Frictional Forecasts*

Isaac Baley Lukas Hack Lora Pavlova

Davud Rostam-Afschar Javier Turen

August 10, 2026

Abstract

We argue that reported predictions of professional forecasters are not pure measures of true beliefs because they are subject to reporting frictions; they constitute frictional forecasts. First, forecasts are lumpy. Zero revisions are frequent and non-zero revisions are large. Second, asking professional forecasters directly reveals that reporting frictions rationalize lumpiness. Third, via a novel RCT, we obtain causal evidence in favor of strategic reporting behavior. Finally, we show how both lumpiness and strategic reporting interact.

Keywords: Surveys of professional forecasters, Strategic motives, Bayesian learning
JEL Codes: C53, C93, E32, E66

*We are thankful to Leland Farmer, Luca Gemmi, Alexandre Kohlhas, Manuel Menkhoff, and Oliver Pfäuti as well as numerous participants at various seminars and conferences, for insightful comments and suggestions. We are also grateful to the ZEW Financial Market Survey Team and, in particular, to Peter Buchmann for implementing the experiment. The randomized control trial was pre-registered at the AEA RCT Registry: `https://doi.org/10.1257/rct.17013-1.0`. We have obtained an ethics committee approval certificate No. ZEW-EC-2025-006. This work uses data from the ZEW Financial Market Survey. Any results, analyses, and observations published within this work do not necessarily correspond to the results or analyses of the data providers. Isaac Baley (Universitat Pompeu Fabra, CREI, Barcelona School of Economics, and CEPR): isaac.baley@upf.edu. Lukas Hack (University of Tübingen, ETH Zurich): lukas.hack@uni-tuebingen.de. Lora Pavlova (ZEW Mannheim): lora.pavlova@zew.de. Davud Rostam-Afschar (University of Mannheim, IZA, GLO): rostam-afschar@uni-mannheim.de. Javier Turen (Pontificia Universidad Católica de Chile): jturen@uc.cl.

1 Introduction

Professional forecasters play a central role in macroeconomics. Their expectations guide financial markets, inform policymakers, and are widely used to test models of expectation formation (e.g., Farmer, Nakamura, and Steinsson, 2024). In these contexts, the reported forecasts are typically interpreted as pure measures of beliefs: the forecast equals the conditional expectation of the target variable. We challenge this view and study reporting frictions, i.e., forces that drive a wedge between beliefs and reported forecasts.

This paper provides evidence in favor of such frictional forecasts. We study professional forecasts of German real GDP growth. Guided by a simple conceptual framework, we first focus on forecasters' decisions on whether or not to revise a forecast, the extensive margin, before focusing on the intensive margin. On the extensive margin, we document that forecasts are lumpy. Zero revisions are frequent, yet revisions are large when they occur. Asking professional forecasters directly reveals that reporting frictions rationalize lumpiness. On the intensive margin, we provide causal evidence suggesting strategic reporting behavior via a novel survey RCT. Lastly, we show how frictions operating on the intensive and extensive margin interact.

Our study is based on the *ZEW Financial Market Survey* (FMS), a panel of German professional forecasters administered by the ZEW Leibniz Center for European Economic Research in Mannheim. The FMS panel of forecasters was established in 1991 and is currently conducted as an online survey. The participants are forecasters from German firms, especially from the banking and insurance industry.[1] Moreover, the survey responses underlie the ZEW Indicator of Economic Sentiment, a monthly indicator whose release is known to move financial markets (Kerssenfischer and Schmeling, 2024). Since participants are aware of this, the survey constitutes a high-stakes environment.

The panel features 150 to 180 regular respondents, considerably more than in most other professional forecasting panels, providing us with a unique setup in which a survey RCT

[1] More details about the survey can be found in Brückbauer and Schröder (2023).

among professional forecasters is feasible.[2] We complement the regular survey with two special modules: a module aimed at understanding the lumpiness of forecasts, fielded in February 2026, and the RCT, fielded in November 2025.

We first turn to the extensive margin and ask how often forecasters revise their forecasts, and by how much? In the panel data, inaction is frequent: the share of exactly zero revisions for a fixed event forecast ranges from 17.4% to 25.7% across forecast horizons. At the same time, revisions are large conditional on being non-zero, with average absolute revisions of 0.40 to 0.65 percentage points, which is substantial. Thus, forecasts are lumpy. Similar patterns have been documented for Bloomberg and U.S. SPF forecasts (Baley and Turén, 2025a), suggesting that lumpiness is a general feature of professional forecasts rather than a peculiarity of our survey.

Lumpiness alone, however, does not establish a reporting friction. Forecasts could also be lumpy when forecasters perceive that new information arrives in a lumpy manner. We therefore ask the forecasters directly: our survey module elicits the motives behind forecast inaction as well as the minimum change in the economic outlook required for a forecaster to typically adjust her forecast. Almost all respondents, 91.9%, report the absence of new information as a reason for not revising. However, about three-quarters of respondents, 72.9%, additionally report at least one motive other than new information: adjustment costs, the desire to signal stability to build credibility, a lack of decision relevance, or an unchanged consensus forecast. All of these motives beyond information suggest the presence of reporting frictions. Linking these answers to the panel, we find that forecasters who report motives beyond new information tend to make small revisions less often. This confirms that these frictions affect actual reporting behavior in the panel.

