{2D}\Big[-\lambda_1^2\lambda_{2|A}^2\left(\sigma_C^4+2\sigma_V^2\sigma_S^2+2\sigma_S^2\sigma_C^2\right)\left(\sigma_V^2+\sigma_S^2\right)\sigma_V^2\sigma_C^4 \\
&+(1-\lambda_1)\lambda_{2|I}D^2\left(\sigma_V^2+\sigma_C^2\right)\Big(\left(\sigma_V^2+\sigma_C^2\right)-2\lambda_1\lambda_{2|A}\sigma_C^2 \\
&-(1-\lambda_1)\lambda_{2|I}\left(\sigma_V^2+\sigma_C^2\right)\Big)+\lambda_1\lambda_{2|A}D\left(\sigma_V^2+\sigma_S^2\right)\left(\sigma_V^2+\sigma_C^2\right)\sigma_C^4\Big]
\end{aligned} \tag{2.49}
$$

The long-run momentum parameter is

$$\begin{aligned}\operatorname{cov}(V-p_2,p_2-p_0)=&\bar{K}\Big[(1-\lambda_1)\,D\left(\sigma_V^2+\sigma_C^2\right)\left(\left(1-\lambda_{2|I}\right)\sigma_V^2+\sigma_S^2\right)\\&-\lambda_1\left(\sigma_V^2\sigma_S^2+\left(1-\lambda_{2|A}\right)\sigma_V^2\sigma_C^2+\sigma_S^2\sigma_C^2\right)\left(\sigma_V^2+\sigma_S^2\right)\left(\sigma_S^2-\sigma_C^2\right)\Big].\end{aligned} \tag{2.50}$$

2.B Robustness: Jensen-Kelly-Pedersen Momentum Factor

2.B.1 Momentum and Volatility

In this section, we estimate the same regression we estimated in Section 2.5.2:

$$r_{WML,t}=\alpha+\alpha_\Delta z_{t-1}+\varepsilon_t \tag{2.51}$$

where $r_{WML,t}$ are the time t momentum returns and z_{t-1} is a variable related to realized or implied volatility volatility.

In this section, we use the return of the $t=12$ to $t=1$ momentum factor from the Jensen-Kelly-Pedersen data depository to conduct the tests. Our findings are reported in Table 2.B.1. The results in the table are qualitatively identical to those in Table 2: momentum delivers positive returns during periods of low volatility and the profitability of the strategy declines significantly during periods of high volatility. The dummy variable specifications show that momentum delivers large negative returns if volatility is within the top 20% (or 10%) of full sample volatility.

2.B.2 Mixed Momentum

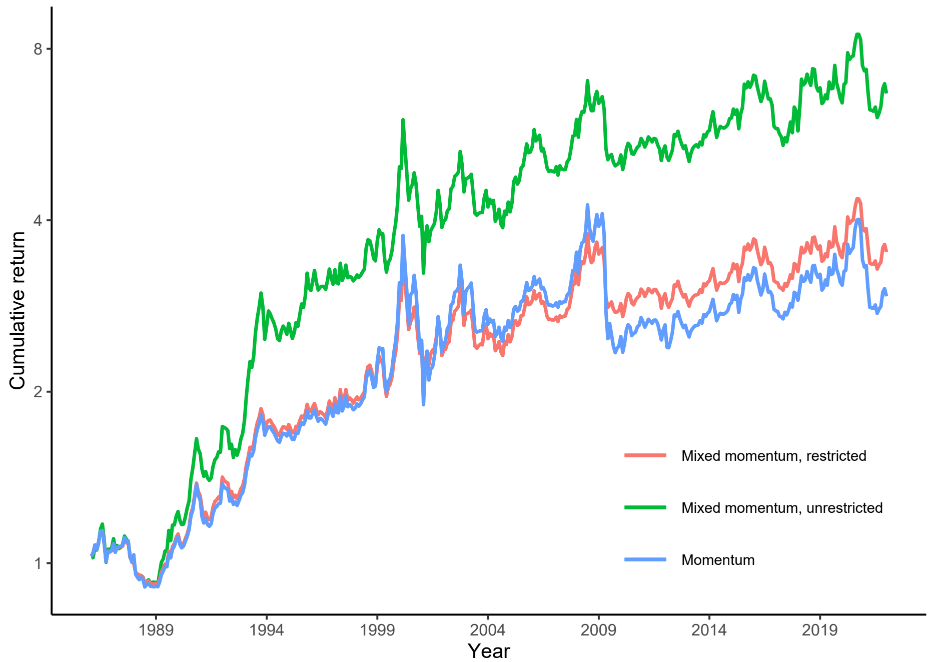
In this section, we implement our mixed momentum strategy using the $t=12$ to $t=1$ and the $t=3$ to $t=1$ momentum factors from the Jensen-Kelly-Pedersen data repository. Figure 2.B.1 depicts the performance of the restricted and unrestricted versions of the mixed momentum strategy relative to the performance of the baseline momentum strategy.

