Overreaction and Forecast Horizon: Longer-term Expectations Overreact More, Shorter-term Expectations Drive Fluctuations*

Basil Halperin
University of Virginia

J. Zachary Mazlish
Oxford and GPI

August 2025

Abstract

We use survey data on macroeconomic expectations, across 89 countries and going back to 1989, to establish four facts about how forecast biases depend on the time horizon of the forecast. The data cover average expectations and horizons from 0 to 10 years. (1) Expectations underreact at a horizon of one year or less. (2) Expectations overreact at horizons of two years or more. (3) Expectations are "too extreme" at all horizons. (4) Overreaction and over-extremity increase with forecast horizon. These four patterns hold across advanced and emerging economies, and across multiple macroeconomic variables. They are inconsistent with several popular models of overreaction, where the degree of overreaction is independent of forecast horizon. However, we show that a model featuring costly recall, uncertainty about the long-run mean, and sticky-information can match all four of our facts. Finally, although long-term expectations exhibit stronger overreaction, it is short-term expectations that are most strongly associated with fluctuations in GDP, investment, and the stock market.

*_Email:_ basilh@virginia.edu, john.mazlish@economics.ox.ac.uk
We thank Alex Kohlhas, Sergio de Ferra, and participants at the 2025 NBER Summer Institute for helpful comments. We are grateful for generous financial support to the George and Obie Shultz Fund at MIT and to the Institute for Humane Studies under grant number 017435.
First version: January 2025.

1 Introduction

A large empirical literature on macroeconomic expectations consistently documents deviations from rational expectations. But how do these deviations vary *by the time horizon of the forecast*?

In this paper, we document four facts about expectation biases and how they change with forecast horizon. Our results are based on average survey macroeconomic expectations from 89 countries, going back as far as 1989, and are thus robust across a diverse range of macroeconomic environments:

(1) Expectations *under*-revise in the short run (at the one-year horizon or less).

(2) Expectations *over*-revise in the long run (at the two-year horizon or more).

(3) At all horizons, expectations tend to be too extreme: unusually high forecasts are too high, and unusually low forecasts are too low.

(4) Over-revision and over-extremity both increase in forecast horizon.

These facts hold for several core macro variables: expectations of GDP growth, investment growth, consumption growth, and inflation.

We distinguish two dimensions of overreaction: "over-revision" and "over-extremity". Over-revision refers to the case when an *update* in your forecast causes you to update your forecast *too much*. This is the sense of over-reaction introduced by Coibion and Gorodnichenko (2015), and when it is clear we will sometimes refer to this simply as "overreaction". Formally, this is measured by regressing forecast errors on forecast revisions. "Over-extremity" instead captures the notion that when your forecasts are higher than normal, they are typically too high; and when your forecasts are lower than normal, they are typically too low. Formally, this is captured by regressing forecast errors on the lagged forecast level, as in Bordalo, Gennaioli, La Porta, and Shleifer (2024).

Our key finding – that both dimensions of overreaction are increasing in forecast horizon – is inconsistent with several popular models of expectation formation. We show that a model of costly recall (Afrouzi, Kwon, Landier, Ma, and Thesmar 2023) extended to include sticky information can match all four facts.

Beyond documenting the term structure of forecaster biases, we ask how these different biases are related to economic fluctuations. A recent literature has found that, in the US, *long-run* expectations are most strongly associated with subsequent real and financial outcomes at the business cycle frequency (Bordalo, Gennaioli, La Porta, and Shleifer 2024; Bordalo, Gennaioli, La Porta, O'Brien, and Shleifer 2024). In contrast to this literature, we find that – in our broad cross-country sample – *short-term* expectations are most predictive of outcomes at the business cycle frequency.

We now detail our main empirical findings and situate them relative to the literature.

Fact 1: Expectations under-revise at horizons of one year or less. This fact was first documented by Coibion and Gorodnichenko (2015), who made use of the same Consensus Economics data source, but for a subset of 12 advanced countries and with forecasts going out less than two years.

Fact 2: Expectations over-revise at horizons of two years or more. With our data on forecasts extending past the two-year horizon to the ten-year horizon, we show that the Coibion-Gorodnichenko fact flips to *over*-revision at horizons of two years or longer. This finding on average expectations also resonates with Bordalo, Gennaioli, Ma, and Shleifer (2020), who find over-revision using *individual*-level short-term forecasts; and Bordalo, Gennaioli, La Porta, and Shleifer (2024), who find over-revision in three-to-five-year equity earnings growth average forecasts.

Fact 3: Expectations are "too extreme" at all horizons. We find that forecasts tend to be too extreme at all forecast horizons from zero to ten years. As far as we are aware, this fact is novel. Related work includes Kohlhas and Walther (2021); as well as Bordalo, Gennaioli, La Porta, and Shleifer (2024) who show that the five-year-ahead forecasts of equity analysts are too extreme in the same sense but who do not study the result at different forecast horizons.

Fact 4: Expectations overreact more at longer horizons, in both senses – over-revising more and being too extreme. Extending fact 1 and fact 2, we find that overreaction increases smoothly with the time horizon of forecasts. Versions of this fact are found in a number of recent papers (Bordalo, Gennaioli, La Porta, and Shleifer 2024; Angeletos, Huo, and Sastry 2021; d'Arienzo 2020; Afrouzi, Kwon, Landier, Ma, and Thesmar 2023; Fisher, Melosi, and Rast 2025; and Adam, Pfäuti, and Reinelt 2025). To our knowledge, we are the first to show the pattern extends beyond the two-year horizon for primary macroeconomic variables[1], and the first to compare advanced versus emerging economies.

Robustness. The primary contribution of our paper is establishing that these four facts hold across four different macroeconomic variables – GDP growth, inflation, investment growth, and consumption growth – and across a wide range of macroeconomic environments. The basic pattern of the four facts emerges even if we split the sample into advanced and emerging market economies, split the sample in half across time, or remove any forecast covering the 2008 financial crisis. Furthermore, we show that using our forecasts helps with out-of-sample forecasting, and helps more the longer is the forecast horizon. This stands in contrast to the findings of Eva and Winkler (2023) about forecasts with a horizon below one year, and instills further confidence that the four facts we focus on are robust features of macroeconomic expectations across different regimes and information settings.

[1] An exception is the contemporaneous work of Bonaglia, d'Arienzo, Fallico, Gennaioli, and Iovino (2025), who study expectations up to the 10-year horizon, for inflation and in advanced economies.

A model that fits the facts. The headline result that overreaction is smoothly increasing in horizon is inconsistent with the models found in Bordalo, Gennaioli, La Porta, and Shleifer (2024) and Angeletos, Huo, and Sastry (2021), where overreaction increases at medium-horizons, but then decreases back towards zero at longer horizons. That is, there is no overreaction in beliefs infinitely far out into the future.

We show a model of costly recall – calibrated to match parameters measured in existing lab results – augmented with costly information can match the facts. We build on the costly recall framework of Afrouzi, Kwon, Landier, Ma, and Thesmar (2023), where beliefs about the *long-run mean* overreact, to generate overreaction that is increase in horizon. To match the under-revision seen in the data at short horizons, we augment the model with sticky information. To match the evidence that there is over-extremity even at short horizons, *sticky* information rather than *noisy* information is required.[2]

Expectations by horizon, aggregate fluctuations, and equity markets. In addition to documenting the above stylized facts, we also investigate which expectations are associated with subsequent macroeconomic and financial outcomes. We first run local projections which control for a host of lagged macroeconomic variables to identify the effects of changing macroeconomic expectations on investment and GDP growth over the business cycle. We find evidence consistent with Bordalo, Gennaioli, La Porta, O'Brien, and Shleifer (2024) that upward revisions in expectations are associated with short-term (less-than-two-year) "booms" and sharp "busts" right afterwards. However, unlike in Bordalo, Gennaioli, La Porta, O'Brien, and Shleifer (2024), it is changes in *short-term* expectations that are most strongly associated with ensuing booms-and-busts. The pattern of the expansion and reversal of real activity matches the typical business cycle characteristics found in Angeletos, Collard, and Dellas (2020), and the fact that such a business cycle arises in response to under- then over-reacting expectations fits with recent work such as Angeletos, Collard, and Dellas (2018), Bianchi, Ilut, and Saijo (2024a), L'Huillier, Singh, and Yoo (2024), Cai (2024), and Bardóczy and Guerreiro (2023).

Next, we show that high GDP growth expectations predict weak stock market returns. Similar to our results about the real economy, we find that short-term growth expectations are better predictors of up to five-year ahead stock market return reversals than high long-term expectations, in contrast to the opposite findings of Bordalo, Gennaioli, La Porta, and Shleifer (2024). Our results can be reconciled with theirs since in our *US data* we match their finding that long-term expectations are better predictors of weak returns. The finding that belief overreaction is important for explaining stock market movements fits within a long tradition, including important recent contributions such as

[2]A number of other recent papers have developed models *consistent with* overreaction increasing in horizon, including Sung (2024), Bianchi, Ilut, and Saijo (2024b), and Farmer, Nakamura, and Steinsson (2024). Other recent work that provides cognitive foundations for the joint existence of over- and under-reaction depending on the forecast setting includes Augenblick, Lazarus, and Thaler (2025) and Ba, Bohren, and Imas (2024). The appendix of Molavi, Tahbaz-Salehi, and Vedolin (2024) shows that if agents believe the data to follow an AR(1) process when the true process is ARMA(1,1), then there can be joint over- and under-reaction.

Bianchi, Ludvigson, and Ma (2024) and McCarthy and Hillenbrand (2021). The specific boom-bust pattern we find in response to positive shocks matches the model implications in Mei and Wu (2024), and connects to other work on how irrational expectations jointly explain stock market and business cycle outcomes, such as Adam, Marcet, and Beutel (2017) and Winkler (2020). Finally, these findings also relate to recent work on the properties of long-horizon earnings expectations (Sias, Starks, and Turtle 2024, H Décaire and Guenzel 2023), and could explain the premium on near-future cash-flows found by Gormsen and Lazarus (2023).

These findings about the importance of short-term expectations pose some tension with our earlier findings that it is long-term expectations which overreact most. The discrepancy is not necessarily a puzzle, since different models predict substantially different effects of long-term expectations on current activity (Beaudry and Portier 2014; Dupor and Mehkari 2014). Further, our results are consistent with models which dampen the importance of long-term expectations relative to short-term expectations in agents' decision-making, such as Angeletos and Lian (2018) and Angeletos and Huo (2021).

Outline. The rest of the paper proceeds as follows. Section 2 introduces the data. Section 3 establishes our four facts about average macroeconomic expectations. Section 4 shows that a costly-recall model with sticky-information matches our empirical facts. Section 5 shows that short-term expectations are more associated with "boom-bust" cycles in the macroeconomy than long-term expectations. Section 6 shows the same pattern for stock-return predictability. Section 7 briefly concludes.

2 Data and Variable Construction

2.1 Survey Data

We use survey data from Consensus Economics. Consensus Economics surveys professional forecasters working at banks, consultancies, and other firms around the world to elicit a host of different macroeconomic forecasts at the country level.

The data covers 89 countries, with sample lengths varying by country and with the longest series extending back to 1989 for the G7 countries. The variety of macroeconomic environments across our sample of economies provides a substantial advantage for our analysis compared to the bulk of the literature, which focuses on the Survey of Professional Forecasters in the United States and a few other marquee surveys. Prior to 2014, Consensus conducted its survey twice per year – at the beginning of April (the start of Q2) and the beginning of October (the start of Q4). Since 2014, the survey has been conducted quarterly.

We make use of Consensus' "long-term" forecasts, which include an estimate at each annual horizon from zero to five years as well as an average forecast across years six through ten. As a concrete example, the April 2025 survey would ask about expected

GDP growth for each year from 2025 to 2030, as well as average annual GDP growth across 2031 through 2035.

The data is only available at the aggregate level: we only have access to the average forecast and the standard deviation of forecasts, but not individual-level forecasts. This is an important limitation for the reasons discussed in Angeletos, Huo, and Sastry (2021).

Consensus Economics collects forecasts on a host of variables that differ in availability by country (and over time). Our analysis focuses on four variables: GDP growth, inflation, consumption growth, and investment growth.

In total, we have 4240 observations of 0-10 year GDP forecasts, 4205 observations of inflation forecasts, 3185 observations of consumption forecasts, and 3185 observations of investment forecasts. Appendix Table 6 provides a full list of the 89 countries in the dataset and describes data availability by country.

The seminal work of Coibion and Gorodnichenko (2015) also used data from Consensus Economics. That paper made use of the benchmark data product from Consensus Economics, which consists of quarterly forecasts out to a two-year horizon. We make use of Consensus' "long-term" forecasts, which are annual and cover the zero-to-ten-year horizon. Thus, the long-term forecasts are necessary for our focus on how bias varies across short-horizon versus long-horizon forecasts. In addition, the work of Coibion and Gorodnichenko (2015) used data covering 12 countries, from the G7 and western Europe; we use data covering 89 countries from across the world, so that our forecasts span a wide range of macroeconomic environments.

To measure realized outcomes for each variable, we use the World Bank's World Development Indicators (WDI) database.

2.2 Constructing Forecast Errors

For macroeconomic variable x in country c, let $\mathbb{E}_t\left(x_{c,t+h}\right)$ denote the time-t forecast for the variable at time $t+h$. Define the corresponding forecast error for that forecast as the forecast minus the realized value:

$$e_{c,t+h} \equiv x_{c,t+h} - \mathbb{E}_t\left(x_{c,t+h}\right)$$

Forecasts x are z-scored with respect to macro variable, country, and forecast horizon, where the z-score is taken with an expanding window.

Why z-score forecasts? One reason is that z-scoring prevents countries or variables with greater volatility in forecast errors from dominating the results.[3]

For example, consider if the data consisted exclusively of inflation forecasts for Venezuela and Switzerland. Inflation expectations in Venezuela have average forecast errors that are *far* more variable than in Switzerland, leading Venezuela to dominate the variance in the

[3]As we will see, it is also important to z-score to prevent the volatility of forecast *revisions* from one country or variable from dominating the results.

data and be correspondingly "overweighted" in regressions. Adding country fixed effects would not alleviate this issue, since the fixed effect would only adjust for the *average* forecast error in Venezuelan, but not the fact that the variability of Venezuelan forecast errors swamps that of Switzerland.[4]

A second advantage of z-scoring is that the average forecast error in each country is automatically set to zero, which separates out underreaction or overreaction driven by a bias in the *mean* forecast. For example, consider US short-term interest rate forecasts. For most of the period from 1981 to 2019, forecasters consistently predicted that the short-term interest rate would rise higher than it actually did. Since interest rates mostly moved down over this period, this looks like a case of underreaction (forecasts revised down, but not as much as they should have). However, as shown in Farmer, Nakamura, and Steinsson (2024), this pattern of forecast errors can stem from having a misspecified prior belief about the mean of the interest rate process and slow learning – rather than any systematic tendency to underreact. Z-scoring variable separates out this issue by setting the mean of each series to zero.

We use a z-score with an expanding window to prevent a lookahead bias. That is, forecasts are z-scored with respect to *only* the observations up to that point in time.[5]

3 Measuring Forecast Biases

3.1 Testing for Overreaction

Our primary test for overreaction follows the approach in Bordalo, Gennaioli, La Porta, and Shleifer (2024), generalized to a panel setting as we study forecasts across multiple countries. We estimate panel regressions of the following form:

$$e_{c,t+h} = \alpha + \beta_1 \Delta \mathbb{E}_t(x_{c,t+h}) + \beta_2 \mathbb{E}_{t-1}(x_{c,t+h}) + f_{c,x} + \epsilon_{c,t} \tag{1}$$

We pool across countries c and our four outcome variables x (GDP growth, inflation, consumption growth, and investment growth) to maximize power. We will later show robustness of results by individual country and outcome variable.

The first regressor, $\Delta \mathbb{E}_t(x_{c,t+h})$, is the *revision in expectations*. We can consider the revision over different lag lengths – e.g. the revision over the past 12 months – but we focus on the revision since the most recent prior survey. The second regressor, $\mathbb{E}_{t-1}(x_{c,t+h})$, is the *lagged forecast* (again from the most recent prior survey).

Finding that either β_1 or β_2 is statistically distinguishable from zero rejects full-information rational expectations (FIRE): since both regressors – forecast revision and lagged forecast – are known to forecasters at time t, neither variable should systematically predict fore-

[4]Similarly, in our baseline analysis we will pool all variables; and for Venezuela, forecast errors for inflation are far more variable than forecast errors for say GDP. Z-scoring allows the two to be comparable.

[5]We drop any observations where there are not at least five previous forecasts to base the z-score on.

cast error under FIRE.[6] In particular, for both coefficients, if β_i is negative, this indicates that the average forecast overreacts; and conversely if β_i is positive then the forecast underreacts. We discuss the interpretation for each β_i in turn.

Over-revision vs. under-revision. The intuition for why $\beta_1 < 0$ indicates overreaction is the following. $\beta_1 < 0$ implies that a positive forecast revision is associated with a negative forecast error. Therefore, the forecast should not have been revised up as much as it was, since the forecast is above the realization (on average) – thus, the revision was an overreaction. Testing for over or underreaction by focusing on the coefficient on forecast revisions was introduced by Coibion and Gorodnichenko (2015). When β_1 is negative, we term this "over-revision" (and correspondingly $\beta_1 > 0$ as "under-revision").

Note that the interpretation of the coefficients needs to account for the fact that we z-score expectations with respect to country and horizon: the β_i coefficients should be interpreted as "relative to what is typical". So, the interpretation of β_1 is: "when expectations revise a large amount *relative to what is typical* for horizon h and country c, are forecast errors large or small *relative to what is typical* for that horizon h and country c?". This corresponds well with an intuitive sense of what a 'systematic' tendency for over or underreaction would mean.

Over-extremity vs. under-extremity. The intuition for why $\beta_2 < 0$ indicates overreaction is the following. $\beta_2 < 0$ implies that a higher *level* of the lagged forecast is associated with the forecast being too high relative to the outcome. This means that high lagged forecasts are too high while low forecasts are too low. Thus we term $\beta_2 < 0$ as "over-extremity" (and correspondingly $\beta_2 > 0$ as "under-extremity").

Why use two senses of overreaction, rather than running separate regressions? There are two answers. One is that each sense captures a different, relevant sense of overreaction. The other answer is that including one variable but not the other would result in an omitted variable bias, since the two variables are correlated.[7]

Fixed effects and standard errors. Variable-by-country fixed effects are denoted $f_{c,x}$, and we show below that results are robust to including time fixed effects and other variations. Given that we have a panel with longitudinal and cross-sectional error correlation, we use Driscoll and Kraay (1998) standard errors, with groupings by both variable and country.

[6]As discussed extensively in Angeletos, Huo, and Sastry (2021), this *aggregate* analysis does not distinguish between a representative irrational agent who systematically underreacts or overreacts to news, versus a population of rational but heterogeneously informed agents.

[7]Appendix figure 10 shows this, pooling across all variables; figure 11 shows the same for each of our four outcome variables. In particular, at the 0-year horizon, there is significant momentum in current-year forecasts: a high-level of the forecast yesterday predicts upward revisions today. Further, at the two-year horizon and beyond, there is significant reversal: a higher value of the forecast predicts a revision downward in forecasts (a sort of "mean-reversion" in forecasts).

3.2 Results

Figure 1 and table 1 present our main results: the estimates from regression (1), for each horizon, pooled across countries and outcome variables.[8] All four facts discussed in the introduction appear in the results:

(i) **Under-revision at short-term horizons.** Fact 1 is that less than one-year forecasts under-revise, which is shown by the significantly positive β_1 coefficient. This is the same fact established by Coibion and Gorodnichenko (2015).

(ii) **Over-revision at long-term horizons.** Fact 2 is that two-year and longer horizon forecasts *over*-revise, as shown by the significantly negative β_1 coefficients at horizons two and beyond.[9]

(iii) **Over-extremity at all horizons.** Fact 3 is that forecasts are too extreme at every horizon, as shown by the fact that β_2 is everywhere significantly less than zero.

(iv) **Over-revision and over-extremity both increasing in horizon.** Fact 4 is that both senses of overreaction are increasing in forecast horizon, as shown by the monotonically decreasing β_1 and β_2 coefficients. The degree of overreaction of six-to-ten year forecasts is more than double that of five-year forecasts, for both senses of overreaction.

Observe that our large sample size allows for substantially tighter estimates than much or all of the literature on macroeconomic forecasts.

[8] Horizons indicate "years ahead", not survey date: recall that the available horizons are annual for years zero to five, and then an average across years six through ten.

[9] That β_1 coefficients flip from under to over-reaction between the one-to-two year horizon (facts 1 and 2) is consistent with the contemporaneous findings of Del Negro (2024) that professional forecasters flip from being under-confident to over-confident at greater than one-year horizons.