On the intensive margin, we focus on one central question: Does the response to peers' forecasts reflect information or strategy? When a forecaster learns about her peers' forecasts,

[2]For example, the U.S. Survey of Professional Forecasters from the Philadelphia Federal Reserve Bank features between 30 and 40 participants. In an RCT with a treatment and control group, it is unlikely to have enough statistical power to detect potentially small treatment effects, especially when not all respondents participate in the RCT.

she may revise for two distinct reasons. First, peers' forecasts carry information about the target variable, so that learning about them is a reason to update one's belief about GDP growth. Second, a forecaster may revise for strategic reasons when she has a preference regarding how her forecast compares with peers. Both channels imply that reported forecasts respond to peer information, so that the response alone does not reveal its source.

Thus, we design and implement a novel RCT among professional forecasters that allows us to disentangle the two channels as follows. Before any information is provided, we elicit the relevant prior beliefs, i.e., forecasts, consensus beliefs, and the perceived rank in the forecast distribution. Then, we inform *all* participants about the lagged consensus, thereby fixing the information about the first moment of the forecast distribution in both groups. Only participants in the treatment group receive information about their rank. As a result, we can compute the rank perception gap as the distance between the prior rank belief and the truth. Our identifying assumption is first-moment sufficiency: conditional on the consensus, the remaining forecast distribution contains no additional fundamental information. Under this assumption, the treatment varies rank information while holding fundamental information fixed, so that differential updating in the treatment group isolates peer effects beyond the information content of peers' forecasts.

Our experiment was fielded in November 2025. In total, we have slightly more than 100 respondents who participated in our experiment, with treatment and control groups of similar size. Before analyzing the experimental data, we conducted several balancing tests to investigate whether the treatment and control groups are comparable. This is particularly important in our setting because the sample size is smaller than in other experiments (e.g., Haaland, Roth, and Wohlfart, 2023).[3] Across all variables and forecast horizons, we find that treatment and control groups are comparable in terms of the distributions of priors. This is confirmed by formal statistical tests, which fail to reject the null of identical distributions.

[3]Our sample is small due to the choice of our subject of interest, which are professional forecasters. Having an experiment with more than 100 participants is much more than what is feasible in other comparable panels of professional forecasters.

Given that we pass these balancing tests, we leverage our experimental variation and estimate Bayesian updating regressions as in Weber, Candia, Afrouzi, Ropele, Lluberas, Frache, Meyer, Kumar, Gorodnichenko, Georgarakos et al. (2025), augmented to account for heterogeneous treatment intensity. Our main results pertain to real GDP forecasts. We find that the control group has a weight on the prior below unity, with point estimates ranging from 0.75 to 1.02, suggesting some updating. However, often, we cannot reject the null of unity. A weight below unity suggests that the control group also updates its beliefs to some extent. In our active control group setting, this is to be expected. Moreover, updating of the control group is a common finding in the literature, where similar magnitudes are reported (e.g., Coibion and Gorodnichenko, 2026).

Yet, the treatment group puts considerably less weight on their priors. The change in the weight on the prior ranges from -0.20 to -0.65. On average across forecast horizons, we obtain a coefficient of -0.38, which is statistically significant at the 1% level. These effects are primarily driven by shorter forecast horizons, i.e., 2025Q4 (nowcast) and 2026Q1 (1-quarter ahead). The effects for horizons further out are insignificant, but the implied weight on the prior in the treatment group is remarkably stable across forecast horizons.

The discussed estimates correspond to the updating in the treatment group at the average perception gap. We also account for heterogeneous treatment intensities, allowing the weight on the prior to also depend on the (absolute) perception gap, which is the distance between the prior beliefs about the cumulative distribution function and the truth. Consistent with the average effects discussed above, a perception gap that is one standard deviation above the average implies a change in the coefficient on the prior between -0.11 and -0.33. As before, these estimates are highly statistically significant for short forecast horizons but not for the longer ones. Importantly, this change in the weight on the prior comes on top of the average effect.

Lastly, we reconcile updating with the RCT with the presence of fixed adjustment costs: in the experiment, treated forecasters revise their reported forecasts within a few minutes,

whereas the panel data show inaction between two distinct survey waves. We find that both frictions interact. Survey participants who report a higher inaction threshold respond less to the treatment.

Related literature. Our primary contribution is to the literature that studies professional forecasts. This influential literature has a long tradition focusing on anomalies in professional forecasts, starting with Ehrbeck and Waldmann (1996), who find evidence for behavioral biases. Professional forecasts are also used as a benchmark to test theories of expectation formation beyond full-information rational expectations. Using the U.S. Survey of Professional Forecasters, Coibion and Gorodnichenko (2015) document underreaction to news at the consensus level, suggesting information frictions, whereas Bordalo, Gennaioli, Ma, and Shleifer (2020) find overreaction at the forecaster level, in line with behavioral frictions. Several alternative explanations for these seminal findings have been proposed. Gemmi and Valchev (2025) argue that both facts can be rationalized with strategic motives as originally analyzed in Ottaviani and Sørensen (2006). In contrast, Broer and Kohlhas (2024) and Adam, Kuang, and Xie (2025) argue that overconfidence in private information plays a key role in explaining the above-mentioned anomalies. Alternatively, Juodis and Kučinskas (2023) find that measurement error is important, and Farmer et al. (2024) argue that some anomalies may be explained because underlying processes are "hard to learn". Finally, Baley and Turén (2025a,b) argue that fixed forecast adjustment costs and strategic considerations are prevalent among professional forecasters, the former being consistent with earlier evidence from Andrade and Le Bihan (2013).

While all of these papers significantly advance our understanding of expectation formation among professional forecasters, none of them elicits forecasters' motives directly or presents credible causal evidence. We go beyond this work in two ways. First, our survey module provides direct evidence on the motives behind lumpy forecasts, complementing the observational evidence on lumpiness in Baley and Turén (2025a) and Andrade and Le Bihan

(2013). Second, and to the best of our knowledge, we present the first causal evidence on peer effects among professional forecasters. This causal evidence suggests that the entire distribution of forecasts matters for each individual forecaster. Strategic motives or intrinsic preference over the relative position in the forecasting distribution are promising candidates to explain this finding.