Table 2.B.1
Momentum and Volatility

	Dependent variable:						
	$r_{\text{WML},t}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		RV			IV		
$\hat{\sigma}_{\text{mkt}}$		-1.230^{*}			-0.925^{**}		
		[−1.834]			[−2.227]		
$\mathbf{1}_{\text{HV}}$			-0.014^{*}	-0.010		-0.015^{**}	-0.033^{***}
			[−1.915]	[−0.894]		[−2.128]	[−2.935]
Constant	0.004	0.016^{**}	0.006^{***}	0.005^{**}	0.019^{***}	0.006^{***}	0.007^{***}
	[1.610]	[2.578]	[3.175]	[2.017]	[3.082]	[3.414]	[3.318]
Observations	432	432	432	432	422	422	422
$P(\mathbf{1}_{\text{HV}} = 1)$	–	–	0.20	0.10	–	0.20	0.10

This table reports the results for the following regression: $r_{\text{WML},t} = \alpha + \alpha_{\Delta} z_{t-1} + \varepsilon_t$. In columns (2) and (5) z_{t-1} is realized and implied volatility, respectively. In columns (3), (4), (6), and (7) z_{t-1} is a high volatility indicator. The CBOE $S\&P100$ Volatility Index is used as a measure of implied volatility and the volatility of value-weighted CRSP is used as a measure of realized volatility. The standard errors are computed using the Newey and West (1987) methodology with six lags. The corresponding t-statistics are in square brackets. Significance at the 1%, 5%, and 10% is denoted by $***$, $**$, and $*$, respectively. The momentum return data is obtained from the Jensen-Kelly-Pedersen data repository and covers the period between 01/1986 and 12/2021.

Figure 2.B.1. Mixed Momentum Strategy



The figure above depicts the cumulative returns of the restricted and unrestricted versions of the mixed momentum trading strategy developed in Section 2.5.2.1, along with the returns of the baseline momentum strategy. We obtain the momentum data from the Jensen-Kelly-Pedersen data repository. We consider the time period between 01/1986 and 12/2021.

The difference in performance between the restricted mixed momentum strategy and the baseline momentum strategy is not as stark as the difference in Figure 5. If we invested \$1 in the two strategies in 01/1986, the baseline momentum strategy would generate \$2.94 and the restricted mixed strategy would generate \$3.52. However, the mixed strategy works as intended and lessens the extent of momentum crashes during high-volatility episodes. The unrestricted version of the mixed momentum strategy outperforms both the restricted mixed strategy and the baseline strategy by levering up and taking on relatively more risk during low-volatility periods.

2.C XGBoost

Formally, the XGBoost algorithm involves minimizing the following objective function for the i-th observation at the t-th iteration:

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} \ell\left(y_i, \widehat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega\left(f_t\right) \tag{2.52}$$

by greedily adding tree f_t that most improves the model.