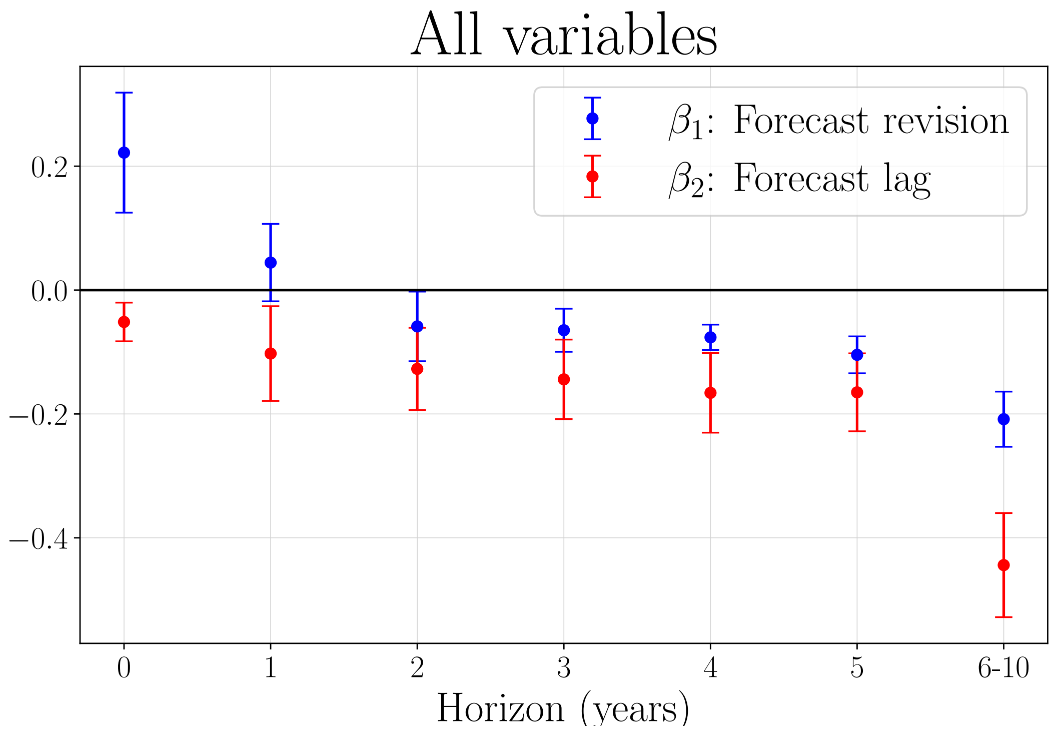


Figure 1: The figure plots the blue β_1 and red β_2 coefficients from the regression $e_{c,t+h} = \alpha + \beta_1 \Delta\, \mathbb{E}_t(x_{c,t+h}) + \beta_2\, \mathbb{E}_{t-1}(x_{c,t+h}) + f_{c,x} + \epsilon_{c,t}$, where $e_{c,t+h}$ are forecast errors, $\Delta\, \mathbb{E}_t(x_{c,t+h})$ are consecutive survey forecast revisions, and $\mathbb{E}_{t-1}(x_{c,t+h})$ is the previous survey's forecast. All forecasts are pooled and are z-scored with respect to variable, country, and horizon, with expanding window z-scores for the lagged forecast. Covid is removed from the sample. Country-variable fixed effects are included and standard errors are Driscoll-Kraay with country-variable groupings.

Table 1: Overreaction by Horizon: Pooling across all variables and countries

Horizon	0	1	2	3	4	5	6-10
Revision	0.16***	0.04	-0.06**	-0.06***	-0.08***	-0.10***	-0.21***
	(0.03)	(0.03)	(0.03)	(0.02)	(0.01)	(0.02)	(0.02)
Lag	-0.06***	-0.10***	-0.13***	-0.14***	-0.17***	-0.17***	-0.44***
	(0.02)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
Observations	8513	7776	7074	6400	5735	5067	3221

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

The fact that forecasts are z-scored also allows for a natural interpretation of the magnitudes of the coefficients. Consider, for example, the point estimate of $\beta_1 = -0.21$ for the six-to-ten-year forecast horizon. This implies that a positive *one standard deviation revision* in six-to-ten-year expectations tends to produce a positive *0.21 standard deviation forecast error*: the forecast will end up being 0.21 standard deviations above the realized outcome.

What does it mean when one coefficient is positive while the other is negative? At the 0-year horizon, forecast errors are positively predicted by forecast revisions ($\beta_1 = 0.16$) but negatively predicted by the lagged forecast value ($\beta_2 = -0.06$), both with statistical significance. The interpretation is that forecasts *under*-revise but are *overly* extreme. That is, when the average forecast updates, it tends to not update enough. But when the forecast is high relative to its typical value (up to that point in time), it tends to be too high. This reflects that there is not one, single definition of "overreaction" in the literature: there are multiple senses in which a forecast a can under or overreact, and each coefficient reflects a different sense.

3.3 Robustness

Results by forecast variable. As just described, the main specification pools together the four outcome variables (GDP growth, inflation, consumption growth, and investment growth) to maximize precision. We now show that examining each variable individually, the pattern of results is broadly very similar to that of the pooled results.

Figure 2 shows the results for each forecast variable individually. The four facts are robustly supported:

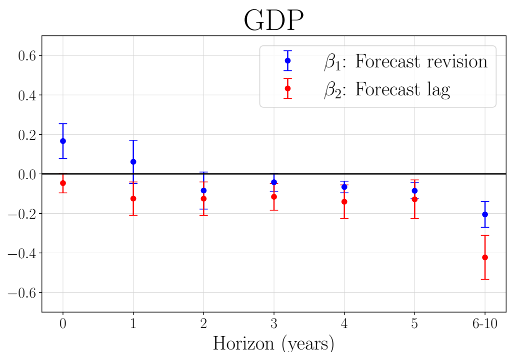
(i) **Short-term under-revision.** For every variable, β_1 is positive at short horizons, indicating under-revision, and is significant for every variable other than consumption.

(ii) **Long-term over-revision.** For horizon 2 and longer, β_1 is always negative and is significant for all but three observations, for which it is very close to statistically significant (horizons 2 and 3 for GDP; horizon 2 for inflation).

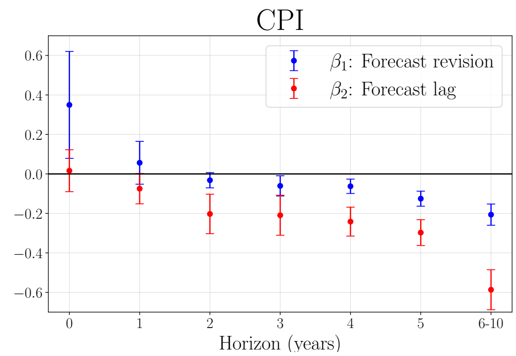
(iii) **Over-extremity.** All point estimates of β_2 are negative, and all but three of those estimates are significant (inflation at horizons one and two, consumption at horizon two).

(iv) **Over-revision and over-extremity increasing in horizon.** Both β_1 and β_2 have a downward trend for each variable. The coefficients are no longer decreasing perfectly monotonically, but overreaction clearly tends to increase in horizon, and six-to-ten horizon forecasts exhibit the greatest degree of overreaction for both β_1 and β_2, for every variable.

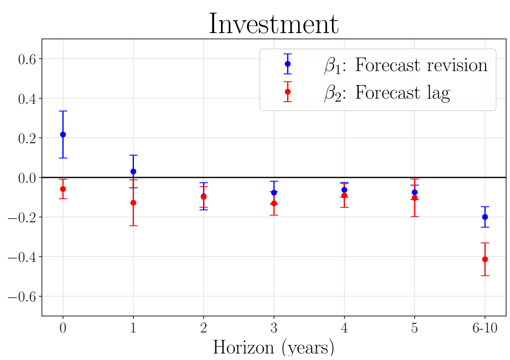
As the figures evince, the patterns of how under and overreaction change with forecast horizon are consistent across variables. This helps justify the choice of pooling in our primary specification, and it supports the interpretation that facts 1 through 4 are general patterns of forecaster over (and under) reaction – not patterns which are variable-specific, at least in our sample.



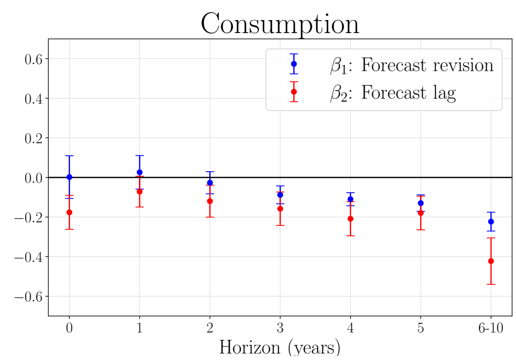
(a) GDP



(b) Inflation



(c) Investment



(d) Consumption

Figure 2: Regression coefficients from $e_{c,t+h} = \alpha + \beta_1 \Delta \mathbb{E}_t[x_{c,t+h}] + \beta_2 \mathbb{E}_{t-1}[x_{c,t+h}] + f_{c,x} + \varepsilon_{c,t}$ for different forecast variables x. Country-variable fixed effects are included and standard errors are Driscoll-Kraay with country-variable groupings.

Time splits. Figure 3 shows that the facts are robust to different ways of splitting the sample along the time dimension.

The first two subplots, figures 3a and 3b, show all four facts hold after splitting the sample into "halves" by forecast horizon, so that each half has an equal number of observations. The date of the sample split varies by horizon.[10]

In cases where certain coefficients don't reach statistical significance, they approach significance thresholds, suggesting the issue is reduced statistical power in the smaller samples rather than absence of the effect. While the first half of the sample shows a slightly less pronounced monotonic decrease in coefficients, both subsamples still display a clear downward trend in both β_1 and β_2, with maximum overreaction occurring at the six-to-ten year horizon.

The third subplot, figure 3c, conservatively removes anything from the sample that may have been affected by the 2008 financial crisis by taking out *all* observations where the forecast horizon spans 2008. That is, all forecasts made *before* 2008 *for the year of 2008 or after* are dropped. For longer horizon forecasts, this reduces the sample quite a bit (especially considering we also drop samples containing Covid years): for instance, the six-to-ten year forecast sample falls from 3223 observations to 793. Despite the stringency of this test, all four facts continue to hold strongly.[11]

Country splits. Figure 4 splits the sample into 23 advanced economies and 63 emerging markets (all non-advanced economies in the sample) to show that the pattern of results is not specific to forecasting certain types of economies. In both samples, the four facts hold exactly. Once again, the downward trend in β_2 coefficients is only evident once the six-to-ten year horizon is included, but the six-to-ten year forecasts clearly exhibit the most over-extremity. Appendix figure 16 shows the distribution of coefficients across individual countries.

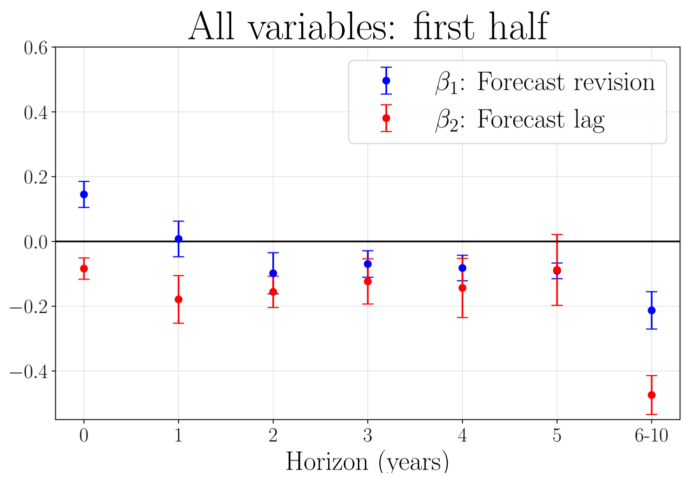
Variations in fixed effects. Figure 5 shows that each of the four facts holds across four different ways of including fixed effects in regression (1). The top-left figure 5a is a repeat of our main specification, which uses group fixed effects, where the groups are country-by-variable.

The top-right figure 5b drops the country-by-variable fixed effects and uses time fixed effects, where the time periods are survey dates (i.e. quarters). The bottom-left figure 5c includes both country-by-variable and time fixed effects. The bottom-right figure 5d uses no fixed effects.

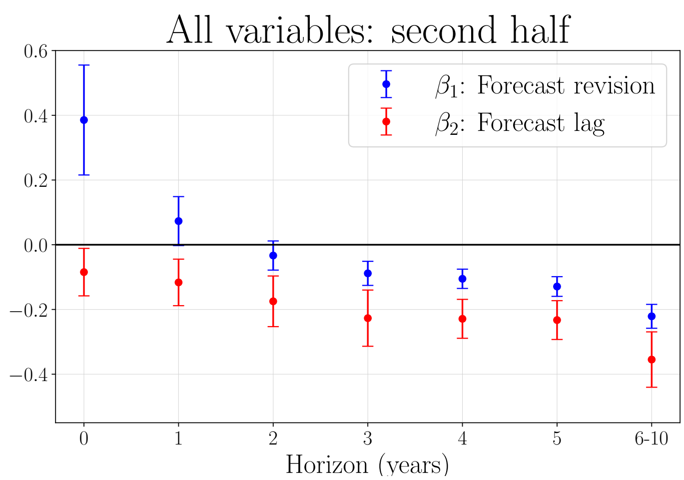
The results are broadly unchanged across all four different specifications. The biggest difference across results is that the two specifications which use time fixed effects find slightly higher β_2 coefficients at short horizons (though the point estimate remains ev-

[10] Appendix figures 15a and 15b show the results are robust to instead splitting the sample by a fixed date (July 2007) and allowing for a different number of observations before and after that date.

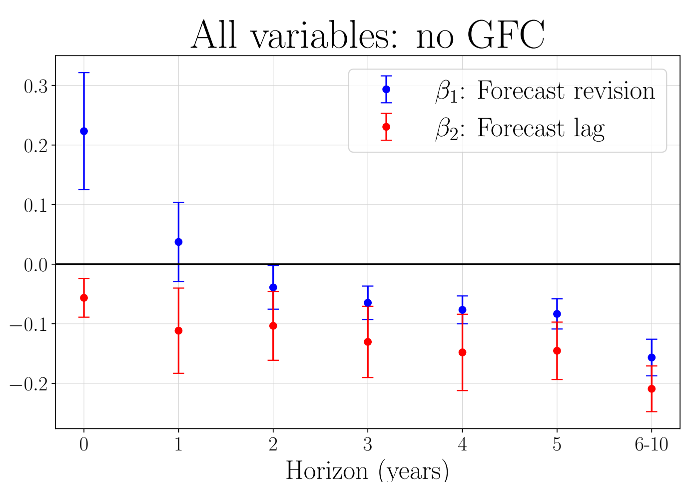
[11] Appendix figure 14 shows the distribution of coefficients when running the regression for each individual year.



(a) First Half

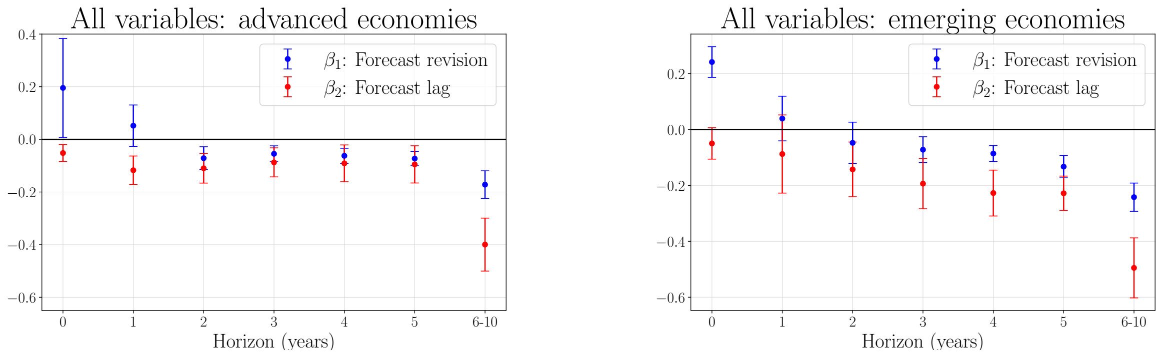


(b) Second Half



(c) No GFC

Figure 3: Regression coefficients from $e_{c,t+h} = \alpha + \beta_1 \Delta \mathbb{E}_t[x_{c,t+h}] + \beta_2 \mathbb{E}_{t-1}[x_{c,t+h}] + f_{c,x} + \varepsilon_{c,t}$ for different splits of the sample across the time dimension, with x pooled across all forecast variables. Country-variable fixed effects are included and standard errors are Driscoll-Kraay with country-variable groupings.



(a) Advanced economies

(b) Emerging economies

Figure 4: Regression coefficients from $e_{i,t+h} = \alpha + \beta_1 \Delta \mathbb{E}_t[x_{i,t+h}] + \beta_2 \mathbb{E}_{t-1}[x_{i,t+h}] + f_{i,t,x} + \varepsilon_{i,t}$, splitting the sample into advanced and emerging economies, with x pooled across all forecast variables. Country-variable fixed effects are included and standard errors are Driscoll-Kraay with country-variable groupings.

erywhere negative). Conceptually, we prefer the specifications without time fixed effects because if forecasters are making a similar forecast error across countries/variables in a given time period, that is something a measure of overreaction *should* capture.

Is noise in long-horizon expectations an issue? de Silva and Thesmar (2024) show that if longer-horizon forecasts are noisier than shorter-horizon forecasts, the β_1 coefficient in a regression of forecast errors on forecast revisions is biased downwards. If Consensus long-horizon forecasts are noisier, that would present an issue for our estimates.

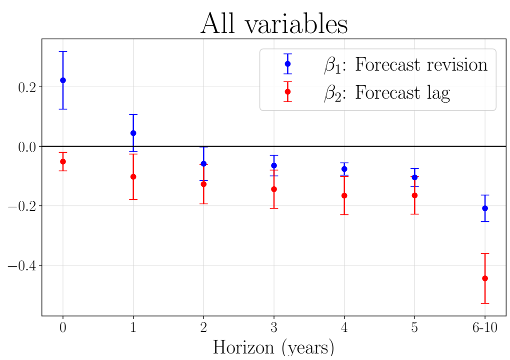
However, appendix figures 12 and 13 plot the mean and median of the standard deviation of forecasts, across forecasters, for a given variable at each horizon, as well as the mean squared error of forecasts at different horizons (by variable).[12] In our data, the standard deviation of forecasts if anything *decrease* in horizon. The mean squared error is slightly elevated at medium horizons, but not noticeably larger at six-to-ten year horizons than at zero-year horizons.

This pattern in our data on macroeconomic forecasts contrasts with the pattern in de Silva and Thesmar (2024), who instead study stock-level earning forecasts. We leave it to future work to examine whether this is due to differences in the true data generating processes, differences in the types of forecasters, or differences in biases.[13]

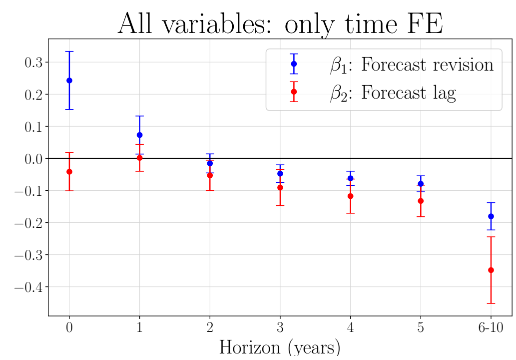
Our results are also consistent with Ahn and Farmer (2024), who find inflation forecast "noise" (forecaster disagreement) decreasing in the horizon of the forecast.

[12] Consensus Economics directly provides the standard deviation of forecasts across forecasters at each horizon, rather than the underlying individual forecasts which could be used to calculate them.

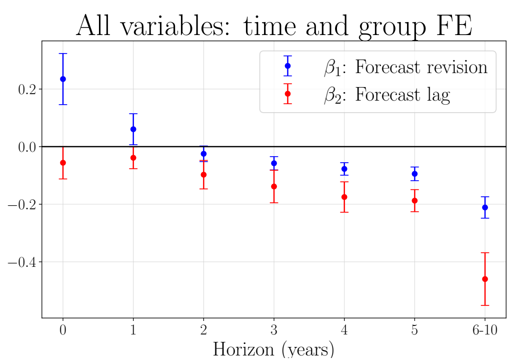
[13] Patton and Timmermann (2010), also find that forecaster disagreement increases in horizon when going from one-month-ahead forecasts to 24-month-ahead forecasts, at a monthly frequency. Our appendix figures 12 and 13 based on *annual* data also typically show a modest rise in forecaster disagreement when going from a 0-year to a 1-year horizon, but typically decreasing beyond the 1-year horizon.



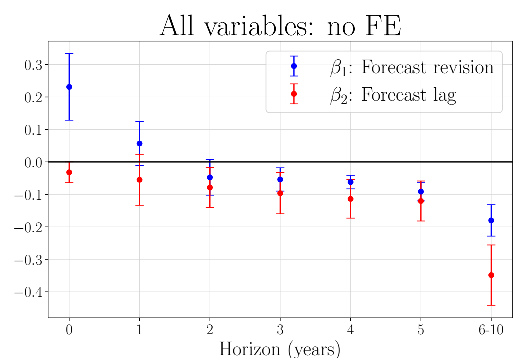
(a) Country-by-variable fixed effects



(b) Time fixed effects



(c) Country-by-variable and time fixed effects



(d) No fixed effects

Figure 5: Regression coefficients from $e_{c,t+h} = \alpha + \beta_1 \Delta \mathbb{E}_t[x_{c,t+h}] + \beta_2 \mathbb{E}_{t-1}[x_{c,t+h}] + f_{c,x} + f_t + \varepsilon_{c,t}$ for different combinations of fixed effects. Standard errors are Driscoll-Kraay with country-variable groupings.

Restricting to the biannual sample. Recall that prior to 2014, forecasts are available biannually (in Q2 and Q4); since, forecasts have been available quarterly. Appendix figure 17 shows that if the data is restricted to the biannual (Q2 and Q4) sample, then the same pattern of results is evident. Appendix figure 18 shows the same when restricting to simply the first survey of the year.

3.4 Out-of-sample Forecasting

Eva and Winkler (2023) argue that in order to reject rational expectations, bias-adjusted forecasts should have lower errors than raw forecasts, out of sample. Yet, for the forecast biases they test – such as the Coibion and Gorodnichenko (2015) aggregate under-revision bias – adjusting for the bias does not consistently improve out-of-sample forecasts. This is an important challenge to the literature arguing for systematic departures from FIRE, since it suggests the biases found by researchers are not very stable over time.

Eva and Winkler (2023) use exclusively US forecast data with a horizon of one year or less. Therefore, they are not able to test whether the cross-country and long-horizon biases we have documented work out of sample. Here we conduct such an exercise and show that our forecast biases *do* help predict forecast errors out of sample, providing additional support for the robustness and importance of long-horizon overreaction.

Our out-of-sample forecasting test follows the procedure in Eva and Winkler (2023) closely. At each survey date, we take an expanding z-score at the country-horizon-variable level. Then, on this "training" data we run regression (1).