Beyond the literature on professional forecasters, we also relate to the experimental literature on forecasting. Assenza, Bao, Hommes, and Massaro (2014) provide an overview. The closest paper from this literature is Afrouzi, Kwon, Landier, Ma, and Thesmar (2023), which focuses on the result of overreaction with an experiment using MTurk participants. Instead, we focus directly on professional forecasters and potential distributional considerations.

Lastly, we relate to the experimental literature on Bayesian updating (e.g., Coibion, Gorodnichenko, and Kumar, 2018; Coibion, Gorodnichenko, and Ropele, 2020). Many of these works are synthesized in Weber et al. (2025) to investigate whether the existing experimental evidence is in line with rational expectations. We contribute to this literature by bringing the RCT approach to professional forecasting.

2 Conceptual framework

The premise of this paper is that reported forecasts may also reflect frictions in reporting. To fix ideas, we begin with a simple framework of forecast reporting. Its purpose is to organize the empirical analysis in the remainder of the paper.

Notation. Consider forecasters indexed by $i = 1, \ldots, N$ and periods $t = 1, \ldots, T$. In each period, forecaster i reports a survey forecast f_t^i. The forecast target variable, real GDP growth in our empirical application, is denoted by y.[4] The forecaster's true belief is $m_t^i = \mathbb{E}_t^i[y]$, where $\mathbb{E}_t^i[\cdot]$ denotes expectations, conditional on her information set at time

[4]To economize on notation, we omit the forecast horizon so that the following setup can be understood to apply to any generic forecast horizon.

t.[5] Our key distinction is between the forecaster's true belief, m_t^i, and the reported survey forecast, f_t^i.

The reporting problem. We consider two distinct reasons (frictions) why $m_t^i \neq f_t^i$: a positioning motive and a fixed adjustment cost. Given these potential frictions, a stylized forecast reporting problem is given by

$$\min_{f_t^i} \quad \underbrace{\left(f_t^i - \mathbb{E}_t^i\, y\right)^2}_{\text{accuracy motive}} + \underbrace{\rho\, \mathcal{S}(f_t^i, \mathbb{E}_t^i\, R_t)}_{\text{positioning motive}} + \underbrace{\kappa\, \mathbf{1}\{f_t^i \neq f_{t-1}^i\}}_{\text{fixed adjustment cost}}, \tag{2.1}$$

where the weight on the accuracy motive is normalized to unity. The function $\mathcal{S}(f_t^i, \mathbb{E}_t^i\, R_t)$ captures the positioning motive with respect to the expected forecast distribution, summarized by its cross-sectional CDF R_t, and ρ is the relative weight on this motive. Lastly, the fixed cost κ accrues whenever the reported forecast deviates from the inherited report f_{t-1}^i.

Interpretation. The reporting problem nests the standard interpretation of survey forecasts. If $\rho = \kappa = 0$, then forecasters care only about accuracy and report their beliefs, i.e., $f_t^i = m_t^i$. If instead $\rho \neq 0$ or $\kappa \neq 0$, then reporting frictions drive a wedge between beliefs and reported forecasts. The positioning motive captures any preference over the position of the own forecast relative to the distribution of all forecasts. It gives rise to frictional *action*: conditional on adjusting, forecasters trade off accuracy against how the own forecast compares with those of their peers, e.g., due to strategic considerations (Gemmi and Valchev, 2025). The fixed adjustment cost can be interpreted, e.g., as the cost of rewriting the narrative that accompanies a revised forecast, the cost of recomputing forecasts, or a desire to signal stability (Baley and Turén, 2025a). It gives rise to frictional *inaction*: small changes in beliefs or positioning of the own forecast do not justify incurring κ, so that reported forecasts become lumpy.

In the remainder of the paper, Section 3 studies frictional inaction, Section 4 turns to

[5] We make no assumption on whether beliefs are formed rationally or subjectively.

frictional action to isolate pure positioning motives, and Section 5 focuses on the interaction of both.

3 The extensive margin: understanding lumpiness

In this section, we introduce the professional forecasts from Germany that we study and provide additional institutional context. We document that lumpiness is a pervasive feature of the data. Asking professional forecasters directly, we confirm that lumpiness reflects reporting frictions.

3.1 Institutional setting and data

The German economy. We fielded two special survey modules: the RCT in November 2025 (see Section 4) and a module on the lumpiness of forecasts in February 2026 (introduced below). Thus, to provide additional context, we briefly recapitulate the economic situation of the German economy around and before the modules were fielded.

The German economy has been stagnating for several years. Following two recession years in 2023 and 2024, annual real GDP growth turned positive again in 2025 with a growth rate of 0.2%. Quarterly data show that real GDP in the third quarter of 2025 remained unchanged from the previous quarter after seasonal and calendar adjustments, with equipment investment developing positively, while exports declined compared with the previous quarter. At the time our survey module was fielded in November 2025, the most recent official data therefore characterized the German economy as broadly stagnant[6], with no sustained expansion visible in high-frequency indicators. A small increase in output was expected for the final quarter of the year, primarily related to consumption and public expenditure, but the magnitude of the projected growth was limited. Overall, the macroeconomic environment during the survey period was characterized by weak real GDP growth.

[6]See, e.g., press release No. 388 of 30 October 2025 of the Federal Statistical Office.