Here $\ell(\cdot)$ represents a loss function that measures the difference between the predicted value of the outcome variable, $\widehat{y}_i$ and the realized value y_i, x_i represents the vector of predictors associated with observation i, and the $\Omega(\cdot)$ function penalizes the complexity of the model.

2.D References

Abarbanell, Jeffery S., and Victor L. Bernard, 1992, Tests of analysts' overreaction/underreaction to earnings information as an explanation for anomalous stock price behavior, *The Journal of Finance* 47, 1181–1207.

Afrouzi, Hassan, Spencer Yongwook Kwon, Augustin Landier, Yueran Ma, and David Thesmar, 2020, Overreaction and working memory, NBER Working Papers 27947, National Bureau of Economic Research.

Andrade, Philippe, and Hervé Le Bihan, 2013, Inattentive professional forecasters, *Journal of Monetary Economics* 60, 967–982.

Andreasen, Martin M, Tom Engsted, Stig V Møller, and Magnus Sander, 2020, The Yield Spread and Bond Return Predictability in Expansions and Recessions, *The Review of Financial Studies* 34, 2773–2812.

Angeletos, George-Marios, Zhen Huo, and Karthik A. Sastry, 2021, Imperfect macroeconomic expectations: Evidence and theory, *NBER Macroeconomics Annual* 35, 1–86.

Auclert, Adrien, Matthew Rognlie, and Ludwig Straub, 2020, Micro jumps, macro humps: Monetary policy and business cycles in an estimated hank model, Working Paper 26647, National Bureau of Economic Research.

Bacchetta, Philippe, and Eric Van Wincoop, 2006, Can information heterogeneity explain the exchange rate determination puzzle?, *American Economic Review* 96, 552–576.

Bacchetta, Philippe, and Eric van Wincoop, 2010, Infrequent portfolio decisions: A solution to the forward discount puzzle, *American Economic Review* 100, 870–904.

Bacchetta, Philippe, and Eric van Wincoop, 2021, Puzzling exchange rate dynamics and delayed portfolio adjustment, *Journal of International Economics* 131, 103460.

Ball, Laurence, N. Gregory Mankiw, and Ricardo Reis, 2005, Monetary policy for inattentive economies, *Journal of Monetary Economics* 52, 703–725, SNB.

Bansal, Ravi, and Magnus Dahlquist, 2000, The forward premium puzzle: different tales from developed and emerging economies, *Journal of International Economics* 51, 115–144.

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

Barroso, Pedro, and Pedro Santa-Clara, 2015, Momentum has its moments, *Journal of Financial Economics* 116, 111–120.

Barroso, Pedro, and Haoxu Wang, 2022, What explains price momentum and 52-week high momentum when they really work?, Working paper.

Ben-Rephael, Azi, Zhi Da, and Ryan D. Israelsen, 2017, It Depends on Where You Search: Institutional Investor Attention and Underreaction to News, *The Review of Financial Studies* 30, 3009–3047.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer, 2022, Belief overreaction and stock market puzzles.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer, 2020, Overreaction in macroeconomic expectations, *American Economic Review* 110, 2748–82.

Borup, Daniel, Jonas N. Eriksen, Mads M. Kjær, and Martin Thyrsgaard, 2023, Predicting bond return predictability, *Management Science* 0, null.

Bouchaud, Jean-Philippe, Philipp Kruger, Augustin Landier, and David Thesmar, 2019, Sticky expectations and the profitability anomaly, *The Journal of Finance* 74, 639–674.

Brooks, Jordan, Michael Katz, and Hanno Lustig, 2018, Post-fomc announcement drift in u.s. bond markets, Working Paper 25127, National Bureau of Economic Research.

Bryzgalova, Svetlana, Sven Lerner, Martin Lettau, and Sven Pelger, 2022, Missing financial data.

Bussiere, Matthieu, Menzie D. Chinn, Laurent Ferrara, and Jonas Heipertz, 2022, The new fama puzzle, Working Paper 24342, National Bureau of Economic Research.