Next, we consider the next survey date in our dataset and compute a "bias-adjusted" forecast using the estimated $\hat{\beta}$ results of the previous regression. Specifically, the bias-adjusted forecast is computed as:

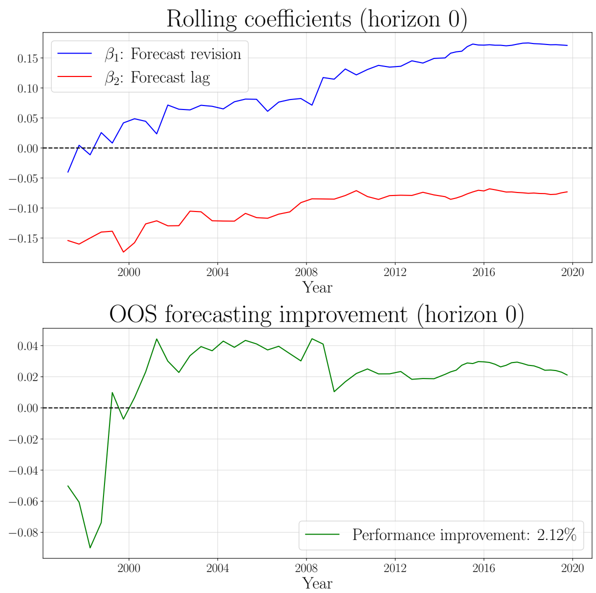
$$\mathbb{E}_t^*[x_{c,t+h}] \equiv \mathbb{E}_t[x_{c,t+h}] + \hat{\beta}_{1,t}\Delta\mathbb{E}_t[x_{c,t+h}] + \hat{\beta}_{2,t}\mathbb{E}_{t-1}[x_{c,t+h}] \tag{2}$$

The asterisk indicates that $\mathbb{E}_t^*[x_{c,t+h}]$ is a bias-adjusted forecast. The adjusted forecast uses only information that a forecaster at the time who had access to the history of forecasts would be able to use.

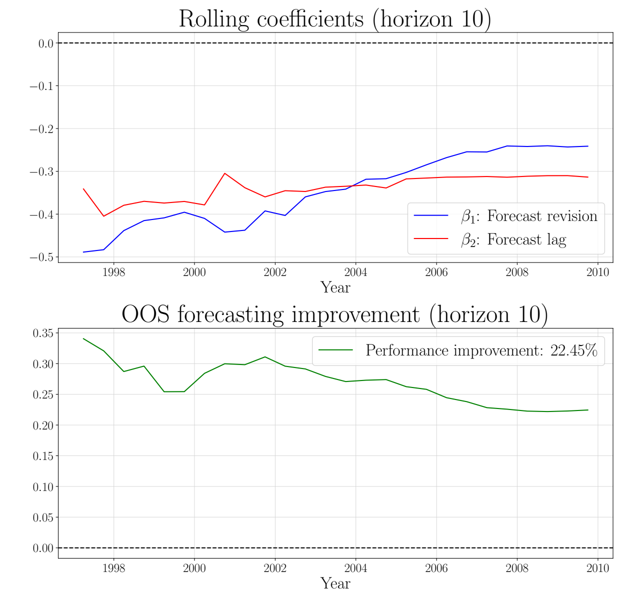
We then compare the performance of the adjusted forecast with the actual forecast and compute the sum of squared errors for each: $\text{SSE}_h = \sum_{c,t} x_{c,t+h} - \mathbb{E}_t[x_{c,t+h}]$ and $\text{SSE}_h^* = \sum_{c,t} x_{c,t+h} - \mathbb{E}_t^*[x_{c,t+h}]$, where SSE_h is the sum of squared errors from the actual (horizon-h) forecasts and SSE_h^* is the sum of squared errors from the bias adjusted forecasts. Finally, we compare the performance of the two by calculating the following relative performance (RP) metric:

$$\text{RP}_h = \frac{\text{SSE}_h - \text{SSE}_j^*}{\text{SSE}_h} \tag{3}$$

The relative performance is the percentage increase in cumulative squared errors due to using the actual forecasts as opposed to the bias-adjusted forecasts. A positive relative performance therefore indicates that actual forecasts are worse than bias-adjusted forecasts.



(a) Horizon 0 OOS Forecasting



(b) Horizon 6-10 OOS Forecasting

Figure 6: The top subplots on the left and right show the rolling estimates of β_1 and β_2 as the sample increases in regression (1). The bottom subplots shows the % difference in sum of squared forecasting errors from using the original forecasts versus the bias-adjusted forecasts, as defined in (3).

Figure 6 compares the out-of-sample results for horizon zero and horizon six-to-ten. The left-hand panel 6a is for horizon zero. The top of the left-hand panel shows the β_1 and β_2 coefficients at each point in time, with the β_1 coefficients in blue and β_2 coefficient in red. The bottom panel shows RP over time, with a positive value indicating the bias-adjusted forecast outperforms the actual forecast.

The idea behind plotting the coefficients over time is to see how stable the over- or under-reaction phenomenon is.[14] Both the zero-year horizon and six-to-ten-year horizon plots show a fair amount of coefficient stability. The β_1 coefficient starts out negative at the very beginning of the horizon zero sample, but otherwise the signs of the coefficients are consistent, and by the last five years or so of the sample the coefficients are very stable.

The horizon zero cumulative performance figure on the bottom-left shows the bias-adjusted forecast slightly outperforming the unadjusted forecast after starting the sample

[14]There are two things to note. First, the dates on the x-axis represent the date of the *forecast*, not the date of the outcome. This is why the x-axis for the six-year horizon stops in 2014: any forecasts made after that point would include Covid, which we drop. Second, we do not use country-variable fixed effects since they might have undue influence in the smaller samples at the beginning of the data. Still, the end of sample coefficients here will be slightly different than the coefficients plotted in 5d, which is the specification without group fixed effects, because they do not include the last year-quarter unit and because of the five data point burn-in period.

under-performing. By the end of the sample the cumulative relative performance improvement is 2.12%. The improvement from bias-adjusting at the ≤ 1 year horizon is better than the negative results in Eva and Winkler (2023), but the small magnitude of the advantage is consistent with their overall message.

The horizon six-to-ten cumulative performance figure in the bottom-right tells a very different story. The bias-adjusted forecast always outperforms the unadjusted forecast by at least 20%, and a 20-30% performance gap remains steady for most of the sample. Outperformance is 22.45% by the end of the sample. Note that a flat line here is consistent with the bias-adjusted forecast continually outperforming, since the y-axis is in percentage terms: if the forecasts started performing equally well, the line would decline towards zero performance difference.

Table 2 shows the relative performance improvement from the bias adjustment roughly increases with foreccast horizon, though unevenly. For forecasts at or beyond the one-year horizon, the bias-adjusted forecast outperforms the unadjusted forecast by over 10%. At horizon two and beyond – where there is overreaction as measured by both β_1 and β_2 – adjusting for overreaction leads to 16-23% better forecasting performance. Appendix figures 19-23 show plots like figure 6 for horizons one through five.

Table 2: Final Cumulative SSE Difference by Horizon (in percentages)

0	1	2	3	4	5	10
2.1%	11.0%	17.8%	16.1%	16.0%	16.1%	22.4%

3.5 Discussion

Across forecasts of GDP, inflation, consumption, and investment; across emerging markets and advanced economies; and from pre- to post-2008, the same four facts hold: forecasts under-revise at short-horizons, over-revise at two-year and longer horizons, feature over-extremity at all forecasts horizons, and the over-revision and over-extremity grow with the time horizon of forecasts. At six-to-ten year horizons – our longest horizon – forecast overreaction is strongest by both measures.

Other papers have shown that longer-horizon expectations overreact (Bordalo, Gennaioli, La Porta, and Shleifer 2024, Angeletos, Huo, and Sastry 2021, Afrouzi, Kwon, Landier, Ma, and Thesmar 2023, d'Arienzo 2020), but as far as we are aware, we are the first to show that overreaction in specifically *macroeconomic* expectations increases with horizon, and does so across a broad and representative cross-country sample. Furthermore, we are the first to show that adjusting for overreaction allows for improved out-of-sample forecasting.

4 A Model that Fits the Facts

As discussed more extensively in Afrouzi, Kwon, Landier, Ma, and Thesmar (2023), the fact that overreaction increases in horizon is inconsistent with popular models of overreaction, where the degree of overreaction is independent of forecast horizon. In particular, the standard model of over-extrapolation and precision over-estimation, as in Angeletos, Huo, and Sastry 2021, does not produce overreaction that increases in horizon. Similarly, the standard model of diagnostic expectations does not produce overreaction that increases in horizon.

To match the moment that overreaction increases in horizon (fact 4) together with facts 1 through 3, we extend the costly-recall model of Afrouzi, Kwon, Landier, Ma, and Thesmar (2023). Their baseline model is already able to match facts 2, 3, and 4. Our first theoretical contribution is to show this is true for the regression framework in (1). Then, we show that adding a *sticky* information friction allow us to jointly match facts 1 and 3: that expectations under-revise at horizons of one year or less but expectations are "too extreme" at all horizons. By contrast, *noisy* information frictions would predict that short-horizon under-reaction is *more* pronounced for the β_2 coefficient than β_1, contrary to the data.

We then go on to show that a calibrated version of the model – using existing parameter estimates from the lab experiments of Afrouzi, Kwon, Landier, Ma, and Thesmar (2023) when available – is able to quantitatively match the data.

4.1 Baseline Costly-Recall Model

Here we introduce in brief the baseline costly-recall model found in Afrouzi, Kwon, Landier, Ma, and Thesmar (2023). Agents forecast the following AR(1) process:

$$x_t = (1-\rho)\mu + \rho x_{t-1} + \epsilon_t \quad \epsilon_t \sim (0, \sigma_\epsilon^2) \tag{4}$$

The state x_t is perfectly observed by agents, but the long-run mean μ is unknown. Agents form forecasts $F_t x_{t+h}$ of future realizations of the state, given the objective $-(F_t x_{t+h} - x_{t+h})^2$. In order to estimate the long-run mean, agents rely on their knowledge of the history of x_t. The costly-recall friction makes it so the most recent observation x_t is freely available to agents, but prior realizations x_{t-h} are costly to retrieve from memory. Agents begin with the prior $\mu \sim N(x_t, \underline{\tau}^{-1})$, and update their prior based on how much historical information S_t the agent chooses to process. Formally, agents face the following costs in processing a set of information S_t:

$$C_t(S_t) \equiv \omega \frac{\exp\left(\gamma \mathbb{I}(S_t, \mu | x_t)\right) - 1}{\gamma} \tag{5}$$

The parameters ω and γ determine the scale and convexity of the cost function, while $\mathbb{I}(S_t, \mu | x_t)$ is Shannon's mutual information function representing the amount of information the agent uses. Afrouzi, Kwon, Landier, Ma, and Thesmar (2023) contain a more detailed explanation of the psychological foundations of this model, but the high-level intuition is straightforward: the long-run mean is unknown, it is costly to process information, and more recent information is less costly to recall.

Given this environment, Afrouzi, Kwon, Landier, Ma, and Thesmar (2023) proposition 1 shows that agents forecasts systematically overreact relative to a rational (and costless information processing) benchmark:

$$F_t x_{t+h} = \underbrace{E_t x_{t+h}}_{\text{rational forecast}} + \underbrace{(1-\rho^h)\min\left\{1, \left(\frac{\omega\underline{\tau}}{(1-\rho^h)^2}\right)^{\frac{1}{1+\gamma}}\right\} x_t}_{\text{overreaction}(\equiv\Delta)} + \underbrace{u_t}_{\text{noise}} \tag{6}$$

If agents forecast according to (6), we can derive the following implications for the regression coefficients in (1) on forecast revisions and lagged forecasts:

Proposition 1: in the regression on forecast revisions and lagged forecasts,

(i) $\beta_1^h = \beta_2^h = -\frac{\Delta_h}{\rho^h+\Delta_h} \leq 0$

(ii) $\frac{d\beta_1(h)}{dh} = \frac{d\beta_2(h)}{dh} < 0 \quad \text{if} \quad \gamma \geq 1.$

Proof: See appendix A.3.1.

Proposition 1 shows that this baseline costly-recall model is already sufficient to match our empirical facts 2, 3, and 4. With no additional frictions, the β_1 and β_2 coefficients are less than or equal to zero and decreasing in horizon, implying that overreaction is increasing in horizon, as long as the costly-recall function is weakly convex, $\gamma > 1$.

4.2 Noisy and Sticky Info

In order to match the fact that the short-horizon β_1 coefficients indicate under-revision, we consider two additional frictions.

(i) Under *noisy* information, instead of perfectly observing x_t, agents observe a noisy signal s_t:

$$s_t = x_t + e_t \quad e_t \sim (0, \sigma_e^2)$$

In this case, agents beliefs about x_t will be a weighted combination of the signal and their prior.

(ii) Under *sticky* information, in each period only a fraction λ of agents update their forecasts.

The following proposition characterizes the key implications of noisy and sticky information for the regression coefficients in (1):

Proposition 2: Under noisy information where agents observe noisy signals s_t of x_t and update their posterior beliefs about x_t in a Bayesian manner,

(i) $\beta_1(h) < \beta_2(h) \quad \forall h$

Additionally, under sticky information where only a fraction $\lambda < 1$ of agents update their forecasts each period,

(i) $\beta_1(h) > \beta_2(h) \quad \forall h$

(ii) $\beta_1(h) > 0$ is possible, while $\beta_2(h) \leq 0 \quad \forall h$

Proof: See appendix A.3.2.

Proposition 2 shows that the sticky information friction is consistent with facts 1 and 3, while the noisy information friction is not. Recall fact 1 is under-revision ($\beta_1 > 0$) at short-horizons, and fact 3 is over-extremity ($\beta_2 \leq 0$) at all horizons. In contrast, the noisy information friction counter-factually predicts that we would see ”more” underreaction in the β_2 coefficient than β_1. Therefore, our results can be interpreted as evidence for sticky information frictions and against noisy information frictions, at least in this population of economic forecasters.

4.3 Model Calibration

We now show that calibrating the sticky information and costly recall model outlined above can deliver quantitative predictions in line with our empirical results.

Note that the model applies to forecasting a single variable with a single long-run mean. Accordingly, while in section 3.2 our benchmark analysis pooled across variables for power, in this section we focus on GDP forecasts. We calibrate the model to match the “average” country's GDP process.

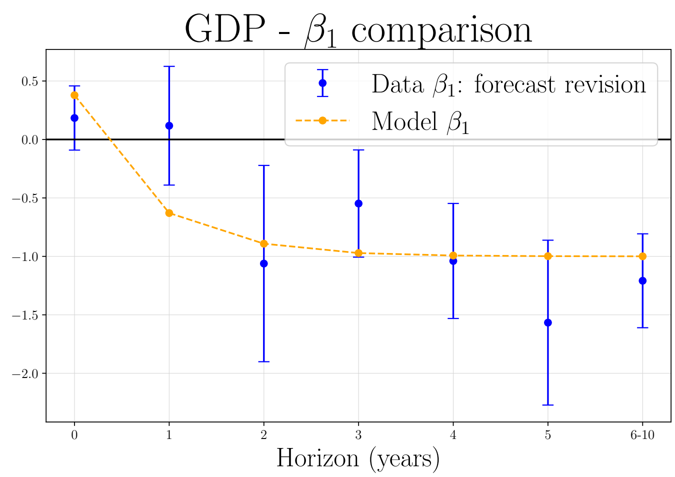
The calibration proceeds as follows:

(i) The values of the scale and convexity parameters of the costly-recall function 5 ω and γ are taken from the main calibration in Afrouzi, Kwon, Landier, Ma, and Thesmar (2023): $\omega = 0.42$ and $\gamma = 2$. These parameters are estimated from subjects in the lab who are asked to forecast simulated GDP growth processes.

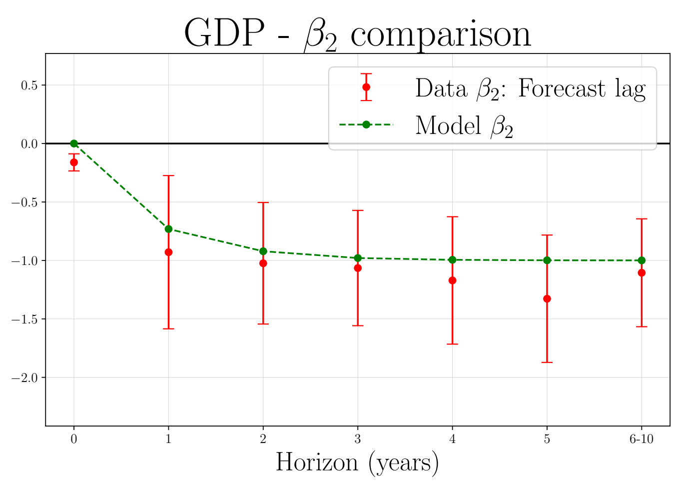
(ii) The persistence of GDP is taken from the regression coefficient of GDP growth on its value one-year prior in a regression with all countries GDP growth data for which they have expectations data, estimated with country fixed effects ($\rho = 0.25$)

(iii) The fraction of forecasters who update their forecast each period is calibrated to minimize the sum of squared deviations between model and data coefficients: $\lambda = 0.725$.

We compare the implied model coefficients to the empirical results from regression (1) when variables are *not* z-scored (since the model's implied coefficients would be transformed by z-scoring, while preserving the results of Proposition 2). Figure 7 displays the results.



(a) Beta 1: FC Revision



(b) Beta 2: Lagged FC

Figure 7: Subplot (a) displays the β_1 coefficient by horizon from the costly-recall and sticky-info model (orange line) against the β_1 coefficient in the data (blue). Subplot (b) contrasts the model (green) and data (red) β_2 coefficients. The data is for GDP forecasts with no z-scoring. Country fixed effects are included and standard errors are Driscoll-Kraay with country groupings. The model is calibrated with $\omega = 0.42, \gamma = 2, \rho = 0.25, \lambda = 0.8$.

Figure 7a shows the β_1 coefficient in the model versus the data. The model is able to replicate under-revision at horizon 0 which flips to over-revision as the horizon increases. The model implied coefficients are within the confidence interval of the data-implied coefficients except for horizon 1, where the model predicts a lower β_1 than seen in the data. Figure 7b shows that the model predicts $\beta_2 = 0$ at horizon 0 and then declines sharply afterwards, matching the data quite well.

Our interpretation is not that this model is the only possible model consistent with our facts. Rather, it provides a natural baseline: a simple extension (sticky information) to a model with external validity in a very different setting (costly recall). It also helps discriminate between sticky-info and noisy-info models, both of which are widely used in the behavioral macroeconomics literature. Other relevant modeling work that could provide cognitive micro-foundations for the forecasting biases we document includes Sung (2024), Bianchi, Ilut, and Saijo (2024b), Augenblick, Lazarus, and Thaler (2025), Ba, Bohren, and Imas (2024), and de Silva, Larsen-Hallock, Rej, and Thesmar (2025).

5 Expectations and the Economy

Bordalo, Gennaioli, La Porta, O'Brien, and Shleifer (2024) use US data to show that changes in expectations of "long-term" stock market earnings growth predict a short-term boom in real variables like growth and investment followed by a subsequent bust.[15]

In this section, we show that, in our broad cross-country sample, the relationship between *output* growth forecasts and subsequent investment and GDP growth follows a slightly different pattern. There are systematic "booms" and "busts" in GDP and investment growth, but it is changes in *short-term* GDP growth forecasts which are most predictive of those booms and busts. This stands in contrast to the long-term forecasts emphasized by Bordalo, Gennaioli, La Porta, O'Brien, and Shleifer (2024). In light of our previous section's findings – that it is long-horizon expectations which overreact most – these results suggest movements in *forecasters'* long-horizon expectations are relatively inert: their movement is not associated with changes in economic agents activity. That could be because forecasters expectations are not representative of economic decision makers' expectations, or because even economic decision makers' short-term actions do not respond significantly to their long-term beliefs.