The survey. The ZEW Financial Market Survey (FMS) is a panel among professional forecasters, mostly financial market experts, in Germany. The FMS targets professionals employed in financial institutions or financial divisions of non-financial firms (Brückbauer and Schröder, 2023) and has 150 to 180 regular respondents. It was introduced in December 1991 and surveys all forecasters in the panel monthly. Moreover, the ZEW publishes a monthly indicator based on the survey responses, the ZEW Indicator of Economic Sentiment, which is known to move financial markets (Kerssenfischer and Schmeling, 2024). This is also known to survey participants, making the survey a high-stakes environment that may support the overall quality of the survey. The core module of the survey focuses predominantly on qualitative assessments and expectations about the overall economic situation, the inflation rate, short- and long-term interest rates, stock market indices, and exchange rates for Germany, the U.S., China, and the Euro Area.

Once per quarter, usually in the first month of each quarter, the professional forecasters provide point forecasts and probabilistic predictions of the Euro Area inflation rate for the current year and for one and two years ahead. Similarly, in the second month of each quarter, participants are asked for their yearly and quarterly real GDP point forecasts for Germany. These questions were added to the survey in 2014. The annual growth rates are elicited for the current and the two subsequent years. The quarterly growth rates are elicited for the current quarter (a nowcast) and the three subsequent quarters. The three-quarter-ahead forecast has been added only in 2024 and replaced a so-called backcast, which was often included before. A backcast is a forecast for the previous quarter elicited before the first official estimate (the so-called flash release).

For our analysis, we use the resulting panel of real GDP growth forecasts over the maximum available sample, which covers 2014Q4 to 2026Q1.

Survey module on lumpiness. To measure the motives behind forecast inaction directly, we added a survey module to the February 2026 wave, placed after the regular quarterly real

GDP growth questions. We use two questions from this module. The first question asks for the reasons why a forecaster typically does not revise:

> *When you do not revise your economic growth forecast for a given quarter: Which of the following reasons for your decision not to make a revision* typically *apply?*
>
> - *I change my forecast only if it is decision-relevant*
> - *There is no change in the average forecast among all participants (the consensus forecast)*
> - *My forecast is more credible if it is adjusted less frequently*
> - *The workload or the costs of the forecast adjustment are too high if there are only minor changes in the economic outlook*
> - *There is no new information that suggests a change in the economic outlook*
> - *Other:* ☐

Respondents rate each reason on a scale with the options *applies*, *tends to apply*, *undecided*, *tends not to apply*, and *does not apply*; they may also choose *no answer*. The second question elicits a minimum revision size to actually revise a forecast:

> *How large must the change in the economic outlook be for you to typically adjust your forecast in the survey?*
>
> *The change in the economic outlook should be so large that my forecast changes at least by:*
>
> ±*0.10 percentage points*
>
> ±*0.20 percentage points*
>
> ±*0.30 percentage points*
>
> ±*0.40 percentage points*
>
> ±*0.50 percentage points*
>
> ± ☐ *#.##* ☐ *percentage points*
>
> *I always adjust my forecast when the economic outlook changes*

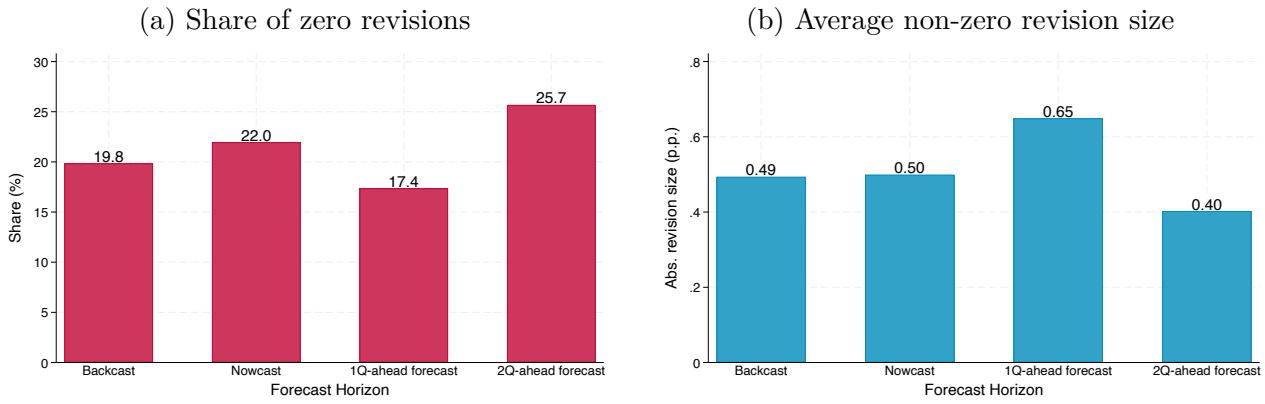
Respondents select exactly one option, where the second-to-last option allows them to enter a number of their choice. The full questionnaire of the special module is shown in Appendix C.

3.2 Main results

Lumpiness in the panel. We begin by documenting revision behavior in the panel to answer the following question: How often do forecasters leave their forecasts unchanged, and how large are revisions when they occur? Figure 1 reports these statistics separately by forecast horizon. Panel (a) shows the share of zero revisions between two consecutive survey waves for a fixed event, e.g., GDP growth in 2026Q1. Panel (b) shows the average absolute revision size conditional on a non-zero revision.[7] We find that inaction is pervasive. The share of zero revisions ranges from 17.4% for the 1-quarter ahead forecast to 25.7% for the 2-quarter ahead forecast, with 19.8% for the backcast and 22.0% for the nowcast. At the same time, revisions are large conditional on being non-zero: the average absolute revision size ranges from 0.40 to 0.65 percentage points across horizons, which is substantial relative to typical quarterly growth rates of the German economy.