Campbell, John Y., and Robert J. Shiller, 1988, Stock prices, earnings, and expected dividends, *The Journal of Finance* 43, 661–676.

Candian, Giacomo, and Pierre De Leo, 2022, Imperfect exchange rate expectations, Working paper.

Cao, Kai, and Haifeng You, 2021, Fundamental analysis via machine learning, Working paper.

Chauvet, Marcelle, and Jeremy Piger, 2008, A comparison of the real-time performance of business cycle dating methods, *Journal of Business Economic Statistics* 26, 42–49.

Chen, Tianqi, and Carlos Guestrin, 2016, Xgboost: A scalable tree boosting system, *CoRR* abs/1603.02754.

Chernov, Mikhail, and Drew Creal, 2020, The PPP View of Multihorizon Currency Risk Premiums, *The Review of Financial Studies* 34, 2728–2772.

Chordia, Tarun, Avanidhar Subrahmanyam, and Qing Tong, 2014, Have capital market anomalies attenuated in the recent era of high liquidity and trading activity?, *Journal of Accounting and Economics* 58, 41–58.

Clarida, Richard, Josh Davis, and Niels Pedersen, 2009, Currency carry trade regimes: Beyond the fama regression, *Journal of International Money and Finance* 28, 1375–1389, The Global Financial Crisis: Causes, Threats and Opportunities.

Coibion, Olivier, and Yuriy Gorodnichenko, 2015, Information rigidity and the expectations formation process: A simple framework and new facts, *American Economic Review* 105, 2644–78.

Cooper, Michael J., Roberto C. Gutierrez Jr., and Allaudeen Hameed, 2004, Market states and momentum, *The Journal of Finance* 59, 1345–1365.

Da, Zhi, Umit G. Gurun, and Mitch Warachka, 2014, Frog in the Pan: Continuous Information and Momentum, *The Review of Financial Studies* 27, 2171–2218.

Dahlquist, Magnus, and Julien Pénasse, 2022, The missing risk premium in exchange rates, *Journal of Financial Economics* 143, 697–715.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor psychology and security market under- and overreactions, *The Journal of Finance* 53, 1839–1885.

Daniel, Kent, and Tobias J. Moskowitz, 2016, Momentum crashes, *Journal of Financial Economics* 122, 221–247.

de Silva, Tim, and David Thesmar, 2023, Noise in expectations: Evidence from analyst forecasts, Working paper.

Dellavigna, Stefano, and Joshua M. Pollet, 2009, Investor inattention and friday earnings announcements, *The Journal of Finance* 64, 709–749.

Diether, Karl B., Christopher J. Malloy, and Anna Scherbina, 2002, Differences of opinion and the cross section of stock returns, *The Journal of Finance* 57, 2113–2141.

Driscoll, John C., and Aart C. Kraay, 1998, Consistent covariance matrix estimation with spatially dependent panel data, *The Review of Economics and Statistics* 80, 549–560.

Du, Wenxin, Alexander Tepper, and Adrien Verdelhan, 2018, Deviations from covered interest rate parity, *The Journal of Finance* 73, 915–957.

Engel, Charles, 2014, Exchange rates and interest parity, in Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff, eds., *Handbook of International Economics*, volume 4 of *Handbook of International Economics*, 453–522 (Elsevier).

Engel, Charles, Ekaterina Kazakova, Mengqi Wang, and Nan Xiang, 2021, A reconsideration of the failure of uncovered interest parity for the u.s. dollar, Working Paper 28420, National Bureau of Economic Research.

Engelberg, Joseph, R. David McLean, and Jeffrey Pontiff, 2018, Anomalies and news, *The Journal of Finance* 73, 1971–2001.

Fama, Eugene F., 1984, Forward and spot exchange rates, *Journal of Monetary Economics* 14, 319–338.

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

Flynn, Joel P., and Karthik Sastry, 2021, Attention cycles, Working paper.

Frankel, Jeffrey A., and Kenneth A. Froot, 1987, Using survey data to test standard propositions regarding exchange rate expectations, *The American Economic Review* 77, 133–153.