5.1 Local Projections Approach

In order to examine whether changes in growth expectations influence business cycle fluctuations, we follow the local projections approach used in Bordalo, Gennaioli, La Porta, O'Brien, and Shleifer (2024). Specifically, we estimate panel local projections of the

[15]Where the measure of "long-term" expectations covers average annual earnings growth over the next three-to-five years.

following form:

$$x_{c,t+h} = \alpha + \zeta_h \Delta_1 \mathbb{E}_t(g_{c,t+h}) + \beta \mathbf{Z}_{c,t} + f_c + \epsilon_{c,t} \tag{7}$$

The dependent variable is any macroeconomic variable of interest x, for country c, at horizon h. The coefficient of interest is ζ_h, which multiplies the one-year revision in h-year ahead GDP growth expectations, $\Delta_1 \mathbb{E}_t(g_{t+h})$. We compare two horizons:

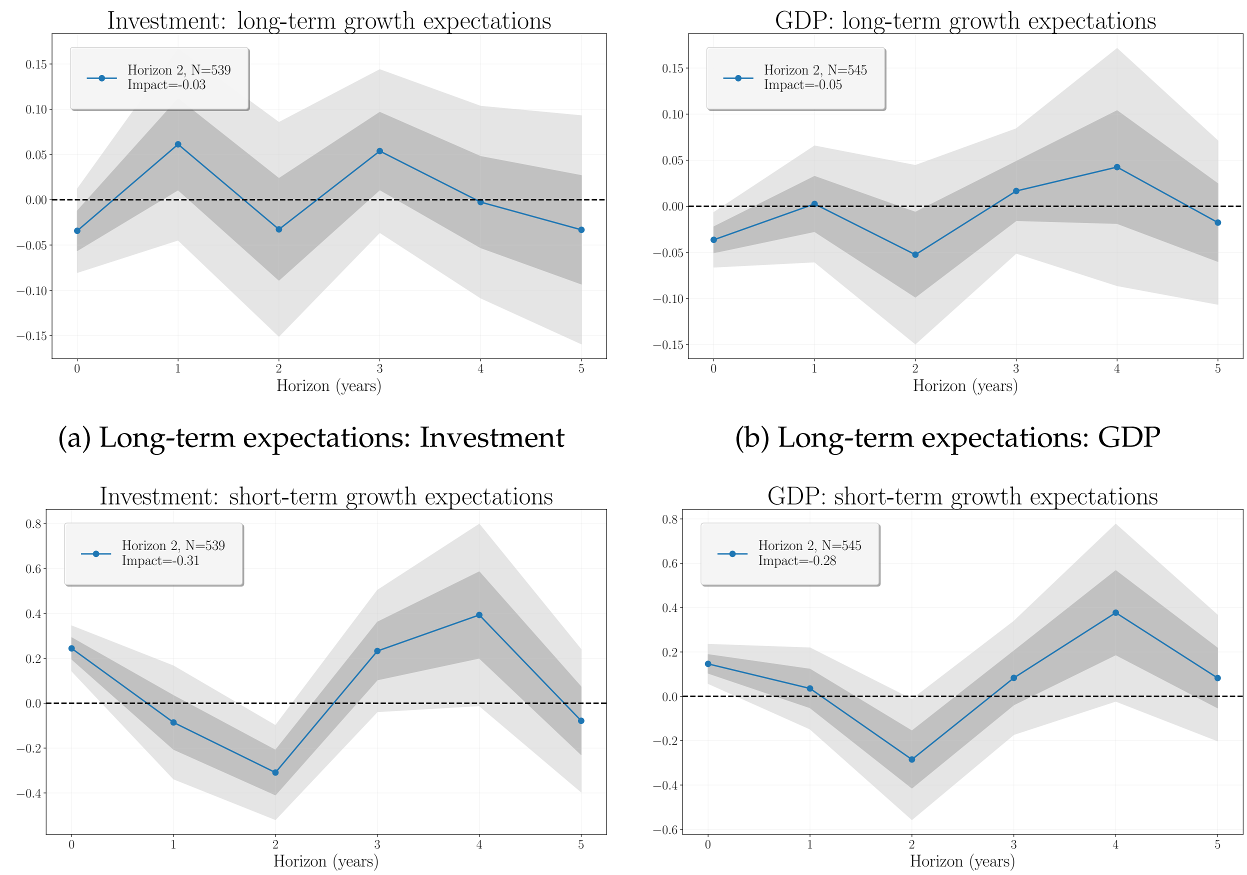
(i) Long-term growth expectations: six-to-ten-year ahead annual GDP growth expectations.

(ii) Short-term growth expectations: average annual zero-to-two-year GDP growth expectations.

The vector of controls $\mathbf{Z}_{c,t}$ follow Bordalo, Gennaioli, La Porta, O'Brien, and Shleifer (2024), with the addition of the current-year forecast of the dependent variable, $\mathbb{E}_t(y_t)$. Controlling for contemporaneous forecasts is important to account for the fact that if the forecast revision is measured in the middle of the year, there is information about current-year economic conditions that is not controlled for by our other lagged controls. The rest of the controls are lagged macroeconomic variables to control for standard business cycle dynamics: the contemporaneous 10-year real rate; the one-year change in the 10-year real rate and one-year stock market return up to the forecast date; the one-year lag of the country's GDP growth, investment growth, inflation, and stock market return; the change in GDP growth and investment growth from $t-2$ to $t-1$ and $t-3$ to $t-2$; the two-year lag of inflation and the stock market return; and the $t-2$ to $t-1$ and $t-3$ to $t-2$ change in the country's 10-year real interest rate.[16] In robustness checks, we use a smaller subset of these control variables – just the one-year lag of GDP, investment, and stock market returns as well as the $t-2$ to $t-1$ change in investment and GDP. These controls help isolate the role that changes in expectations alone have on business cycle dynamics. Any causal influence of past macroeconomic aggregates on the change in expectations that also affects future macroeconomic aggregates independently of their effect on expectations will be stripped out.

We continue to z-score all variables by country and use country fixed-effects f_c. As a result, the coefficients represent the impact of the independent variable on the dependent variable in standard deviation units. We again use Driscoll-Kraay standard errors.

[16]The ten-year real-interest rates we use are *ex-ante* real interest rates constructed in two different ways: 1) where available, we use inflation-linked 10-years, like TIPS in the US, downloaded from Bloomberg. 2) where inflation-linked 10-years are unavailable, we use 10-year nominal rates from the OECD "long-term rates" data set and subtract 10-year inflation expectations from the Consensus survey. Stock market data comes from the WRDS "Daily World Indices" database.



(a) Long-term expectations: Investment

(b) Long-term expectations: GDP

(c) Short-term expectations: Investment

(d) Short-term expectations: GDP

Figure 8: Regression coefficients from $y_{c,t+j} = \alpha + \beta_1 \Delta_1 \mathbb{E}_t(g_{c,t+h}) + \beta_2 \mathbb{E}_t(y_t) + \beta_3 \mathbf{X}^*_{c,t} + f_c + \epsilon_{c,t}$ for different horizon-forecasts g. The top-row uses the one-year change in six-to-ten year ahead average annual GDP growth expectations; the bottom-row uses the one-year change in average annual GDP growth expectations for the year of the forecast and the following two years. The dependent variable on the LHS column is investment, and on the RHS is GDP. Country-variable fixed effects are included and standard errors are Driscoll-Kraay with country-variable groupings.

5.2 Results

Figure 8 is a two-by-two panel. In the left-hand column, the dependent variable is investment growth – the traditional propagator of "animal spirit"-driven booms and busts. In the right-hand column, the dependent variable is GDP growth. Dark gray bands represent 68% confidence intervals, and light gray bands represent 95% confidence intervals.

In the first row, the independent variable is the one-year change in long-term growth expectations (the six-to-ten-year horizon). In the second row, the independent variable is the one-year change in "short-term growth expectations" (the zero-to-two-year horizon).

The first panel row shows, at best, a very mild boom-bust pattern from changes in long-term expectations. There is no statistically detectable response of either investment or GDP growth to changes in long-term growth expectations at the 95% confidence level. Furthermore, the magnitude of the "boom" and "bust" predicted by changing long-term expectations is less than one-tenth of a standard deviation.

By contrast, the second row shows that changes in short-term growth expectations are strongly associated with booms and busts. Both investment and GDP boom in a highly significant fashion in the initial year, and both of them bust dramatically two years later.[17] The magnitudes of the booms and busts are a bit over a third of a standard deviation – nearly five times larger than the association for long-term expectations. The pattern of results in this panel matches well the "main business cycle" shock identified in Angeletos, Collard, and Dellas (2020), which can result from fluctuations in short-term economic confidence (Angeletos, Collard, and Dellas 2018).

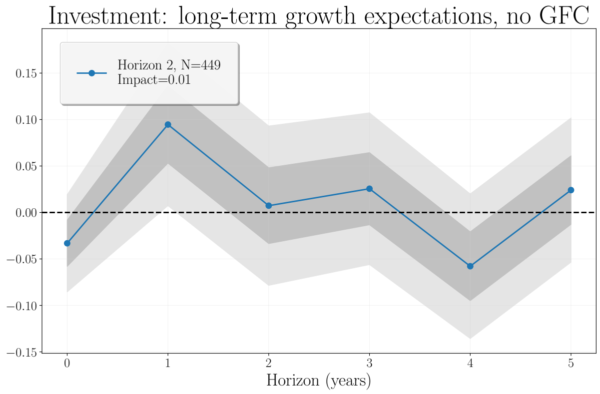
5.2.1 No GFC

We next show that the "bust" portion of these results is fairly reliant on including the global financial crisis (GFC) in the sample. Figure 9 shows the exact same type of results, except with one-year expectation changes in 2005, 2006, and 2007 removed from the sample.

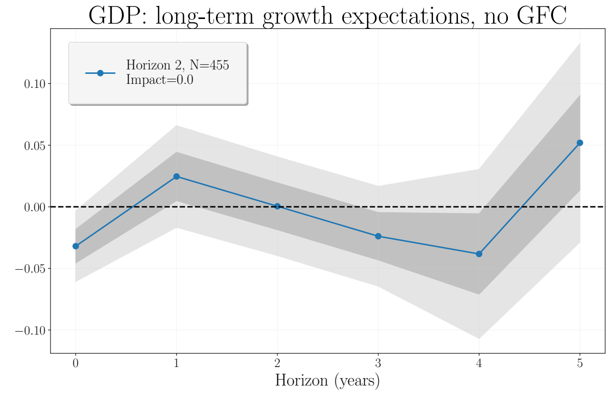
The top row shows that, once the GFC is excluded, there continues to be no significant association between changes in long-term growth expectations and booms and busts. The bottom row shows that upward revisions in short-term growth expectations continue to be associated with booms in investment and growth – even within the current year where the forecast is controlled for – consistent with the finding that average short-term forecasts under-revise. However, the association between upward movements in short-term growth expectations and two-year-ahead busts is significantly dampened: the two-year horizon coefficients for both investment growth and GDP growth are about one-fourth to one-third of their magnitude when the GFC is included, and they are no longer statistically significant.

The attenuation of the bust results when the GFC is excluded is not necessarily decisive evidence *against* the boom-bust mechanism: arguably the GFC is the prime example in the sample of exactly the sort of boom-bust dynamics consistent with an animal spirits story. Given that the non-GFC results point in the same direction, even if insignificant, we prefer to interpret our results as showing that changes in short-term growth expectations have a tendency to create boom-bust dynamics, but by no means always necessitate subsequent booms and busts.

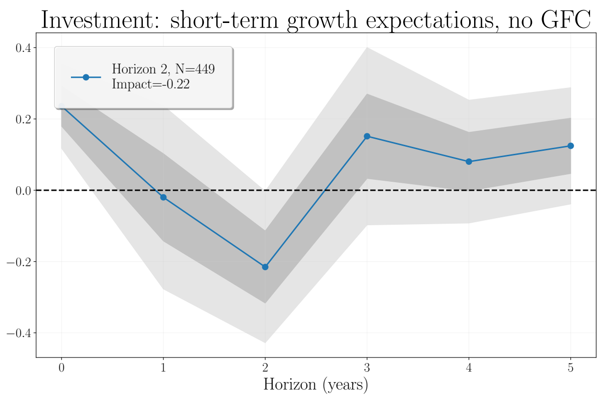
[17] The p-value on the two-year horizon coefficient where GDP is the dependent variable is .0516.



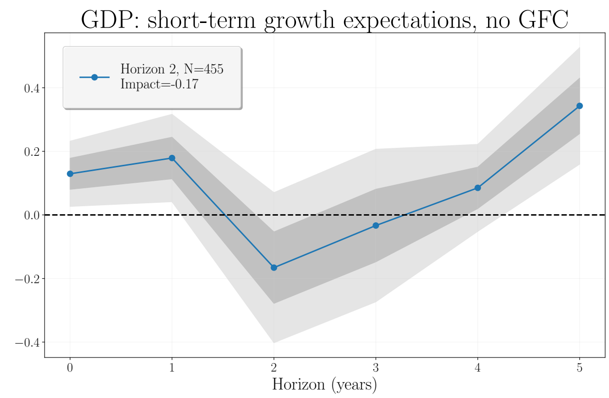
(a) No GFC, LT expectations: Investment



(b) No GFC, LT expectations: GDP



(c) No GFC, ST expectations: Investment



(d) No GFC, ST expectations: GDP

Figure 9: No GFC: Regression coefficients from $y_{c,t+j} = \alpha + \beta_1 \Delta_1 \mathbb{E}_t(g_{c,t+h}) + \beta_2 \mathbb{E}_t(y_t) + \beta_3 \mathbf{X}^*_{c,t} + f_c + \epsilon_{c,t}$ for different horizon-forecasts g. The top-row uses the one-year change in six-to-ten year ahead average annual GDP growth expectations; the bottom-row uses the one-year change in average annual GDP growth expectations for the year of the forecast and the following two years. Changes in expectations from 2005, 2006, and 2007 are removed from the sample. The dependent variable on the LHS column is investment, and on the RHS is GDP. Country-variable fixed effects are included and standard errors are Driscoll-Kraay with country-variable groupings.

Table 3 presents the regression results for horizons zero through four years later, for both the with and without GFC samples. The main entries are the β_1 coefficient estimates and the parentheses below the coefficient are (X, Y) where X is the number of observations for that regression and Y is the (Driscoll-Kraay) standard error of the β_1 coefficient.[18]

[18]Due to all the control variables used, the sample is significantly smaller here. Appendix figure 24 shows that if you use *no* control variables other than the current-year forecast you still get a similar pattern of results; however, we do not think the quantitative part of those results should be paid much attention.

Table 3: Changes in growth expectations on actual investment and growth

	Horizon				
Variable	h=0	h=1	h=2	h=3	h=4
LT Growth → GDP	−0.02	0.03	−0.07	−0.01	0.03
	(713, 0.02)	(630, 0.04)	(548, 0.06)	(486, 0.03)	(444, 0.07)
LT Growth → Investment	−0.02	0.08	−0.05	0.03	−0.00
	(710, 0.03)	(624, 0.07)	(542, 0.07)	(481, 0.04)	(440, 0.05)
ST Growth → GDP	0.28***	0.34*	−0.32*	−0.08	0.20
	(715, 0.06)	(631, 0.18)	(548, 0.17)	(486, 0.12)	(444, 0.14)
ST Growth → Investment	0.33***	0.22	−0.31**	0.08	0.30**
	(712, 0.06)	(625, 0.15)	(542, 0.15)	(481, 0.11)	(440, 0.14)
No GFC					
LT Growth → GDP	−0.01	0.04	0.00	−0.01	−0.05
	(622, 0.02)	(539, 0.03)	(457, 0.02)	(395, 0.02)	(353, 0.04)
LT Growth → Investment	−0.01	0.11*	0.01	0.04	−0.06
	(619, 0.03)	(533, 0.06)	(451, 0.04)	(390, 0.04)	(349, 0.04)
ST Growth → GDP	0.26***	0.42**	−0.08	$3.35e-03$	0.03
	(624, 0.06)	(540, 0.16)	(457, 0.10)	(395, 0.11)	(353, 0.07)
ST Growth → Investment	0.33***	0.26**	−0.12	0.14	0.10
	(621, 0.06)	(534, 0.12)	(451, 0.09)	(390, 0.11)	(349, 0.08)

Notes: "LT" indicates long-term expectations, and "ST" indicates short-term expectations. Values in parentheses show (N, SE) where N is the number of observations and SE is the standard error.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

6 Stock Return Predictability

Bordalo, Gennaioli, La Porta, and Shleifer (2024) show that the same index of "long-term" expected earnings growth predicts negative stock market returns on five-year horizons in the US better than short-term expectations, and they claim this is evidence that overreacting long-term expectations help explain a number of stock market puzzles.

In this section, we show that *in US data* the relationship between *GDP* growth forecasts and subsequent stock market returns is similar – but differs for other countries. Outside of the US, high *short-term* GDP growth forecasts are most predictive of subsequent local country stock market returns, due to their association with short-term *weak* returns. This result is consistent with our previous section that changes in short-term expectations are most strongly associated with booms and busts in the business cycle.

6.1 Results in the US

First, we show a similar pattern of results as Bordalo, Gennaioli, La Porta, and Shleifer (2024) when looking at just US data. The US data runs from April 1990 to July 2023.[19] Stock market returns for the US are real cum-dividend returns on the S&P 500. Nominal returns are deflated by the World Bank's World Development Indicators CPI measure, for the US and for all other countries.

At each date at which we have a new set of consensus expectations, we construct the one-year, three-year, five-year, and the one-year-four-year-forward (1y4y) real cum-dividend return. The latter is the return from $t + 1$ to $t + 5$, where $t + 1$ begins one year from the date of the forecast.

Table 4 examines return predictability at these different horizons relative to three different expectations: short-term growth expectations (zero-to-two years), average growth expectations over the next ten years, and long-term growth expectations (six-to-ten years ahead). We include both zero-to-ten year and six-to-ten year average forecasts is because while six-to-ten year average forecasts are arguably the right measure of "long-term" growth expectations, the zero-to-ten year horizon is more comparable to the expectations measure used in Bordalo, Gennaioli, La Porta, and Shleifer (2024).

The key result in the table is that, within the US, *long-term* expectations – the six-to-ten year ahead average growth expectations – are the strongest return predictors for horizons beyond one-year. For 3-year, 5-year, and 1y4y year returns the regression using six-to-ten year average GDP growth has the largest R^2 and the largest coefficients, in absolute value. Since all variables are z-scored the coefficients are directly comparable across independent variables.

[19] The results we present in this section include 2020, unlike our results from previous sections, since the rapid stock market recovery in the wake of its Covid crash makes it less likely that Covid distorts this data. Appendix tables 7 and 8 show that all the results in this section retain the same pattern when Covid is excluded from the sample.

Table 4: United States: Return Predictability Regressions

	Return Horizon			
	1-year	3-year	5-year	1y4y
2-year avg GDP growth	-0.33^{***}	-0.21^{*}	-0.19	-0.14
	(0.10)	(0.13)	(0.14)	(0.13)
	[13.5%]	[3.7%]	[3.2%]	[1.7%]
10-year avg GDP growth	-0.33^{***}	-0.33^{***}	-0.23^{**}	-0.13
	(0.09)	(0.10)	(0.11)	(0.11)
	[14.2%]	[14.1%]	[6.6%]	[2.1%]
6-10 year avg GDP growth	-0.26^{***}	-0.39^{***}	-0.27^{***}	-0.17^{**}
	(0.09)	(0.08)	(0.08)	(0.08)
	[10.5%]	[23.8%]	[11.3%]	[4.7%]

Notes: Standard errors in parentheses, R^2 in square brackets. *** p<0.01, ** p<0.05, * p<0.1

As in Bordalo, Gennaioli, La Porta, and Shleifer (2024), there is a strong pattern of high-growth expectations predicting subsequently low returns. This holds across all return and expectation horizons, but is strongest for the long-term growth expectations, which predict low returns on both a one-year horizon and one-year-four-year-forward horizon. The level of six-to-ten year ahead GDP growth expectations explains 24% of the variance of the next three year's returns and 11% of the variance of the next five year's returns.

6.2 All countries

We now turn to the results when using a broad sample of 34 countries for which we have both expectations and stock market return data. All stock market return data are cum-dividend returns and come from WRDS's "Daily World Indices" database. Returns are constructed relative to the exact date of the forecast, the same as they were in the US, and are deflated by the WDI's CPI measure. The US remains part of this sample.

Table 5: 34 countries: Return Predictability Regressions

	Return Horizon			
	1-year	3-year	5-year	1y4y
2-year avg GDP growth	-0.30^{***}	-0.14	-0.22^{***}	-0.03
	(0.05)	(0.09)	(0.06)	(0.06)
	[9.4%]	[1.4%]	[3.8%]	[0.1%]
10-year avg GDP growth	-0.31^{***}	-0.13	-0.14^{**}	0.08
	(0.08)	(0.10)	(0.06)	(0.06)
	[6.9%]	[1.1%]	[1.5%]	[0.5%]
6-10 year avg GDP growth	-0.15^{*}	-0.10	-0.04	0.12
	(0.09)	(0.08)	(0.07)	(0.08)
	[1.5%]	[0.7%]	[0.1%]	[1.1%]

Notes: Standard errors in parentheses, R^2 in square brackets. *** p<0.01, ** p<0.05, * p<0.1

Table 5 shows the results. The key result is that, unlike in the US, the long-term six-to-ten year GDP growth expectations no longer significantly predict returns at *any* horizon, while the short-term zero-to-two year growth forecasts are the strongest return predictor. As before, return predictability comes from high growth expectations predicting subsequently *low* returns, but now the relationship is strongest between short-term growth expectations and weak subsequent returns.

The regression results also reveal two other interesting findings. First, the fact that short-term expectations predict weak returns is driven by the relationship between high short-term growth expectations and immediate weak returns in the following year. There is no meaningful relationship between short-term growth expectations and 1y4y forward returns. Second, unlike in the US, there is a *positive* relationship between long-term growth expectations and 1y4y returns. It is not significant, but the fact that it is positive while one-year returns are negatively associated with long-term growth expectations explains why there is no longer-horizon return predictability.

These results suggest that long-term expectations are not a systematic and universal explainer of stock market puzzles, but rather, that US long-term growth expectations have happened to be off in a way that strongly predicted subsequent returns. That could still very well be due to the sort of mechanisms outlined by Bordalo, Gennaioli, La Porta, and Shleifer (2024), but these results show that more work is needed to understand *when* and *why* those sorts of mechanisms kick-in, and whether there are features of the US stock market that makes it uniquely responsive to long-term growth expectations.

One important caveat is it could be that in the US, firms earnings growth expectations are more connected with aggregate GDP growth expectations than in other countries, making it so that long-term *earnings* growth expectations are a consistent explainer of return anomalies, but that our data is unable to capture such a relationship around the world.[20] Appendix tables 9 and 10 show that our results are not significantly different across advanced and emerging market economies, but more work needs to be done to fully examine this hypothesis.

Our finding that short-term growth expectations explain more of subsequent stock market returns than long-term expectations is consistent with that of De la O and Myers (2024), though we emphasize that here we use GDP growth expectations, while they focus on firms' earnings-growth expectations.

7 Conclusion

This paper leverages the Consensus Economics long-term survey – a large cross-country panel of the macroeconomic expectations of professional forecasters out to a ten year horizon – to establish four facts about average macroeconomic expectations:

(i) Less than one-year ahead expectations under-revise. This is the under-revision of average expectations documented by Coibion and Gorodnichenko (2015).

(ii) Two or more year ahead expectations over-revise. There has been other evidence of over-revision at longer horizons, but we are the first to document at precisely which horizons over-revision prevails for macroeconomic expectations.

(iii) Expectations tend to be too extreme at all horizons. The fact that over-extremity tends to be the case at all horizons is novel.