As a result, we conclude that forecasts are lumpy. Similar patterns have been documented for Bloomberg and U.S. SPF forecasts (Baley and Turén, 2025a), suggesting that lumpiness is a general feature of professional forecasts rather than a peculiarity of our survey.

Figure 1: Lumpiness in the panel



Notes: This figure shows statistics computed from the panel data by forecast horizon, excluding 2020. Backcast refers to the forecast for the previous quarter (before the official estimate is released). Nowcast refers to the prediction for the current quarter (in which the survey is conducted. Panel (a) shows the share of revisions between two consecutive survey waves for a fixed event. Panel (b) shows the average absolute revision size, conditional on a non-zero revision.

[7] We exclude the year 2020 from these statistics because of the extraordinarily large revisions during the COVID-19 pandemic. The results including 2020 are shown in Figure A.1.

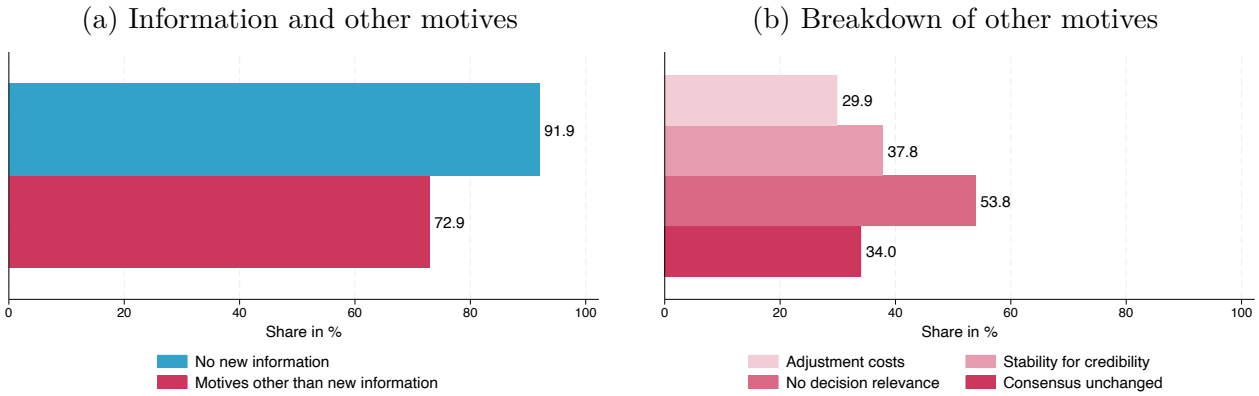
Motives for lumpiness. Lumpiness could reflect reporting frictions. However, forecasts could also be lumpy when forecasters perceive that new information arrives in a lumpy manner. Since these two explanations are observationally equivalent in the panel data, we ask the forecasters directly, using the survey module introduced in Section 3.1. The conventional, frictionless view makes a sharp prediction for these answers: the only reason for not revising is that no new information has arrived since the last survey wave, so respondents should report the absence of new information as the sole motive and nothing else.

Figure 2 shows the percentage share of forecasters who report that a motive applies as a reason to not revise a forecast.[8] Panel (a) displays the share of respondents who report the absence of new information and the share who report at least one other motive. Panel (b) offers a breakdown into each of the other motives. Almost all respondents, 91.9%, indeed report the absence of new information as a reason for not revising. However, about three quarters of respondents, 72.9%, additionally report at least one motive other than new information. Among the other motives, 29.9% report adjustment costs, which may capture the burden of writing the narratives, reports, and explanations that accompany a revised forecast. 37.8% report stability for credibility, i.e., they avoid revisions in order not to appear erratic. More than half of the respondents, 53.8%, report a lack of decision relevance, meaning that their employer does not require an updated forecast when it is not needed for internal purposes. Lastly, 34.0% report an unchanged consensus. This last motive can reflect the consensus as an information source or a positioning motive relative to the consensus, a distinction that we revisit in Section 4.

The bottom line contrasts with the frictionless view: the forecasters themselves reveal that reporting frictions matter for their inaction.

[8] We aggregate the five-point Likert scale into a binary variable as follows: We compute the shares of respondents who report that a motive *applies* or *tends to apply* relative to all respondents that take a stance, i.e., report *applies*, *tends to applies*, *tends to not apply*, or *does not apply*. The disaggregated responses are shown in Figure A.2.