Frankel, Richard, and Charles M.C. Lee, 1998, Accounting valuation, market expectation, and cross-sectional stock returns, *Journal of Accounting and Economics* 25, 283–319.

Froot, Kenneth A., and Richard H. Thaler, 1990, Anomalies: Foreign exchange, *Journal of Economic Perspectives* 4, 179–192.

Gabaix, Xavier, and Matteo Maggiori, 2015, International Liquidity and Exchange Rate Dynamics, *The Quarterly Journal of Economics* 130, 1369–1420.

Gârleanu, Nicolae, and Stavros Panageas, 2015, Young, old, conservative, and bold: The implications of heterogeneity and finite lives for asset pricing, *Journal of Political Economy* 123, 670–685.

Goulding, Christian L., Campbell R. Harvey, and Michele Mazzoleni, 2022, Momentum turning points, Working paper.

Gourinchas, Pierre-Olivier, and Aaron Tornell, 2004, Exchange rate puzzles and distorted beliefs, *Journal of International Economics* 64, 303–333.

Gómez-Cram, Roberto, 2022, Late to recessions: Stocks and the business cycle, *The Journal of Finance* 77, 923–966.

Haddad, Valentin, and Tyler Muir, 2021, Do intermediaries matter for aggregate asset prices?, *The Journal of Finance* 76, 2719–2761.

Hansen, Lars Peter, and Robert J. Hodrick, 1980, Forward exchange rates as optimal predictors of future spot rates: An econometric analysis, *Journal of Political Economy* 88, 829–853.

Harvey, Campbell R., Yan Liu, and Heqing Zhu, 2015, . . . and the Cross-Section of Expected Returns, *The Review of Financial Studies* 29, 5–68.

He, Zhiguo, Bryan Kelly, and Asaf Manela, 2017, Intermediary asset pricing: New evidence from many asset classes, *Journal of Financial Economics* 126, 1–35.

Hirshleifer, David, Sonya Seongyeon Lim, and Siew Hong Teoh, 2009, Driven to distraction: Extraneous events and underreaction to earnings news, *The Journal of Finance* 64, 2289–2325.

Hodrick, Robert J., and Edward C. Prescott, 1997, Postwar u.s. business cycles: An empirical investigation, *Journal of Money, Credit and Banking* 29, 1–16.

Hong, Harrison, Terence Lim, and Jeremy C. Stein, 2000, Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies, *The Journal of Finance* 55, 265–295.

Hong, Harrison, and Jeremy C. Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets, *The Journal of Finance* 54, 2143–2184.

Hou, Kewei, Chen Xue, and Lu Zhang, 2014, Digesting Anomalies: An Investment Approach, *The Review of Financial Studies* 28, 650–705.

Hribar, Paul, and Nicole Thorne Jenking, 2004, The effect of accounting restatements on earnings revisions and the estimated cost of capital, *Review of Accounting Studies* 9, 337–356.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *The Journal of Finance* 48, 65–91.

Jensen, Theis Ingerslev, Bryan T Kelly, and Lasse Heje Pedersen, 2022, Is there a replication crisis in finance?, *Journal of Finance, Forthcoming* .

Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp, 2016, A rational theory of mutual funds' attention allocation, *Econometrica* 84, 571–626.

Kalemli-Özcan, Sebnem, and Liliana Varela, 2021, Five facts about the uip premium, Working Paper 28923, National Bureau of Economic Research.

Kohlhas, Alexandre N., and Ansgar Walther, 2021, Asymmetric attention, *American Economic Review* 111, 2879–2925.

La Porta, Rafael, 1996, Expectations and the cross-section of stock returns, *The Journal of Finance* 51, 1715–1742.

Londono, Juan M., and Hao Zhou, 2017, Variance risk premiums and the forward premium puzzle, *Journal of Financial Economics* 124, 415–440.