(iv) Over-revision and over-extremity increase in the horizon of the forecast. This matches the experimental evidence about abstract AR(1) processes in Afrouzi, Kwon, Landier, Ma, and Thesmar (2023), suggesting that this pattern is a general feature of human expectation formation.

These facts hold for each of four different macroeconomic variables, over time, and across advanced and emerging economies. By adjusting for these biases in forecasts, we are able to improve out-of-sample forecasting of six-to-ten year ahead macroeconomic outcomes by over 20%.

We then show that an extension of the Afrouzi, Kwon, Landier, Ma, and Thesmar (2023) costly-recall model which also incorporates sticky-information is consistent with our four facts, while a noisy-information extension is not.

Despite long-term expectations overreacting more, it is short-term expectations which are most strongly associated with "booms and busts" in investment, GDP, and the stock market. This result stands in contrast to other recent work which has emphasized the role of long-run overreaction in explaining business cycle and stock market fluctuations.

[20] Allen, Qian, Shan, and Zhu (2024) presents evidence that China's stock market has a different relationships with *realized* GDP growth than other countries, but the patterns across non-China countries are fairly similar.

References

Adam, Klaus, Albert Marcet, and Johannes Beutel (2017). "Stock price booms and expected capital gains". In: *American Economic Review* 107.8, pp. 2352–2408.

Adam, Klaus, Oliver Pfäuti, and Timo Reinelt (2025). "Subjective housing price expectations, falling natural rates, and the optimal inflation target". In: *Journal of Monetary Economics* 149, p. 103647.

Afrouzi, Hassan, Spencer Y Kwon, Augustin Landier, Yueran Ma, and David Thesmar (2023). "Overreaction in expectations: Evidence and theory". In: *The Quarterly Journal of Economics* 138.3, pp. 1713–1764.

Ahn, Hie Joo and Leland Farmer (2024). "Diagreement about the Term Structure of Inflation Expectations". In.

Allen, Franklin, Jun Qian, Chenyu Shan, and Julie Lei Zhu (2024). "Dissecting the Long-Term Performance of the Chinese Stock Market". In: *The Journal of Finance* 79.2, pp. 993–1054.

Angeletos, George-Marios, Fabrice Collard, and Harris Dellas (2018). "Quantifying confidence". In: *Econometrica* 86.5, pp. 1689–1726.

Angeletos, George-Marios, Fabrice Collard, and Harris Dellas (2020). "Business-cycle anatomy". In: *American Economic Review* 110.10, pp. 3030–3070.

Angeletos, George-Marios and Zhen Huo (2021). "Myopia and anchoring". In: *American Economic Review* 111.4, pp. 1166–1200.

Angeletos, George-Marios, Zhen Huo, and Karthik A Sastry (2021). "Imperfect macroeconomic expectations: Evidence and theory". In: *NBER Macroeconomics Annual* 35.1, pp. 1–86.

Angeletos, George-Marios and Chen Lian (2018). "Forward guidance without common knowledge". In: *American Economic Review* 108.9, pp. 2477–2512.

Augenblick, Ned, Eben Lazarus, and Michael Thaler (2025). "Overinference from weak signals and underinference from strong signals". In: *The Quarterly Journal of Economics* 140.1, pp. 335–401.

Ba, Cuimin, J Aislinn Bohren, and Alex Imas (2024). "Over-and underreaction to information". In: *Available at SSRN 4274617*.

Bardóczy, Bence and Joao Guerreiro (2023). "Unemployment insurance in macroeconomic stabilization with imperfect expectations". In: *Manuscript, April*.

Beaudry, Paul and Franck Portier (2014). "News-driven business cycles: Insights and challenges". In: *Journal of Economic Literature* 52.4, pp. 993–1074.

Bianchi, Francesco, Cosmin Ilut, and Hikaru Saijo (2024a). "Diagnostic business cycles". In: *Review of Economic Studies* 91.1, pp. 129–162.

Bianchi, Francesco, Cosmin L Ilut, and Hikaru Saijo (2024b). *Smooth diagnostic expectations*. Tech. rep. National Bureau of Economic Research.

Bianchi, Francesco, Sydney C Ludvigson, and Sai Ma (2024). *What Hundreds of Economic News Events Say About Belief Overreaction in the Stock Market*. Tech. rep. National Bureau of Economic Research.

Bonaglia, Edoardo, Daniele d'Arienzo, Nunzio Fallico, Nicola Gennaioli, and Luigi Iovino (2025). "The Volatility of Inflation Expectations and Interest Rates". In.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, Matthew O'Brien, and Andrei Shleifer (2024). "Long-Term Expectations and Aggregate Fluctuations". In: *NBER Macroeconomics Annual* 38.1, pp. 311–347.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer (2024). "Belief overreaction and stock market puzzles". In: *Journal of Political Economy* 132.5, pp. 1450–1484.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer (2020). "Overreaction in macroeconomic expectations". In: *American Economic Review* 110.9, pp. 2748–2782.

Cai, Michael (2024). *Explaining the Macroeconomic Inertia Puzzle*. Tech. rep. Working Paper, Northwestern University.

Coibion, Olivier and Yuriy Gorodnichenko (2015). "Information rigidity and the expectations formation process: A simple framework and new facts". In: *American Economic Review* 105.8, pp. 2644–2678.

d'Arienzo, Daniele (2020). "Maturity increasing overreaction and bond market puzzles". In: *Available at SSRN 3733056*.

De la O, Ricardo and Sean Myers (2024). "Which Subjective Expectations Explain Asset Prices?" In: *The Review of Financial Studies* 37.6, pp. 1929–1978.

de Silva, Tim, Eugene Larsen-Hallock, Adam Rej, and David Thesmar (2025). "Expectations Formation with Fat-Tailed Processes: Evidence and Theory". In.

de Silva, Tim and David Thesmar (2024). "Noise in expectations: Evidence from analyst forecasts". In: *The Review of Financial Studies* 37.5, pp. 1494–1537.

Del Negro, Marco (2024). "Can Professional Forecasters Predict Uncertain Times?" In: *Federal Reserve Bank of New York, Liberty Street Economics*.

Driscoll, John C and Aart C Kraay (1998). "Consistent covariance matrix estimation with spatially dependent panel data". In: *Review of economics and statistics* 80.4, pp. 549–560.

Dupor, Bill and M Saif Mehkari (2014). "The analytics of technology news shocks". In: *Journal of Economic Theory* 153, pp. 392–427.

Eva, Kenneth and Fabian Winkler (2023). "A Comprehensive Empirical Evaluation of Biases in Expectation Formation". In.

Farmer, Leland E, Emi Nakamura, and Jón Steinsson (2024). "Learning about the long run". In: *Journal of Political Economy* 132.10, pp. 3334–3377.

Fisher, Jonas DM, Leonardo Melosi, and Sebastian Rast (2025). *Long-run inflation expectations*. Tech. rep. Working Paper.

Gormsen, Niels Joachim and Eben Lazarus (2023). "Duration-driven returns". In: *The Journal of Finance* 78.3, pp. 1393–1447.

H Décaire, Paul and Marius Guenzel (2023). "What Drives Very Long-Run Cash Flow Expectations?" In: *Marius, What Drives Very Long-Run Cash Flow Expectations*.

Kohlhas, Alexandre N and Ansgar Walther (2021). "Asymmetric attention". In: *American Economic Review* 111.9, pp. 2879–2925.

L'Huillier, Jean-Paul, Sanjay R Singh, and Donghoon Yoo (2024). "Incorporating diagnostic expectations into the New Keynesian framework". In: *Review of Economic Studies* 91.5, pp. 3013–3046.

McCarthy, Odhrain and Sebastian Hillenbrand (2021). "Heterogeneous beliefs and stock market fluctuations". In: *Available at SSRN 3944887*.

Mei, Pierfrancesco and Lingxuan Wu (2024). "Thinking about the Economy, Deep or Shallow?" In.

Molavi, Pooya, Alireza Tahbaz-Salehi, and Andrea Vedolin (2024). "Model complexity, expectations, and asset prices". In: *Review of Economic Studies* 91.4, pp. 2462–2507.

Patton, Andrew J and Allan Timmermann (2010). "Why do forecasters disagree? Lessons from the term structure of cross-sectional dispersion". In: *Journal of Monetary Economics* 57.7, pp. 803–820.

Sias, Richard, Laura Starks, and Harry Turtle (2024). "Long-term Beliefs and Financial Choices". In.

Sung, Yeji (2024). "Macroeconomic expectations and cognitive noise". In: Federal Reserve Bank of San Francisco.

Winkler, Fabian (2020). "The role of learning for asset prices and business cycles". In: *Journal of Monetary Economics* 114, pp. 42–58.

A Appendix

A.1 Additional Figures

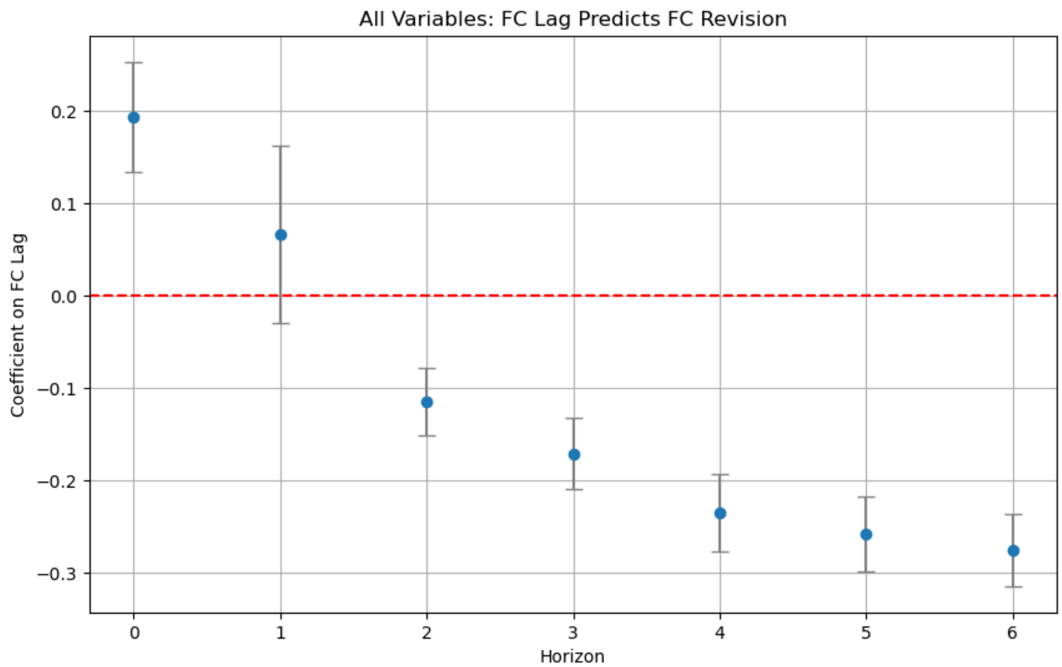
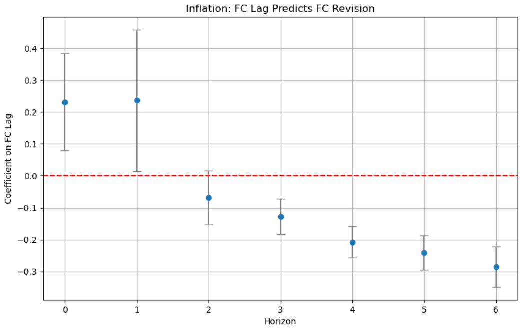
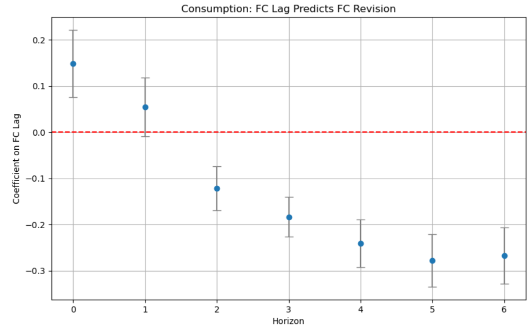


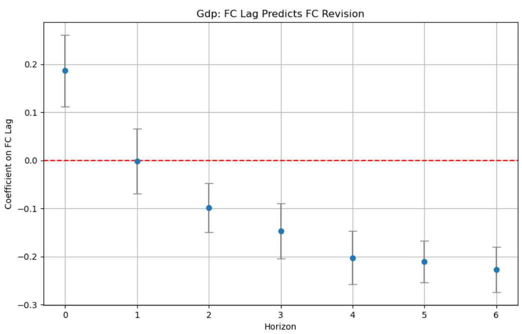
Figure 10: The regression coefficients in the figure are from the regression $\Delta\,\mathbb{E}_t(x_{c,t+h}) = \alpha + \beta\,\mathbb{E}_{t-1}(x_{c,t+h}) + \epsilon_{c,t}$, run separately for each horizon. GDP, inflation, consumption, and investment forecasts are pooled together. The forecast here is the forecast of average annualized growth (of that variable) between t and $t+h$, which is constructed by cumulating the individual-year growth forecasts. All forecasts are z-scored with respect to variable, country, and horizon, with lagged forecasts z-scored using an expanding window that drops the first 10 observations. Standard errors are Driscoll-Kraay with country-variable groupings.



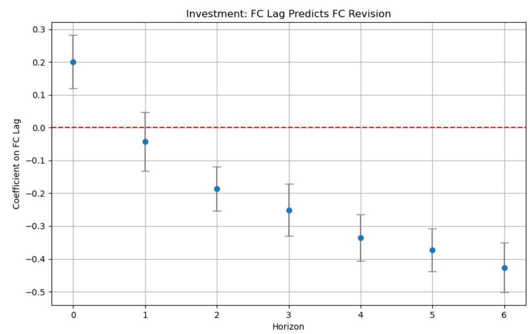
(a) Inflation



(b) Consumption

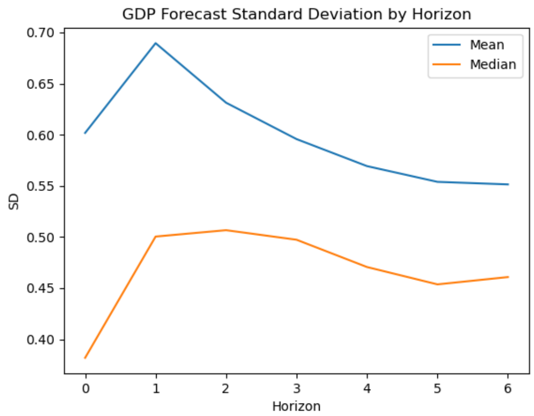


(c) GDP

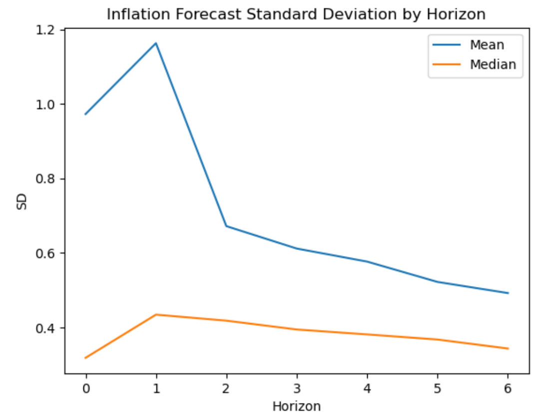


(d) Investment

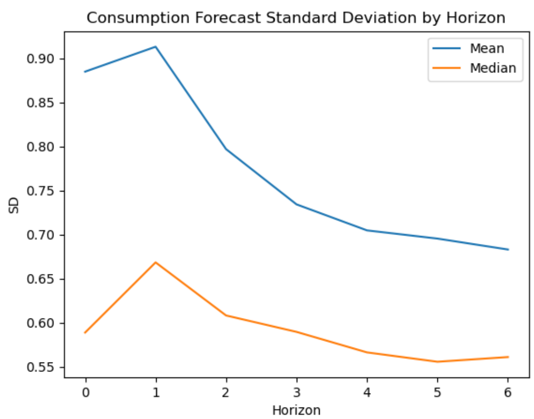
Figure 11: Regression coefficients from $\Delta\mathbb{E}_t[x_{c,t+h}] = \alpha + \beta\mathbb{E}_{t-1}[x_{c,t+h}] + \varepsilon_{c,t}$ for different forecast variables x.



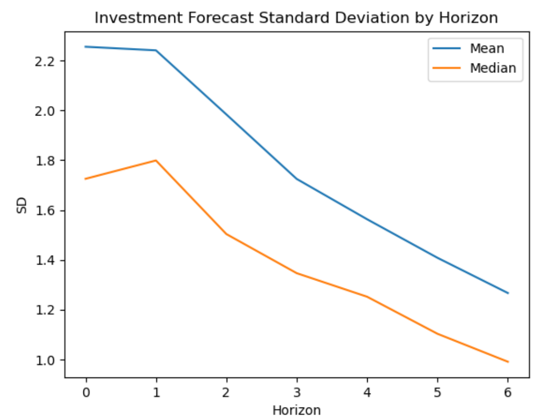
(a) GDP



(b) Inflation



(c) Consumption



(d) Investment

Figure 12: These plots show the mean and median standard deviation of forecasts by horizon and variable.

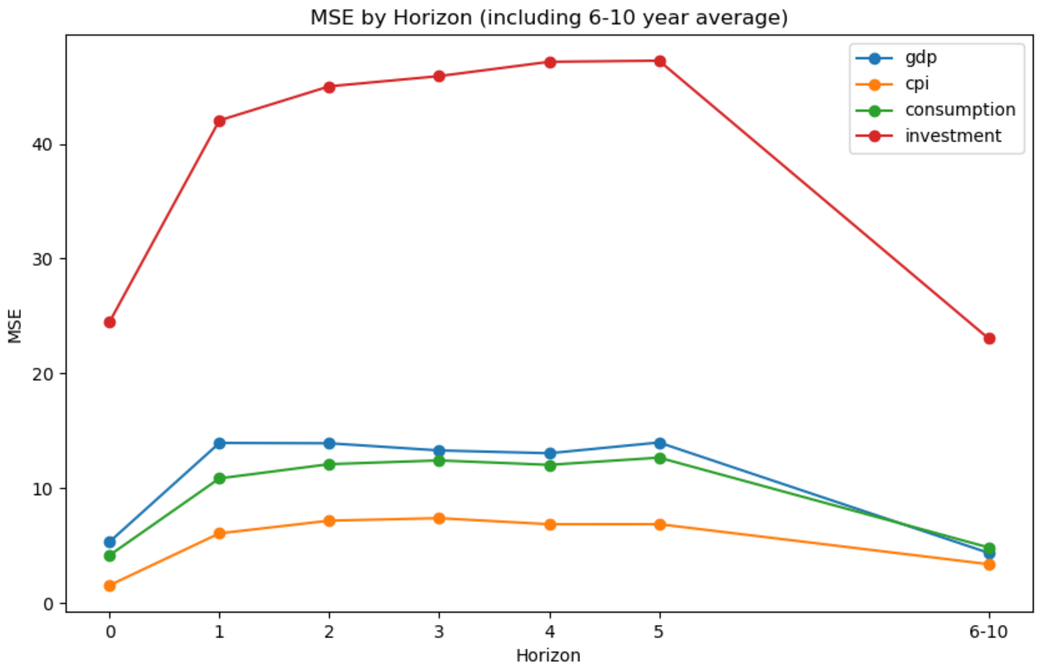


Figure 13: MSE by Horizon

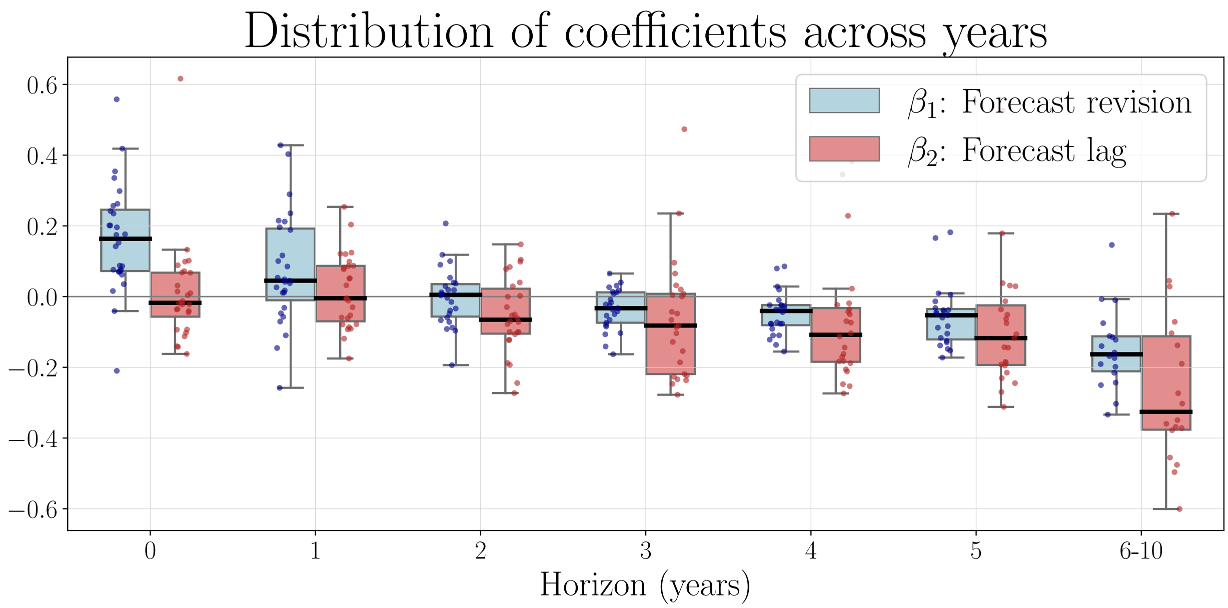
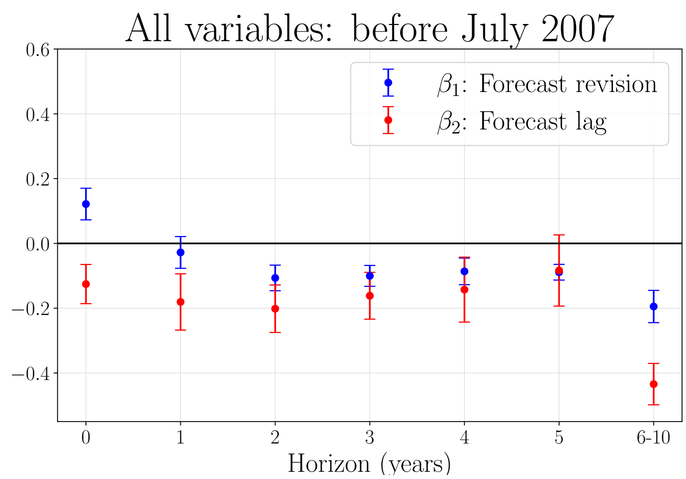
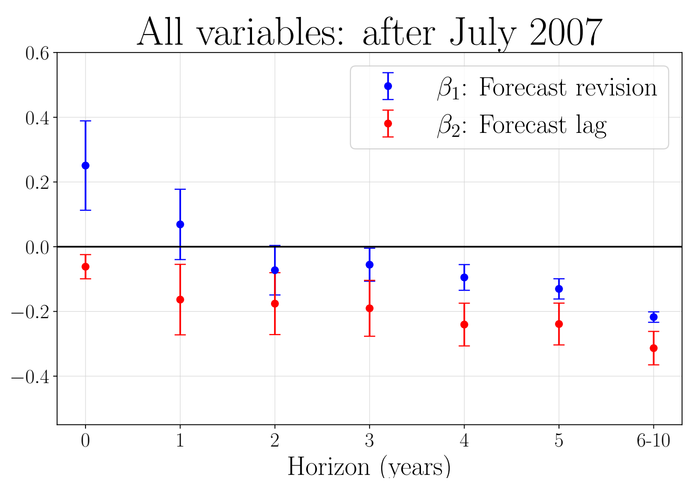


Figure 14: **Coefficients by year:** The figure plots the blue β_1 and red β_2 coefficients from the regression $e_{c,t+h} = \alpha + \beta_1 \Delta\, \mathbb{E}_t(x_{c,t+h}) + \beta_2\, \mathbb{E}_{t-1}(x_{c,t+h}) + f_{c,x} + \epsilon_{c,t}$, where $e_{c,t+h}$ are forecast errors, $\Delta\, \mathbb{E}_t(x_{c,t+h})$ are consecutive survey forecast revisions, and $\mathbb{E}_{t-1}(x_{c,t+h})$ is the previous survey's forecast. Each circle represents the coefficient from running the regression on a single year's worth of data. Covid is removed from the sample. Country-variable fixed effects are included and standard errors are Driscoll-Kraay with country-variable groupings.