Figure 2: Motives for lumpiness

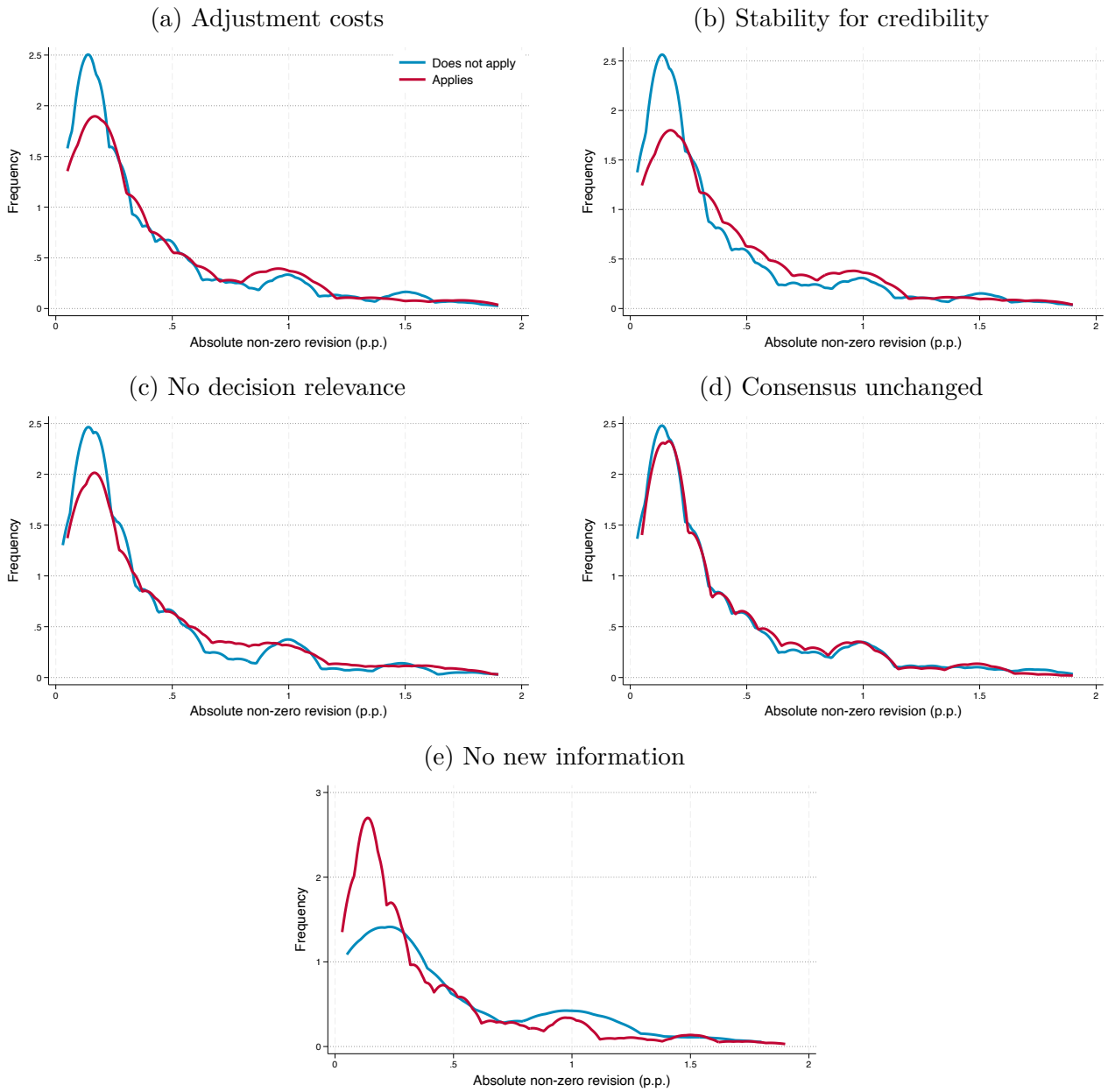


Notes: This figure shows reported motives for not revising a forecast. We display the shares of respondents who report that a motive *applies* or *tends to apply* relative to all respondents that take a stance, i.e., report *applies*, *tends to applies*, *tends to not apply*, or *does not apply*, for details, see Appendix C. Panel (a) shows the shares of forecasters that report *no new information* and at least one *other motive* as a justification for not revising a forecast, respectively. Panel (b) shows the breakdown into the underlying other motives.

Motives and lumpiness in the panel. A natural question is whether the answers to our special module questions are consistent with actual revision behavior in the panel. If the reported motives capture genuine frictions, forecasters who report them should make small revisions less often. Thus, Figure 3 shows kernel density estimates of absolute non-zero revisions, splitting the sample into forecasters who report that a given motive applies and those who report that it does not. For the motives other than new information in panels (a) to (d), forecasters who report that the motive applies indeed tend to make small revisions less often, i.e., the density of their revisions has less mass at small revision sizes. For the absence of new information in panel (e), the pattern reverses: forecasters who report this motive make small revisions more often, although this distinction is driven by a small number of forecasters who report that the absence of new information is not a reason for not revising.

In summary, the reported motives line up with revision behavior in the panel in a way that is consistent with reporting frictions, whereas a lumpy arrival of new information alone is unlikely to explain the lumpiness of forecasts.

Figure 3: Motives for lumpiness



Notes: This figure shows density estimates of absolute non-zero forecast revisions, based on the Epanechnikov kernel. Each panel splits the sample into forecasters who report that a motive *applies* or *tends to apply*, relative to all respondents that take a stance, i.e., report *applies*, *tends to applies*, *tends to not apply*, or *does not apply*, and vice versa for "does not apply". For details, see Appendix C.

3.3 Additional results

We briefly discuss several additional results with the corresponding results shown in appendices A and B.

Inaction thresholds. Our survey module also elicits the minimum change in the economic outlook required for a forecaster to typically adjust her forecast. We interpret this minimum revision size as a self-reported inaction threshold. If fixed adjustment costs are at work, such thresholds should be positive for most forecasters and correlate with the motives discussed before. The inaction thresholds should also predict actual revision sizes in the panel. We examine each implication in turn. First, Figure A.3 shows the distribution of the thresholds: 100 out of 136 respondents report a strictly positive threshold, and the average threshold is 0.16 percentage points across all respondents and 0.22 percentage points when excluding zeros. Second, Figure A.4 relates the thresholds to the reported motives. Forecasters who state that at least one motive other than new information applies report larger thresholds, with the difference being most pronounced for the lack of decision relevance. Lastly, regressing absolute non-zero revisions in the panel on the self-reported threshold delivers an average coefficient of 0.42, which is statistically significant at the 5% level, see Table B.1 and Table B.2.