Loungani, Prakash, Herman Stekler, and Natalia Tamirisa, 2013, Information rigidity in growth forecasts: Some cross-country evidence, *International Journal of Forecasting* 29, 605–621.

Luo, Jiang, Avanidhar Subrahmanyam, and Sheridan Titman, 2020, Momentum and Reversals When Overconfident Investors Underestimate Their Competition, *The Review of Financial Studies* 34, 351–393.

Mackowiak, Bartosz, and Mirko Wiederholt, 2009, Optimal sticky prices under rational inattention, *American Economic Review* 99, 769–803.

Mankiw, N. Gregory, and Ricardo Reis, 2002, Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve, *The Quarterly Journal of Economics* 117, 1295–1328.

Martineau, Charles, 2023, Rest in peace post-earnings announcement drift, *Critical Finance Review* .

Meier, Stephan, and Charles D. Sprenger, 2015, Temporal stability of time preferences, *The Review of Economics and Statistics* 97, 273–286.

Mincer, Jacob, and Victor Zarnowitz, 1969, The evaluation of economic forecasts, in *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, 3–46 (National Bureau of Economic Research, Inc).

Mody, Ashoka, Damiano Sandri, and Franziska L. Ohnsorge, 2012, Precautionary savings in the great recession, Working Paper 2012/04, IMF.

Molavi, Pooya, Alireza Tahbaz-Salehi, and Andrea Vedolin, 2021, Model complexity, expectations, and asset prices, Working paper.

Moreira, Alan, and Tyler Muir, 2017, Volatility-managed portfolios, *The Journal of Finance* 72, 1611–1644.

Moskowitz, Tobias J., Yao Hua Ooi, and Lasse Heje Pedersen, 2012, Time series momentum, *Journal of Financial Economics* 104, 228–250, Special Issue on Investor Sentiment.

Mäkinen, Taneli, and Björn Ohl, 2015, Information acquisition and learning from prices over the business cycle, *Journal of Economic Theory* 158, 585–633, Symposium on Information, Coordination, and Market Frictions.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Novy-Marx, Robert, 2013, The other side of value: The gross profitability premium, *Journal of Financial Economics* 108, 1–28.

Opp, Christian C.G., 2015, Intertemporal information acquisition and investment dynamics, Working paper.

Palmrose, Zoe-Vonna, Vernon J. Richardson, and Susan Scholz, 2004, Determinants of market reactions to restatement announcements, *Journal of Accounting and Economics* 37, 59–89.

Pavlova, Anna, and Roberto Rigobon, 2007, Asset Prices and Exchange Rates, *The Review of Financial Studies* 20, 1139–1180.

Pouget, Sebastien, Julien Sauvagnat, and Stephane Villeneuve, 2016, A Mind Is a Terrible Thing to Change: Confirmatory Bias in Financial Markets, *The Review of Financial Studies* 30, 2066–2109.

Reis, Ricardo, 2006, Inattentive Producers, *The Review of Economic Studies* 73, 793–821.

Schneider, Andres, 2022, Risk-sharing and the term structure of interest rates, *The Journal of Finance* 77, 2331–2374.

So, Eric C., 2013, A new approach to predicting analyst forecast errors: Do investors overweight analyst forecasts?, *Journal of Financial Economics* 108, 615–640.

Teräsvirta, Timo, 1994, Specification, estimation, and evaluation of smooth transition autoregressive models, *Journal of the American Statistical Association* 89, 208–218.

Valente, Joao Paulo, Kaushik Vasudevan, and Tianhao Wu, 2022, The role of beliefs in asset prices: Evidence from exchange rates, Working paper.

Van Binsbergen, Jules H, Xiao Han, and Alejandro Lopez-Lira, 2022, Man versus Machine Learning: The Term Structure of Earnings Expectations and Conditional Biases, *The Review of Financial Studies* .

Verdelhan, Adrien, 2010, A habit-based explanation of the exchange rate risk premium, *The Journal of Finance* 65, 123–146.

White, Halbert, 1980, A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica* 48, 817–838.