(a) Before July 2007



(b) Post July 2007

Figure 15: The figure plots the blue β_1 and red β_2 coefficients from the regression $e_{c,t+h} = \alpha + \beta_1 \Delta\,\mathbb{E}_t(x_{c,t+h}) + \beta_2\,\mathbb{E}_{t-1}(x_{c,t+h}) + f_{c,x} + \epsilon_{c,t}$, where $e_{c,t+h}$ are forecast errors, $\Delta\,\mathbb{E}_t(x_{c,t+h})$ are consecutive survey forecast revisions, and $\mathbb{E}_{t-1}(x_{c,t+h})$ is the previous survey's forecast. The first sub-plot is pre-July 2007 data, the second sub-plot is post-July 2007 data. Covid is removed from the sample. Country-variable fixed effects are included and standard errors are Driscoll-Kraay with country-variable groupings.

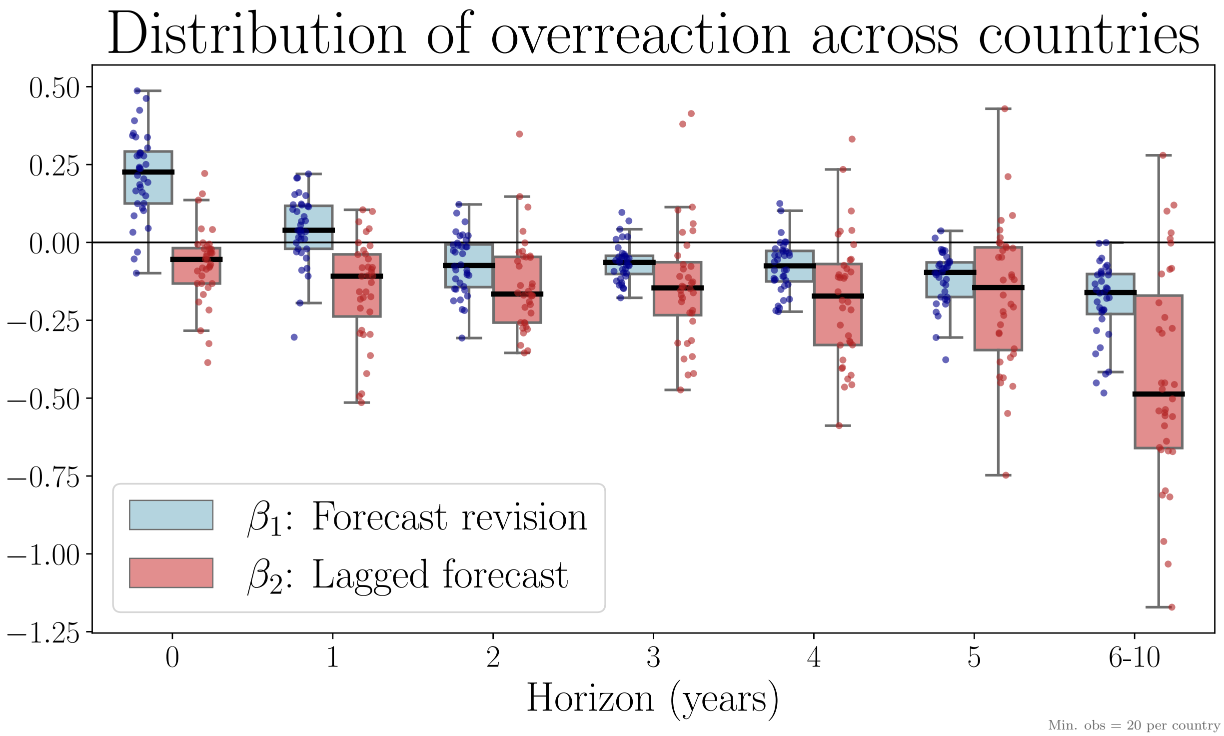


Figure 16: **Coefficients by country:** The figure plots the blue β_1 and red β_2 coefficients from the regression $e_{c,t+h} = \alpha + \beta_1 \Delta\, \mathbb{E}_t(x_{c,t+h}) + \beta_2\, \mathbb{E}_{t-1}(x_{c,t+h}) + f_{c,x} + \epsilon_{c,t}$, where $e_{c,t+h}$ are forecast errors, $\Delta\, \mathbb{E}_t(x_{c,t+h})$ are consecutive survey forecast revisions, and $\mathbb{E}_{t-1}(x_{c,t+h})$ is the previous survey's forecast. Each circle represents the coefficient from an individual country's regression. Covid is removed from the sample. Country-variable fixed effects are included and standard errors are Driscoll-Kraay with country-variable groupings. The sample is restricted to country-horizons with at least 20 observations.

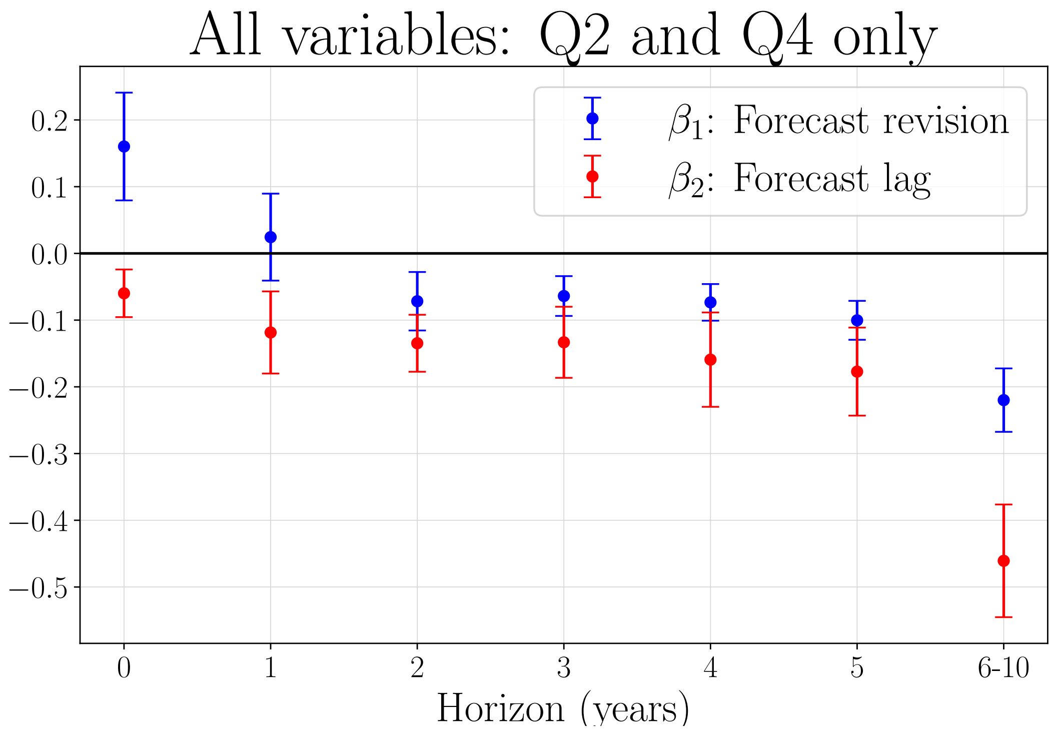


Figure 17: **Q2 and Q4 data only:** The figure plots the blue β_1 and red β_2 coefficients from the regression $e_{c,t+h} = \alpha + \beta_1 \Delta \mathbb{E}_t(x_{c,t+h}) + \beta_2 \mathbb{E}_{t-1}(x_{c,t+h}) + f_{c,x} + \epsilon_{c,t}$, where $e_{c,t+h}$ are forecast errors, $\Delta \mathbb{E}_t(x_{c,t+h})$ are consecutive survey forecast revisions, and $\mathbb{E}_{t-1}(x_{c,t+h})$ is the previous survey's forecast. All forecasts are pooled and are z-scored with respect to variable, country, and horizon, with expanding window z-scores for the lagged forecast. Covid is removed from the sample. Country-variable fixed effects are included and standard errors are Driscoll-Kraay with country-variable groupings.

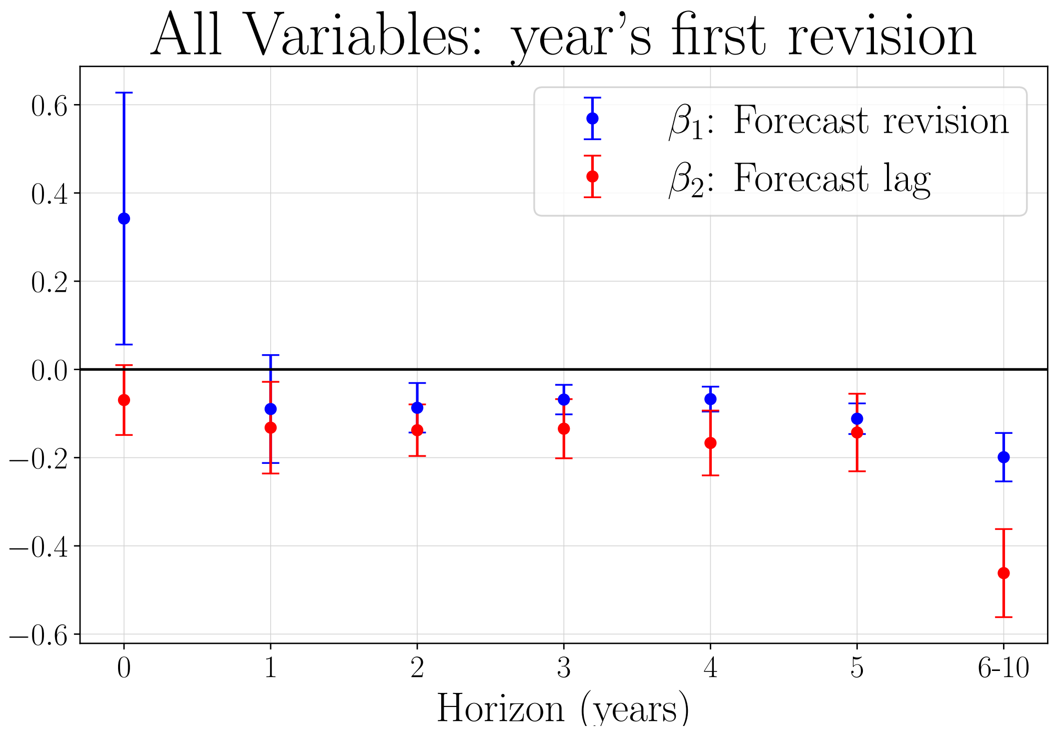


Figure 18: **First revision only:** The figure plots the blue β_1 and red β_2 coefficients from the regression $e_{c,t+h} = \alpha + \beta_1 \Delta\, \mathbb{E}_t(x_{c,t+h}) + \beta_2\, \mathbb{E}_{t-1}(x_{c,t+h}) + f_{c,x} + \epsilon_{c,t}$, where $e_{c,t+h}$ are forecast errors, $\Delta\, \mathbb{E}_t(x_{c,t+h})$ are consecutive survey forecast revisions, and $\mathbb{E}_{t-1}(x_{c,t+h})$ is the previous survey's forecast. All forecasts are pooled and are z-scored with respect to variable, country, and horizon, with expanding window z-scores for the lagged forecast. Covid is removed from the sample. Country-variable fixed effects are included and standard errors are Driscoll-Kraay with country-variable groupings.