Time and forecaster fixed effects One may wonder whether certain forecaster types or certain episodes drive lumpiness. To assess this, we regress the revision indicator on forecaster and time fixed effects and examine their explanatory power, see Table B.5 and Table B.6. Forecaster fixed effects deliver an R^2 of 0.18 on average across forecast horizons, and additionally including time fixed effects raises the R^2 only to 0.21. Thus, fixed effects explain a comparably small fraction of the variation in inaction, suggesting that lumpiness is a broad-based phenomenon.

Short vs. long horizons. Lastly, we examine whether forecasters revise their forecasts jointly across horizons. Typical macroeconomic shocks are persistent, so news that induces a revision of near-term forecasts should usually also trigger revisions at larger forecast horizons. Yet, we often observe that nowcasts and one-quarter-ahead forecasts are adjusted while two- and three-quarter-ahead forecasts are not; see Table B.3. The estimated "pass-through"

from short-horizon to long-horizon revisions is significantly below unity. This imperfect pass-through points to some inaction over and above a lumpy arrival of new information.

4 The intensive margin: information vs. strategy

This section studies the intensive margin of forecast revisions. We focus on disentangling whether forecasters are responsive to peer forecasts because they carry information about fundamentals as opposed to pure strategic considerations about one's relative position in the distribution of peer forecasts. Our RCT is tailored to speak to this question, and the results suggest the presence of a pure positioning motive.

4.1 Identification challenge

We focus on one central question: Does the response to peers' forecasts reflect information or strategy? When a forecaster learns about her peers' forecasts, she may revise for two distinct reasons. First, peers' forecasts carry information about the target variable, so that learning about them is a reason to update one's belief about GDP growth. Second, a forecaster may revise because of the positioning motive from our conceptual framework in Section 2: if she cares about her rank in the forecast distribution, news about this rank triggers a revision even when her belief about GDP growth is unchanged. Both channels imply that reported forecasts respond to peer information, so that the response alone does not reveal its source. Previous work provides no direct evidence on this question and rather assumes that deviations from the consensus can be used as a proxy to study strategic behavior (e.g., Gemmi and Valchev, 2025; Baley and Turén, 2025a,b).

To disentangle both motives, we require measures of the relevant beliefs and exogenous variation in these beliefs. Our RCT provides us with both. We can measure the relevant beliefs and generate exogenous variation in them through randomized information provision.