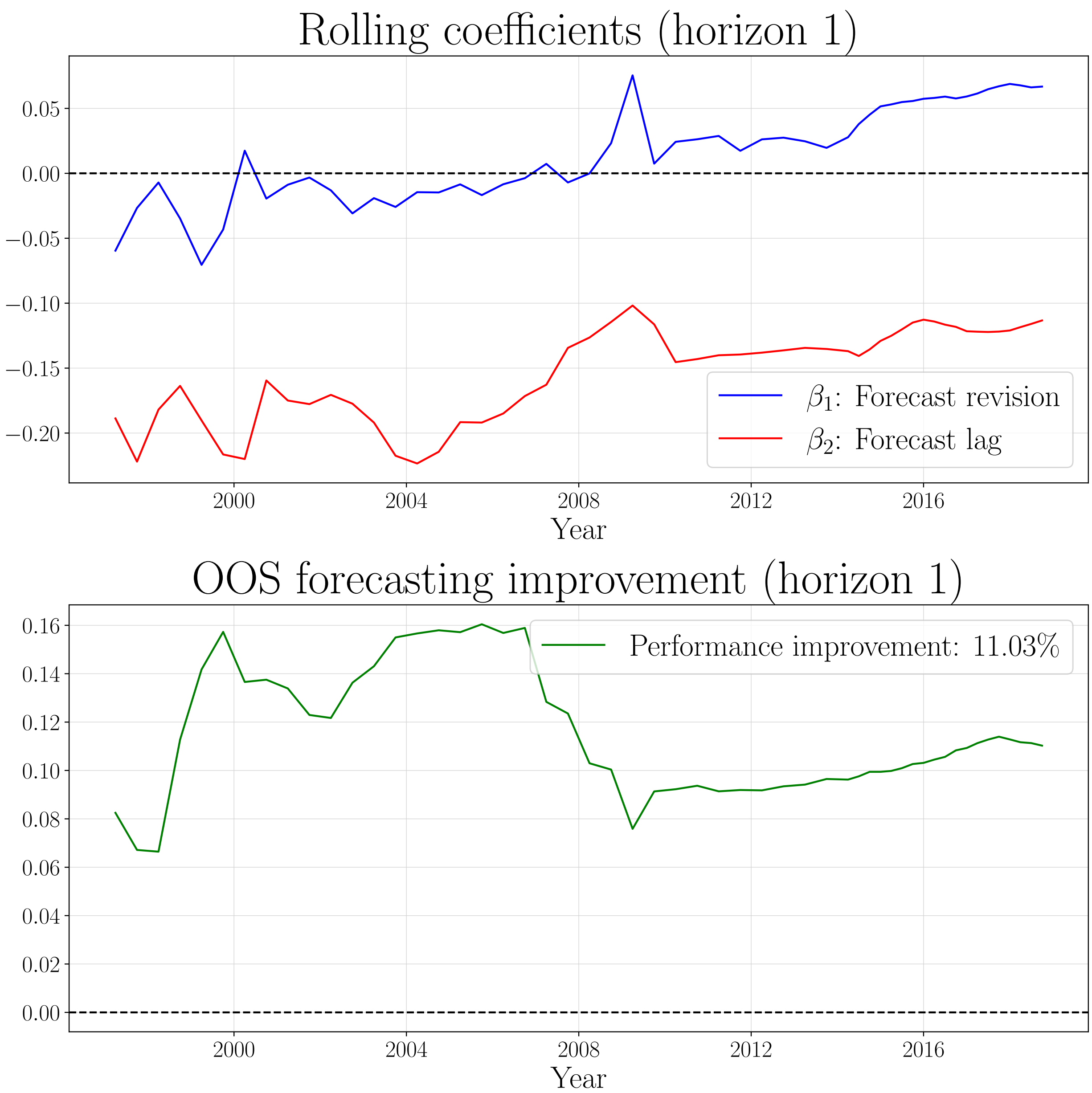


Figure 19: Horizon 1 OOS Forecasting

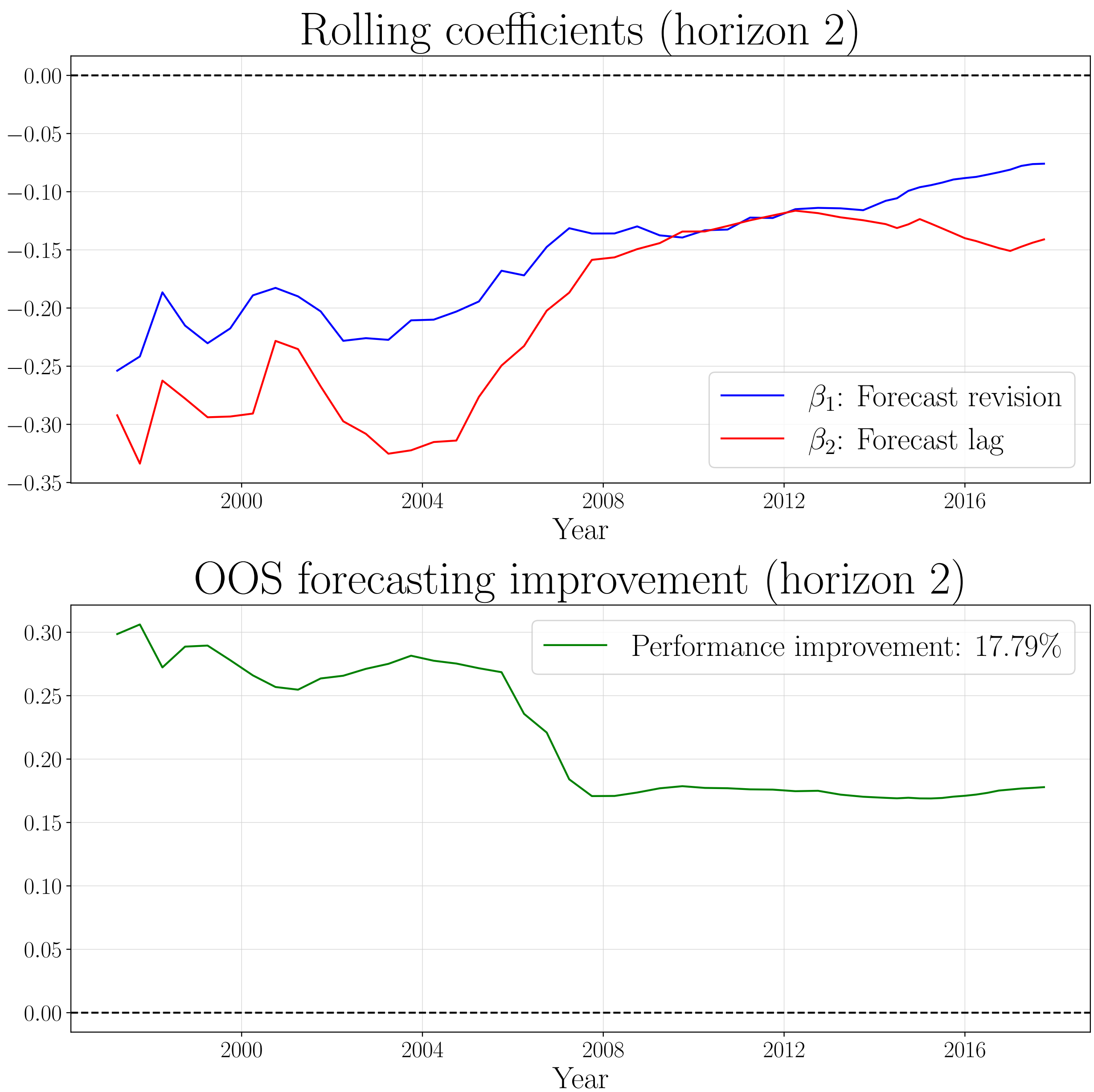


Figure 20: Horizon 2 OOS Forecasting

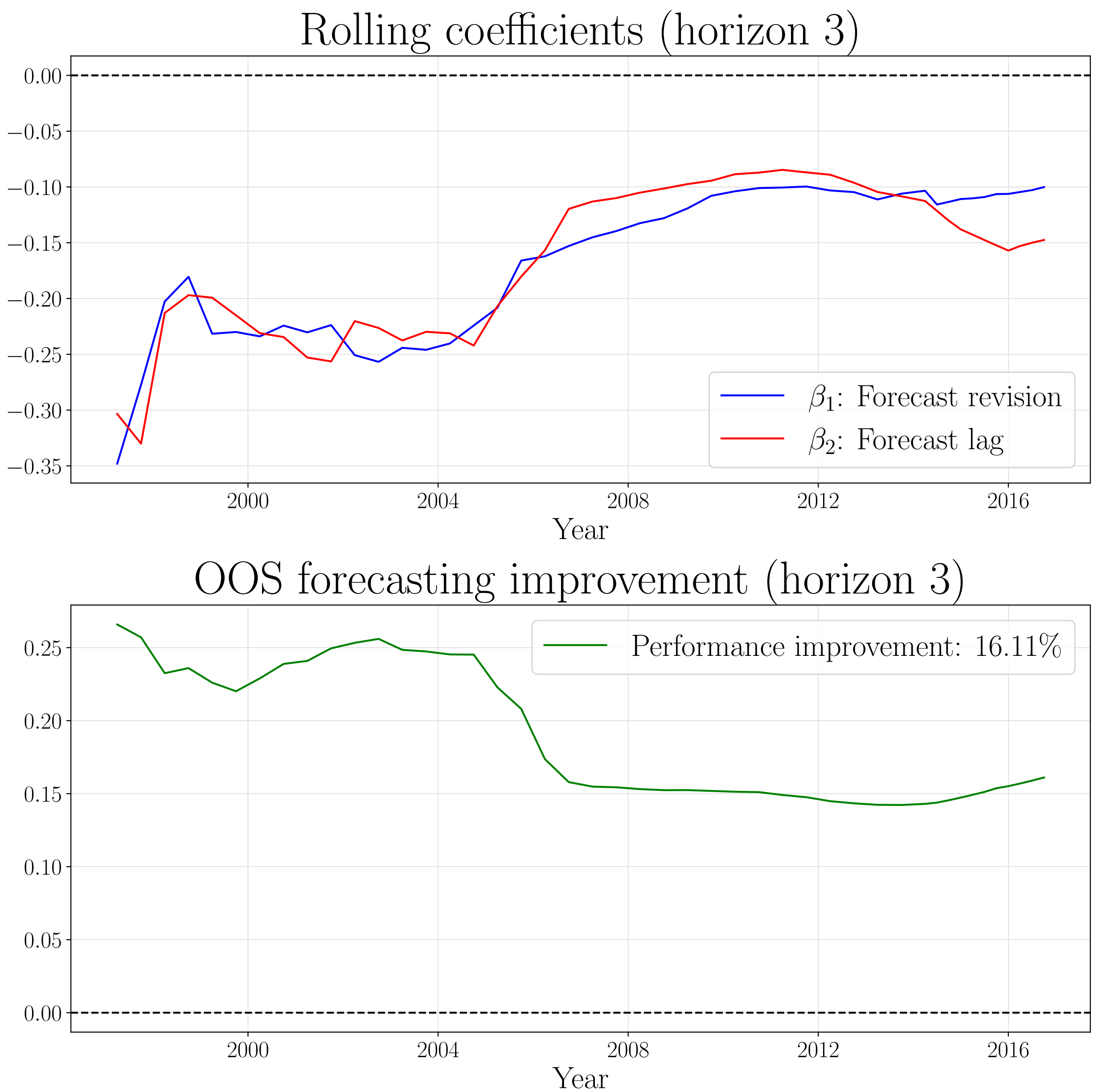


Figure 21: Horizon 3 OOS Forecasting

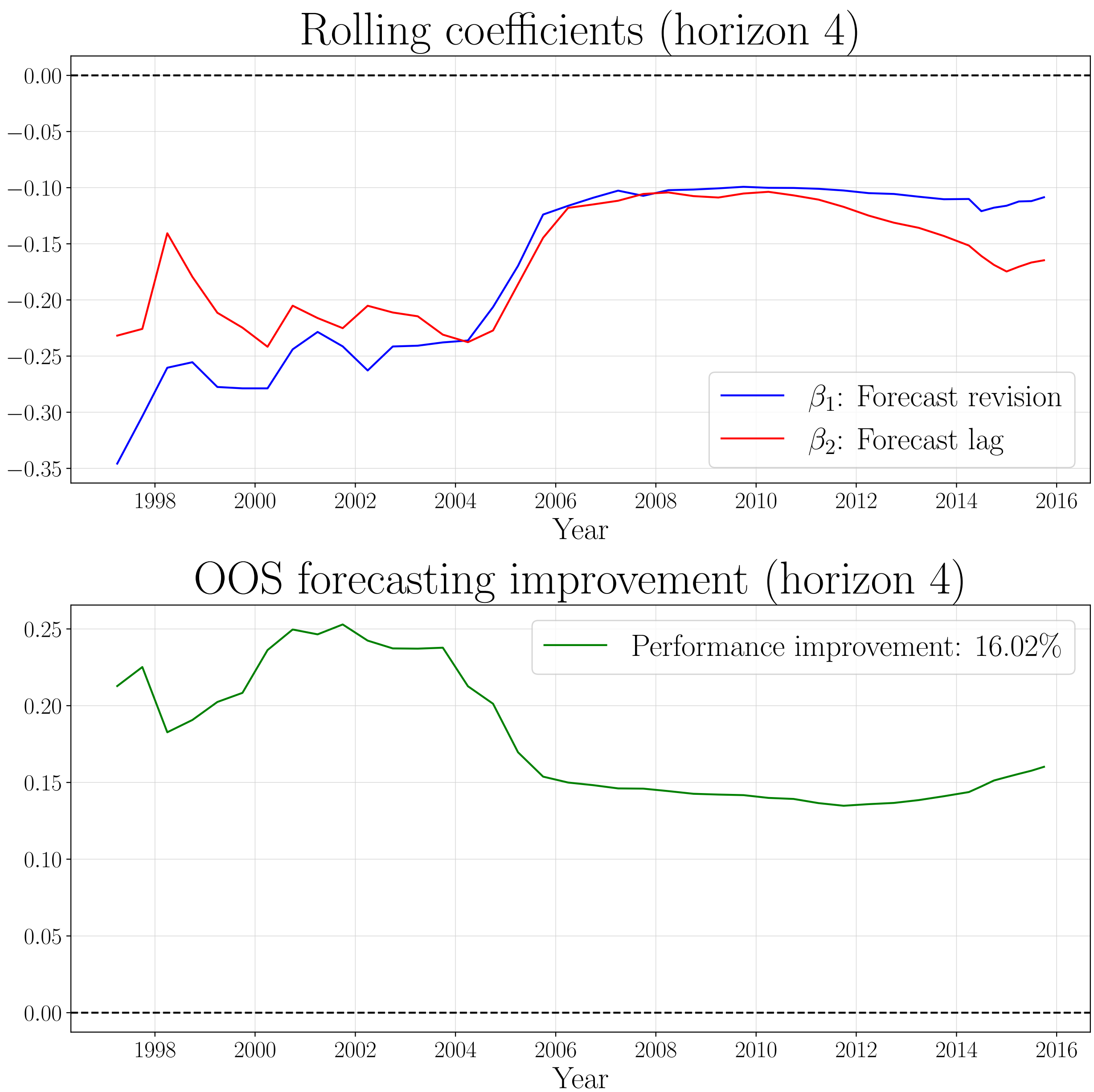
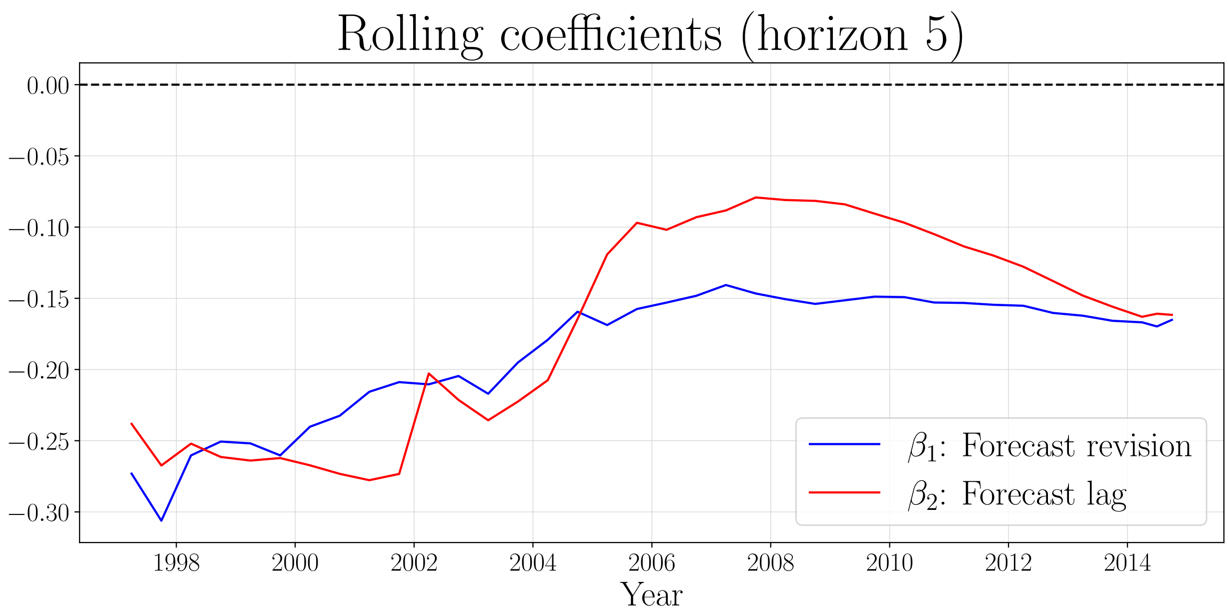


Figure 22: Horizon 4 OOS Forecasting



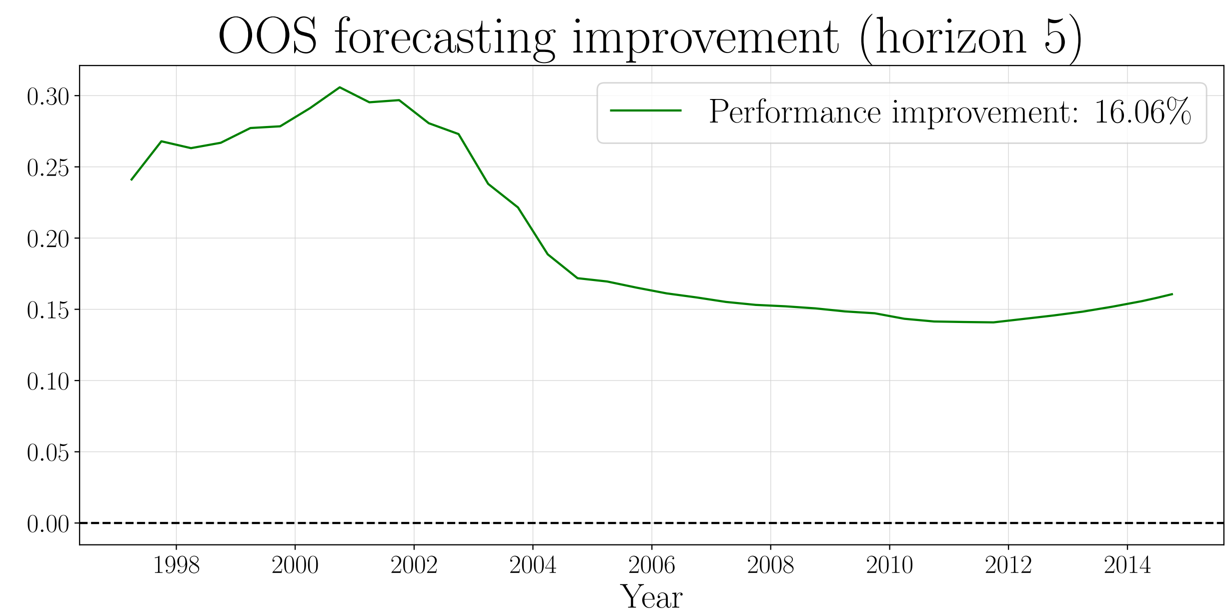
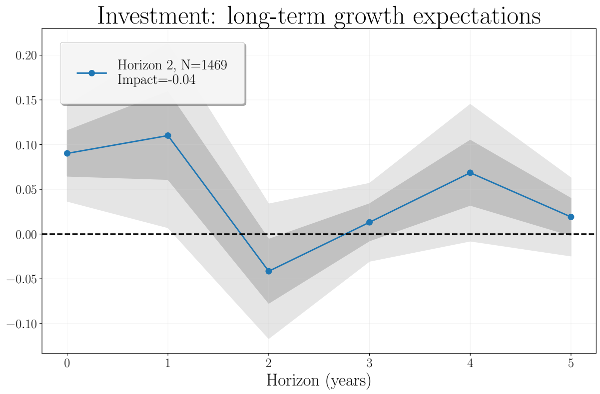
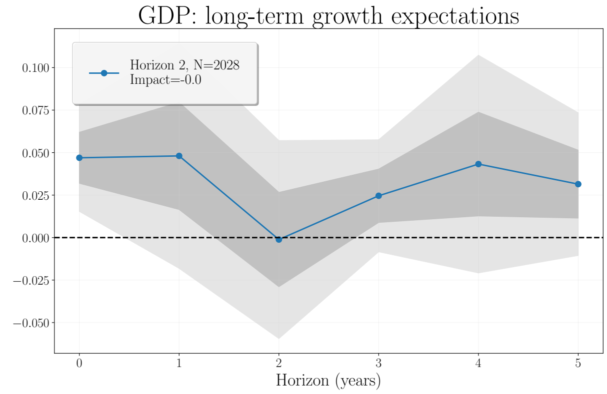


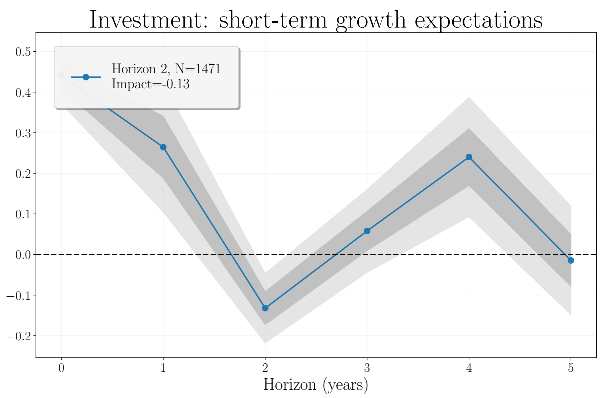
Figure 23: Horizon 5 OOS Forecasting



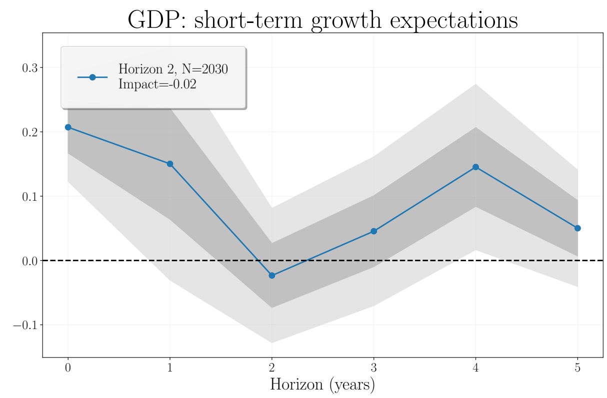
(a) Long-term expectations: Investment



(b) Long-term expectations: GDP



(c) Short-term expectations: Investment



(d) Short-term expectations: GDP

Figure 24: No control regressions: Regression coefficients from $y_{c,t+j} = \alpha + \beta_1 \Delta_1 \mathbb{E}_t(g_{c,t+h}) + \beta_2 \mathbb{E}_t(y_t) + f_c + \epsilon_{c,t}$ for different horizon-forecasts g. In this version of the regression, the only control variable is the current-year forecast $\mathbb{E}_t(y_t)$. The top-row uses the one-year change in six-to-ten year ahead average annual GDP growth expectations; the bottom-row uses the one-year change in average annual GDP growth expectations for the year of the forecast and the following two years. The dependent variable on the LHS column is investment, and on the RHS is GDP. Country fixed effects are included and standard errors are Driscoll-Kraay with country groupings.

A.2 Additional Tables

Table 6: Number of Observations by Country and Variable

Country	Variable			
	GDP	Inflation	Investment	Consumption
Albania	18	18	0	0
Argentina	79	78	79	79
Armenia	18	18	0	0
Australia	84	84	69	84

Continued on next page

Table 6 – continued from previous page

Country	Variable			
	GDP	Inflation	Investment	Consumption
Austria	41	41	0	0
Azerbaijan	18	18	0	0
Bangladesh	18	18	0	0
Belarus	18	18	0	0
Belgium	41	41	0	0
Bolivia	18	18	0	0
Bosnia & Herzegovina	18	18	0	0
Brazil	79	79	79	79
Bulgaria	51	51	51	51
Canada	86	86	69	86
Chile	79	79	79	79
China	76	76	45	0
Colombia	69	69	67	68
Costa Rica	18	18	0	0
Croatia	51	51	51	51
Cyprus	18	18	0	0
Czech Republic	70	70	67	70
Denmark	41	41	0	0
Dominican Republic	18	18	0	0
Ecuador	18	18	0	0
Egypt	22	22	0	0
El Salvador	18	18	0	0
Estonia	51	51	51	51
Euro zone	60	60	56	57
Finland	41	41	0	0
France	86	86	69	86
Georgia	18	18	0	0
Germany	86	86	69	86
Greece	41	41	0	0
Guatemala	18	18	0	0
Honduras	18	18	0	0
Hong Kong	73	73	69	72
Hungary	70	70	67	70
India	76	75	69	34
Indonesia	73	73	68	73
Ireland	41	41	0	0
Israel	22	22	0	0

Continued on next page

Table 6 – continued from previous page

Country	Variable			
	GDP	Inflation	Investment	Consumption
Italy	86	86	69	86
Japan	85	85	68	85
Kazakhstan	18	18	0	0
Kosovo	9	9	0	0
Latvia	51	51	51	51
Lithuania	51	51	51	51
Macedonia	18	18	0	0
Malaysia	76	76	69	75
Mexico	79	79	79	79
Moldova	18	18	0	0
Montenegro	8	8	0	0
Myanmar	18	18	0	0
Netherlands	76	76	68	76
New Zealand	76	76	69	76
Nicaragua	18	18	0	0
Nigeria	22	22	0	0
Norway	66	66	66	66
Pakistan	18	18	0	0
Panama	18	18	0	0
Paraguay	18	18	0	0
Peru	71	71	71	71
Philippines	48	48	48	47
Poland	70	70	67	70
Portugal	41	41	0	0
Romania	70	70	66	69
Russia	70	70	67	70
Saudi Arabia	22	22	0	0
Serbia	18	18	0	0
Singapore	76	76	67	74
Slovakia	70	70	67	70
Slovenia	51	51	51	51
South Africa	22	22	0	0
South Korea	76	76	69	76
Spain	76	76	69	76
Sri Lanka	18	18	0	0
Sweden	76	76	68	75
Switzerland	69	69	69	69

Continued on next page

Table 6 – continued from previous page

Country	Variable			
	GDP	Inflation	Investment	Consumption
Taiwan	76	76	69	76
Thailand	74	73	65	70
Turkey	69	69	66	69
Turkmenistan	18	18	0	0
Ukraine	70	70	48	70
United Kingdom	86	57	69	86
United States	86	86	69	86
Uruguay	18	18	0	0
Uzbekistan	18	18	0	0
Venezuela	75	72	74	74
Vietnam	18	18	15	15
Total	4240	4205	3018	3185

Table 7: USA: Return Predictability Regressions, no Covid

	Return Horizon			
	1-year	3-year	5-year	1-to-5-year
2-year avg GDP growth	-0.21	-0.21	-0.16	-0.12
	(0.19)	(0.14)	(0.14)	(0.13)
	[3.9%]	[3.2%]	[2.0%]	[1.1%]
10-year avg GDP growth	-0.44^{***}	-0.50^{***}	-0.25^{*}	-0.12
	(0.15)	(0.13)	(0.13)	(0.13)
	[16.3%]	[17.8%]	[5.0%]	[1.1%]
6-10 year avg GDP growth	-0.40^{***}	-0.53^{***}	-0.30^{***}	-0.18^{*}
	(0.11)	(0.09)	(0.09)	(0.09)
	[19.3%]	[29.4%]	[10.4%]	[3.7%]

Notes: Standard errors in parentheses, R^2 in square brackets. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 8: 34 countries: Return Predictability Regressions, no Covid

	Return Horizon			
	1-year	3-year	5-year	1-to-5-year
2-year avg GDP growth	-0.36^{***}	-0.22^{*}	-0.28^{***}	-0.07
	(0.11)	(0.12)	(0.07)	(0.08)
	[8.2%]	[3.0%]	[6.0%]	[0.4%]
10-year avg GDP growth	-0.45^{***}	-0.29^{***}	-0.24^{***}	0.04
	(0.09)	(0.10)	(0.06)	(0.06)
	[10.5%]	[4.4%]	[3.7%]	[0.1%]
6-10 year avg GDP growth	-0.32^{***}	-0.24^{***}	-0.11^{*}	0.10
	(0.08)	(0.08)	(0.06)	(0.08)
	[5.3%]	[3.1%]	[0.7%]	[0.6%]

Notes: Standard errors in parentheses, R^2 in square brackets. *** p<0.01, ** p<0.05, * p<0.1

Table 9: Advanced economies: Return Predictability Regressions

	Return Horizon			
	1-year	3-year	5-year	1y4y
2-year avg GDP growth	-0.34^{***}	-0.21	-0.34^{***}	-0.14^{*}
	(0.07)	(0.14)	(0.08)	(0.08)
	[12.2%]	[3.2%]	[9.1%]	[1.5%]
10-year avg GDP growth	-0.25^{**}	-0.19	-0.28^{**}	-0.07
	(0.12)	(0.16)	(0.12)	(0.15)
	[5.8%]	[3.0%]	[5.9%]	[0.4%]
6-10 year avg GDP growth	-0.08	-0.14	-0.16	$6.75e-03$
	(0.14)	(0.17)	(0.15)	(0.18)
	[0.5%]	[1.5%]	[1.6%]	[0.0%]

Notes: Standard errors in parentheses, R^2 in square brackets. *** p<0.01, ** p<0.05, * p<0.1

Table 10: Emerging markets: Return Predictability Regressions

	Return Horizon			
	1-year	3-year	5-year	1y4y
2-year avg GDP growth	-0.33^{***}	-0.22^{**}	-0.23^{***}	$5.91e-03$
	(0.07)	(0.10)	(0.07)	(0.10)
	[11.4%]	[4.3%]	[4.8%]	[0.0%]
10-year avg GDP growth	-0.20^{***}	-0.14	-0.15	0.03
	(0.08)	(0.11)	(0.11)	(0.13)
	[4.0%]	[1.7%]	[2.0%]	[0.1%]
6-10 year avg GDP growth	-0.03	-0.02	-0.02	0.11
	(0.10)	(0.13)	(0.14)	(0.15)
	[0.1%]	[0.0%]	[0.0%]	[0.9%]

Notes: Standard errors in parentheses, R^2 in square brackets. *** p<0.01, ** p<0.05, * p<0.1

A.3 Proofs

A.3.1 Proposition 1

In the regression on forecast revisions and lagged forecasts,

(i) $\beta_1^h = \beta_2^h = -\frac{\Delta_h}{b_h} \leq 0$

(ii) $\frac{d\beta_1(h)}{dh} = \frac{d\beta_2(h)}{dh} < 0 \quad \text{if} \quad \gamma \geq 1.$

Where $b_h = \rho^h + \Delta_h$.

Proof. For convenience, we re-write 4 in demeaned form (does not affect results):

$$u_t = \rho u_{t-1} + \nu_t \quad \nu_t \sim (0, \sigma_\nu^2) \tag{8}$$

Costly recall implies

$$F_t(h) = \rho^h u_t + \Delta^h u_t, \qquad \Delta^h > 0, \qquad b \equiv \rho^h + \Delta^h > 0. \tag{9}$$

We can then characterize forecast errors, revisions, and lagged forecasts as follows:

$$FE_{t+h} = -\Delta_h u_t + \sum_{j=0}^{h-1} \rho^{h-1-j} \nu_{t+1+j},$$
$$R_t(h) = b_h(u_t - u_{t-1}),$$

$$L_{t-1}(h) = b_h u_{t-1}.$$

The shock sum is uncorrelated with all t and $t-1$ variables.

Second moments

$$\begin{aligned}
\text{Var}\big(R_t(h)\big) &= 2b_h^2(1-\rho)\sigma_u^2, \\
\text{Var}\big(L_{t-1}(h)\big) &= b_h^2\sigma_u^2, \\
\text{Cov}\big(R_t(h), L_{t-1}(h)\big) &= -b_h^2(1-\rho)\sigma_u^2, \\
\text{Cov}\big(FE_{t+h}, R_t(h)\big) &= -\Delta_h b_h(1-\rho)\sigma_u^2, \\
\text{Cov}\big(FE_{t+h}, L_{t-1}(h)\big) &= -\Delta_h b_h \rho\sigma_u^2.
\end{aligned}$$

Derivation of the OLS coefficients

Denominator

$$\text{Var}(R)\ \text{Var}(L) = \left[2b_h^2(1-\rho)\sigma_u^2\right]\left[b_h^2\sigma_u^2\right] = 2b_h^4(1-\rho)\sigma_u^4. \tag{10}$$

$$\text{Cov}(R,L)^2 = \left[-b_h^2(1-\rho)\sigma_u^2\right]^2 = b_h^4(1-\rho)^2\sigma_u^4. \tag{11}$$

$$D \equiv \text{Var}(R)\,\text{Var}(L) - \text{Cov}(R,L)^2 = b_h^4\sigma_u^4(1-\rho)(1+\rho) = b_h^4\sigma_u^4(1-\rho^2) > 0. \tag{12}$$

Numerator for β_1

$$\text{Cov}(FE,R)\ \text{Var}(L) = \left[-\Delta_h\, b_h(1-\rho)\sigma_u^2\right]\left[b_h^2\sigma_u^2\right] = -\Delta_h\, b_h^3(1-\rho)\sigma_u^4. \tag{13}$$

$$\text{Cov}(FE,L)\ \text{Cov}(R,L) = \left[-\Delta_h\, b_h\rho\sigma_u^2\right]\left[-b_h^2(1-\rho)\sigma_u^2\right] = \Delta_h\, b_h^3\rho(1-\rho)\sigma_u^4. \tag{14}$$

$$N_1 = -\Delta_h\, b_h^3(1-\rho)\sigma_u^4 - \Delta_h\, b_h^3\rho(1-\rho)\sigma_u^4 = -\Delta_h\, b_h^3(1-\rho^2)\sigma_u^4. \tag{15}$$

Numerator for β_2

$$\text{Cov}(FE,L)\ \text{Var}(R) = \left[-\Delta_h\, b_h\rho\sigma_u^2\right]\left[2b_h^2(1-\rho)\sigma_u^2\right] = -2\Delta_h\, b_h^3\rho(1-\rho)\sigma_u^4. \tag{16}$$

$$\text{Cov}(FE,R)\ \text{Cov}(R,L) = \left[-\Delta_h\, b_h(1-\rho)\sigma_u^2\right]\left[-b_h^2(1-\rho)\sigma_u^2\right] = \Delta_h\, b_h^3(1-\rho)^2\sigma_u^4. \tag{17}$$

$$N_2 = -2\Delta_h\, b_h^3\rho(1-\rho)\sigma_u^4 - \Delta_h\, b_h^3(1-\rho)^2\sigma_u^4 = -\Delta_h\, b_h^3(1-\rho^2)\sigma_u^4. \tag{18}$$

Coefficients

$$\boxed{\beta_1(1) = \beta_2(1) = \frac{N_1}{D} = -\frac{\Delta_h}{\rho^h + \Delta_h} \ < \ 0.} \tag{19}$$

When the min constraint does not bind in 6, we have $\Delta_h = (1-\rho^h)\,\kappa_h$, where κ_h is:

$$\kappa_h = (\omega\tau)^{\frac{1}{1+\gamma}}\,(1-\rho^h)^{-\frac{2}{1+\gamma}}. \tag{20}$$

Define $g(h) \equiv 1-\rho^h$ (note $g' > 0$ because $\ln\rho < 0$). Then

$$\Delta_h = g\,\kappa_h = C\,g^{\frac{\gamma-1}{1+\gamma}}, \qquad C \equiv (\omega\tau)^{\frac{1}{1+\gamma}} > 0. \tag{21}$$

$$\frac{d\Delta_h}{dh} = C\,\frac{\gamma-1}{1+\gamma}\,g^{\frac{\gamma-1}{1+\gamma}-1}\,g' \begin{cases} > 0 & \gamma > 1, \\ = 0 & \gamma = 1, \\ < 0 & \gamma < 1. \end{cases} \tag{22}$$

Hence $d\Delta_h/dh \geq 0$ when $\gamma \geq 1$..

Derivative of $\beta(h)$ Using $\beta(h) = -\Delta_h/b_h$,

$$\beta'(h) = -\frac{\Delta_h'}{b_h} + \frac{\Delta_h\,b_h'}{b_h^2} = -\frac{\Delta_h'}{b_h} + \frac{\Delta_h\,\rho^h \ln\rho}{b_h^2}. \tag{23}$$

Sign when $\gamma \geq 1$ If $\gamma \geq 1$ then $\Delta_h' > 0$ by (22). Because $\ln\rho < 0$ and all other factors are positive, the second term in (23) is *negative*. Both terms therefore contribute the same sign, giving

$$\boxed{\frac{d\beta_1(h)}{dh} = \frac{d\beta_2(h)}{dh} < 0 \quad \text{if} \quad \gamma \geq 1.} \tag{24}$$

□

A.3.2 Proposition 2

If agents observe noisy signals s_t of x_t and update their posterior beliefs about x_t in a Bayesian manner,

(i) $\beta_1(h) < \beta_2(h) \quad \forall\, h$

If only a fraction $\lambda < 1$ of agents update their forecasts each period,

(i) $\beta_1(h) > \beta_2(h) \quad \forall h$

(ii) $\beta_1(h) > 0$ is possible, while $\beta_2(h) \leq 0 \quad \forall h$

Proof. We begin with the noisy signal case:

Notation recap

- Signal: $s_t = x_t + \varepsilon_t,\ \varepsilon_t \overset{iid}{\sim} (0, q)$.
- Kalman gain: $\kappa_0 := \dfrac{\tau_\varepsilon}{\tau_0 + \tau_\varepsilon} \in (0, 1)$.
- Overreaction term: $\kappa_h := \min\Big\{1, (\omega\tau_\varepsilon/(1-\rho^h)^2)^{1/(1+\gamma)}\Big\} \in (0, 1]$.
- Demeaned state: $u_t := x_t - \mu$.
- Posterior: $\hat{u}_t = \kappa_0(u_t + \varepsilon_t)$.

We proceed from appendix E in Afrouzi, Kwon, Landier, Ma, and Thesmar (2023) where they derive the forecasting rule for the noisy signal case:

$$F_t x_{t+h} = \rho^h \hat{x}_t + \big(1-\rho^h\big)\kappa_h \hat{x}_t + \varepsilon_t,$$

where $\hat{x}_t = \mu + \kappa_0(x_t - \mu) + \kappa_0\varepsilon_t,\ \kappa_0 := \tau_\varepsilon/(\tau_0 + \tau_\varepsilon) \in (0, 1)$.

Insert $\hat{x}_t$ and collect x_t

$$\begin{aligned} F_t x_{t+h} = {} & \underbrace{\big[\rho^h\kappa_0 + (1-\rho^h)\kappa_h\kappa_0\big]}_{W_1}\ x_t \\ & + \underbrace{\rho^h(1-\kappa_0)}_{W_2}\,\mu + \underbrace{(1-\rho^h)\kappa_h(1-\kappa_0)}_{W_3}\,\mu \\ & + \underbrace{\big[\rho^h\kappa_0 + (1-\rho^h)\kappa_h\kappa_0 + 1\big]}_{W_\varepsilon}\ \varepsilon_t. \end{aligned}$$

Add-subtract $\rho^h x_t$

$$F_t x_{t+h} = \rho^h x_t + \big[\underbrace{\kappa_0(1-\rho^h)\kappa_h}_{\text{over-reaction}} - \underbrace{(1-\kappa_0)\rho^h}_{\text{under-reaction}}\big] x_t + W_\epsilon \varepsilon_t.$$

Redefine b_h:

$$b_h := \kappa_0\kappa_h(1-\rho^h) + \kappa_0\rho^h, \qquad (\text{note } b_h < \rho^h\kappa_0 \text{ when } \kappa_0 < 1).$$

Then

$$\boxed{F_t^{NS}(h) = b_h\, u_t + W_\epsilon \epsilon_t.}$$

Regression ingredients

$$FE^{NS}_{t+h} = u_{t+h} - F^{NS}_t(h) = \big[\rho^h - b_h\big]u_t - W_\epsilon \varepsilon_t + \sum_{j=1}^{h} \rho^{h-j} \nu_{t+h}, \tag{25}$$

$$R^{NS}_t(h) = F^{NS}_t(h) - F^{NS}_{t-1}(h) = b_h(u_t - u_{t-1}) + W_\epsilon(\varepsilon_t - \varepsilon_{t-1}), \tag{26}$$

$$L^{NS}_{t-1}(h) = b_h\, u_{t-1} + W_\epsilon \varepsilon_{t-1}. \tag{27}$$

Denote $\delta_h := \rho^h - b_h$, $q = \sigma^2_\epsilon$

Second moments

$$\begin{aligned}
\operatorname{Var}(R) &= b_h^2 \underbrace{\operatorname{Var}(u_t - u_{t-1})}_{2(1-\rho)A} + \operatorname{Var}(\varepsilon_t - \varepsilon_{t-1}) = 2b_h^2(1-\rho)A + 2W_\epsilon^2 q, && (28)\\
\operatorname{Var}(L) &= b_h^2 A + W_\epsilon^2 q, && (29)\\
\operatorname{Cov}(R, L) &= b_h^2\big[\operatorname{Cov}(u_t, u_{t-1}) - \operatorname{Var}(u_{t-1})\big] + \operatorname{Cov}(\varepsilon_t - \varepsilon_{t-1}, \varepsilon_{t-1})\\
&= b_h^2(\rho A - A) - W_\epsilon^2 q = -b_h^2(1-\rho)A - W_\epsilon^2 q, && (30)\\
\operatorname{Cov}(FE, R) &= \operatorname{Cov}\big(\delta_h u_t - \varepsilon_t,\ b_h(u_t - u_{t-1}) + \varepsilon_t - \varepsilon_{t-1}\big)\\
&= \delta_h b_h\big[\operatorname{Var}(u_t) - \operatorname{Cov}(u_t, u_{t-1})\big] - \operatorname{Var}(\varepsilon_t)\\
&= \delta_h b_h(1-\rho)A - W_\epsilon^2 q, && (31)\\
\operatorname{Cov}(FE, L) &= \operatorname{Cov}\big(\delta_h u_t - \varepsilon_t,\ b_h u_{t-1} + \varepsilon_{t-1}\big) = \delta_h b_h \rho A. && (32)
\end{aligned}$$

OLS denominator

$$\begin{aligned}
D^{NS} &= \operatorname{Var}(R)\ \operatorname{Var}(L) - \operatorname{Cov}(R, L)^2\\
&= \big[2b_h^2(1-\rho)A + 2W_\epsilon^2 q\big]\ \big[b_h^2 A + W_\epsilon^2 q\big] - \big[-b_h^2(1-\rho)A - W_\epsilon^2 q\big]^2\\
&= \underbrace{b_h^4 A^2(1-\rho^2)}_{\text{all-}u\text{ terms}} + \underbrace{2b_h^2 W_\epsilon^2 q A}_{\text{one } q} + \underbrace{W_\epsilon^4 q^2}_{\text{two } q} > 0.
\end{aligned} \tag{S.2}$$

Numerator N_1^{NS}

$$\begin{aligned}
N_1^{NS} &= \operatorname{Cov}(FE, R)\ \operatorname{Var}(L) - \operatorname{Cov}(FE, L)\ \operatorname{Cov}(R, L)\\
&= \big[\delta_h b_h(1-\rho)A - W_\epsilon^2 q\big]\big[b_h^2 A + W_\epsilon^2 q\big]\\
&\quad - \big[\delta_h b_h \rho A\big]\big[-b_h^2(1-\rho)A - W_\epsilon^2 q\big]\\
&= \delta_h b_h^3 A^2(1-\rho) + \delta_h b_h(1-\rho)AW_\epsilon^2 q - b_h^2 A W_\epsilon^2 q - W_\epsilon^4 q^2\\
&\quad + \delta_h b_h^3 \rho A^2(1-\rho) + \delta_h b_h \rho A W_\epsilon^2 q\\
&= \boxed{\delta_h b_h^3 A^2(1-\rho^2) + \delta_h b_h A W_\epsilon^2 q - b_h^2 A W_\epsilon^2 q - W_\epsilon^4 q^2}.
\end{aligned} \tag{S.3}$$

Numerator N_2^{NS}

$$\begin{aligned}
N_2^{NS} &= \mathrm{Cov}(FE, L)\ \mathrm{Var}(R) - \mathrm{Cov}(FE, R)\ \mathrm{Cov}(R, L) \\
&= \left[\delta_h b_h \rho A\right]\left[2b_h^2(1-\rho)A + 2W_\epsilon^2 q\right] \\
&\quad - \left[\delta_h b_h(1-\rho)A - W_\epsilon^2 q\right]\left[-b_h^2(1-\rho)A - W_\epsilon^2 q\right] \\
&= 2\delta_h b_h^3 \rho A^2(1-\rho) + 2\delta_h b_h \rho A W_\epsilon^2 q \\
&\quad + \delta_h b_h^3(1-\rho)^2 A^2 + \delta_h b_h(1-\rho)AW_\epsilon^2 q - b_h^2(1-\rho)AW_\epsilon^2 q - W_\epsilon^4 q^2 \\
&= \boxed{\delta_h b_h^3 A^2(1-\rho^2) + \delta_h b_h A W_\epsilon^2 q(1+\rho) - b_h^2 A W_\epsilon^2 q(1-\rho) - W_\epsilon^4 q^2}.
\end{aligned}$$

Ordering

$$\begin{aligned}
N_1^{NS} - N_2^{NS} &= -\rho\, b_h W_\epsilon^2 q A\ \left(b_h + \delta_h\right) \\
&< 0 \qquad \left(0 < \rho < 1,\ b_h, \delta_h, q, A, W_\epsilon^2 > 0\right),
\end{aligned}$$

so

$$\boxed{\beta_1^{NS}(h) < \beta_2^{NS}(h) \quad \forall\, h,\ q > 0.}$$

Sticky information If a fraction λ update their forecasts each period, the forecast is given by:

$$\boxed{F_t(h)\ =\ E_t u_{t+h}\ =\ \lambda \sum_{k=0}^{\infty} \pi^k\, b_{h+k}\, u_{t-k}, \quad 0 < \pi := 1 - \lambda < 1.} \tag{D0}$$

Here b_{h+k} is the costly–recall coefficient for a forecast made k periods ago: $b_h = \rho^h + \Delta_h$. Define the cohort tail

$$w_k := \pi^k b_{h+k}, \quad d_t := \sum_{k=1}^{\infty} w_k\, u_{t-k}. \tag{D1}$$

Re-write the current and lagged forecast:

$$F_t(h) = \sum_{k=0}^{\infty} \underbrace{\lambda \pi^k b_{h+k}}_{A_k} u_{t-k}, \qquad F_{t-1}(h) = \sum_{k=0}^{\infty} \underbrace{\lambda \pi^k b_{h+1+k}}_{B_k} u_{t-1-k}. \tag{D b}$$

In terms of d_t

$$F_t(h) = \lambda b_h u_t + \lambda d_t, \tag{D2}$$

Re-index the second sum

Shift the forecast made at $t-1$ so all terms are u_{t-j}:

$$F_{t-1}(h) = \sum_{j=1}^{\infty} B_{j-1}\, u_{t-j} = \sum_{j=1}^{\infty} \frac{A_j}{\pi}\, u_{t-j} \quad (\text{by } B_{j-1} = A_j/\pi). \tag{D c}$$

Subtract term-by-term

$$\begin{aligned} R_t(h) &:= F_t(h) - F_{t-1}(h) \\ &= A_0\, u_t + \sum_{j=1}^{\infty} \Big(A_j - \frac{A_j}{\pi}\Big) u_{t-j} \\ &= \lambda b_h u_t \ - \ \lambda \frac{\lambda}{\pi} \sum_{j=1}^{\infty} \pi^j b_{h+j} u_{t-j}. \end{aligned} \tag{D d}$$

For clarity define $\alpha_0 := \lambda b_h,\ \gamma := \lambda^2/\pi,\ w_j := \pi^j b_{h+j}\ (j \geq 1)$.

$$\boxed{R_t(h) = \alpha_0\, u_t - \gamma \sum_{j=1}^{\infty} w_j\, u_{t-j}} \tag{D e}$$

Lagged belief

$$\boxed{L_{t-1}(h) = F_{t-1}(h) = \frac{\lambda}{\pi}\, d_t = \frac{\lambda}{\pi} \sum_{j=1}^{\infty} w_j\, u_{t-j}.} \tag{D f}$$

Forecast error is given by:

Realised state: $u_{t+h} = \rho^h u_t + \sum_{j=1}^{h} \rho^{h-j} \nu_{t+h}$. Subtract $F_t(h)$ and discard the orthogonal shock sum:

$$\boxed{\mathrm{FE}^*_{t+h} := u_{t+h} - F_t(h) - \sum_{j=1}^{h} \rho^{h-j} \nu_{t+h} = \underbrace{(\rho^h - \lambda b_h)}_{\delta_1} u_t - \lambda d_t.} \tag{D g}$$

Two–dimensional representation

$$(R_t,\ L_{t-1},\ \mathrm{FE}^*)\ \in\ \mathrm{span}\{u_t,\ d_t\}.$$

Two-dimensional moment matrix Let

$$\mathbf{c} := (u_t, d_t)^\top, \quad \Sigma := \mathrm{Var}(\mathbf{c}) = A \begin{pmatrix} 1 & S \\ S & T \end{pmatrix}, \quad S := \frac{\mathrm{Cov}(u_t, d_t)}{A},\ T := \frac{\mathrm{Var}(d_t)}{A}. \tag{C2.2}$$

$$C := \begin{pmatrix} \lambda b_h & -\lambda^2/\pi \\ 0 & \lambda/\pi \end{pmatrix}, \quad \mathbf{z} := C\mathbf{c} = \begin{pmatrix} R_t \\ L_{t-1} \end{pmatrix}, \quad \mathbf{f}^\top := (\delta_1, -\lambda). \tag{C2.3}$$

In order to derive closed-form coefficients, we work with a sympy script (available in online appendix):

$$\beta = \left(C\Sigma C^{\top}\right)^{-1} C\Sigma \mathbf{f} \implies \boxed{\beta_1(h) = -1 + \frac{\rho^h}{b_h \lambda}, \;\; \beta_2(h) = -1 + \frac{\rho^h}{b_h}.} \tag{C2.4}$$

Ordering. We have $\beta_1(h) > \beta_2(h)$ for every $0 < \lambda < 1$.

When $h = 0$, $\beta_1(0) = -1 + \frac{1}{\lambda} > 0$ while $\beta_2(0) = -1 + 1 = 0$. For $h > 0$, $\rho^h < b_h$, so $\beta_2(h) < 0$.

□