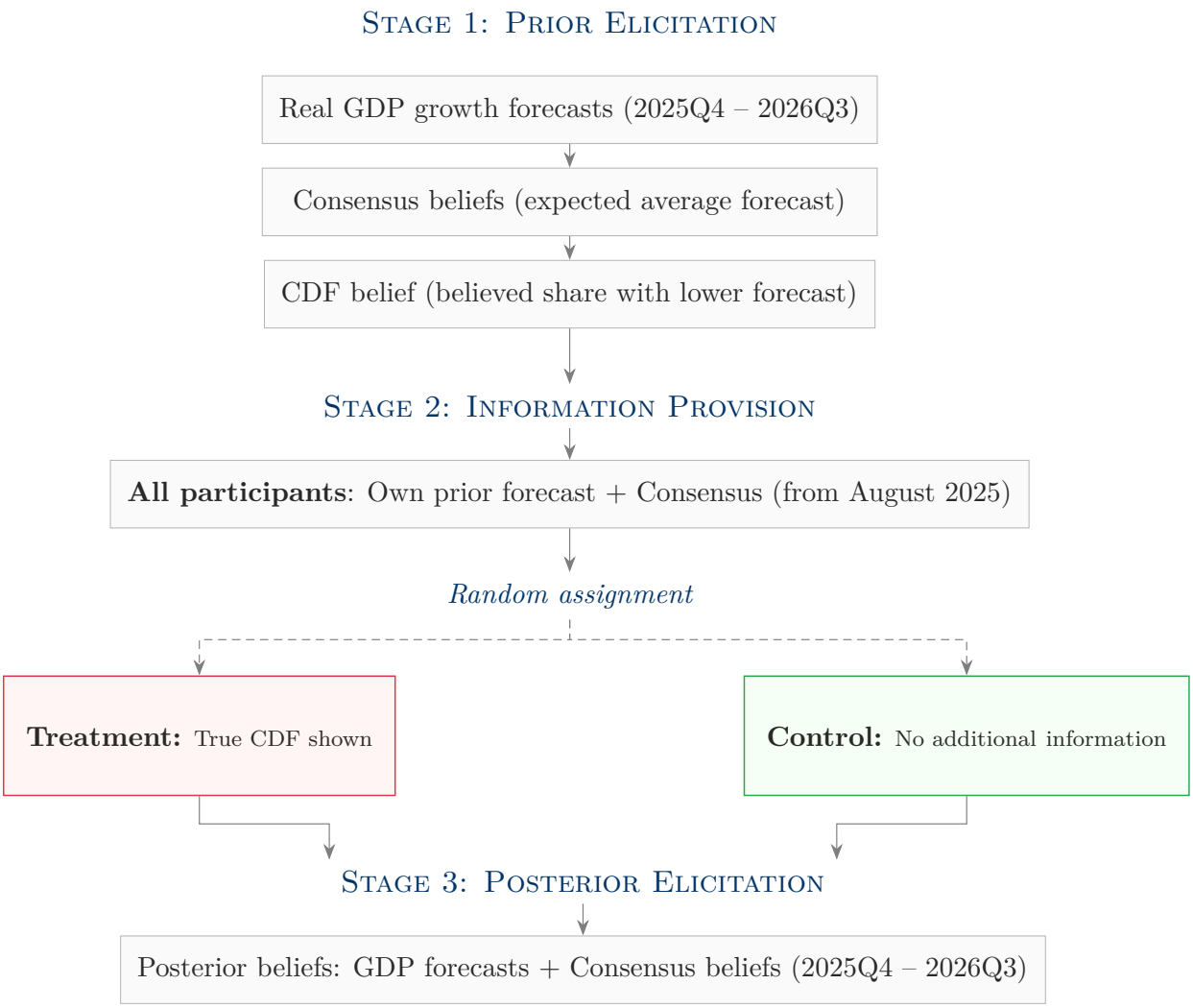
4.2 Experimental design

Our experimental design addresses the identification challenge as follows. Before any information is provided, we elicit the relevant prior beliefs, i.e., forecasts, consensus beliefs, and the perceived rank in the forecast distribution. Then, we inform *all* participants about the lagged consensus, thereby fixing the information about the first moment of the forecast distribution in both groups. Only participants in the treatment group additionally receive information about their rank. Our identifying assumption is first-moment sufficiency: conditional on the consensus, the remaining forecast distribution contains no additional fundamental information. Under this assumption, the treatment varies rank information while holding fundamental information fixed, so that differential updating in the treatment group isolates peer effects beyond the information content of peers' forecasts.

Design components. The survey experiment includes a within-subject and a between-subject component. The within-subject component is implemented by asking about priors and posteriors before and after the information provision experiment, respectively. The between-subject component is implemented by providing some information only to a subset of participants, i.e., the treatment group. Participants are randomly assigned to the treatment and control groups with equal probability. Figure 4 provides an overview of our experimental design, which we explain below in more detail.

Timing and participants. The experiment was included in the November 2025 survey wave, fielded between November 3 and November 10. To implement our information treatment, we used information about the 2025Q4 forecasts provided in previous waves. These forecasts were elicited in February, May, and August 2025. For each participant, we used the most recently provided forecast for 2025Q4. For most participants, we could use the August forecasts, but a few forecasters did not participate in the August wave. In such cases, we used the most recent forecast, either from May or February. Additionally, we must restrict

Figure 4: Overview of the experimental Design



the set of participants to those who have provided a 2025Q4 forecast in at least one of these three prior survey waves.

Prior forecasts. We elicit prior beliefs for two variables before any additional information is shown to the participants. These variables are quarterly real GDP growth forecasts and the belief over the average (consensus) real GDP growth forecast among all participants in the *current* wave, i.e., in the November 2025 wave. GDP growth forecasts are already elicited in the regular survey as follows:

Point forecast of the growth rate of the German GDP

For the quarterly values, please indicate non-annualized quarterly real & seasonally adjusted GDP growth.

Q4 2025 [#.##] *pct* *Q1 2026* [#.##] *pct* *Q2 2026* [#.##] *pct* *Q3 2026* [#.##] *pct*

As indicated, participants may provide their forecasts at basis point precision.[9] The full questionnaire, as implemented in the survey, is shown in Appendix C. Subsequently, we elicit consensus beliefs using a question that we have added to the survey. It asks for the average economic growth forecast among all respondents in the current survey and is formulated analogously to the growth forecasts, i.e., for the same forecast horizons and at the same precision; see Appendix C for the exact wording. To the best of our knowledge, we are the first to elicit these consensus beliefs, as opposed to the realized consensus. We view this as an additional contribution of our paper since most papers proxy the consensus belief only via its realization (e.g., Gemmi and Valchev, 2025; Baley and Turén, 2025a,b). Based on our question, one can assess the extent to which realizations and actual beliefs align.

Prior distributional beliefs. To implement our information provision experiment, we ask one additional prior question to measure treatment intensity. Referring to the last survey wave in which a respondent provided a real GDP growth forecast for 2025Q4, the question asks for the beliefs about the cumulative distribution function evaluated at one's own forecast. In other words, the question asks for the share of more pessimistic forecasters than oneself. Below, we show this question for participants who provided a forecast for 2025Q4 in the August wave:[10]

Previously, in August 2025, you provided a forecast for quarterly growth for Q4 2025.

What do you think about the forecasts of other participants in that previous survey?

The share of all respondents who, in August 2025, stated a lower growth rate for Q4 2025 than

[9] Additionally, the survey asks about annual real GDP forecasts and potential reasons for forecast revisions. However, we do not use these additional questions in our experiment and analysis.

[10] Recall that for each forecaster, we use the most recent wave in which she provided a real GDP growth forecast for 2025Q4. This is the August 2025 wave for most participants.

you, amounted to...

Participants could enter their beliefs by filling in the text box or by moving the slider. We introduced the latter to reduce survey fatigue and enhance participation because participants have filled several text boxes at this point in the survey. Based on this question, we can compute participants' perception gap, i.e., the true value minus their prior beliefs. Participants with a positive perception gap have underestimated the share of forecasters who are more pessimistic than themselves, i.e., they have underestimated the cumulative distribution function of all forecasts evaluated at their own forecast.

Treatment group. In our experiment, we provide certain information to both treatment and control groups, as we explain in the subsequent paragraph. In addition, the perception gap is the information that we provide only to the treatment group in our experiment. Specifically, we inform participants in the treatment group about the actual share of more pessimistic forecasters, their prior belief (which they have entered before), and the implied gap between the two, i.e., the perception gap. This information is provided graphically so as not to be overlooked in the survey; see Figure C.4 in Appendix C.

Control group. Two pieces of information are provided not only to the treatment group but also to the control group. This indicates that we have an "active" control group design, in which the control group may also update its beliefs in response to the provided information. First, both groups are reminded of their 2025Q4 forecasts from the previous wave, i.e., the same wave used to compute the perception gap. We do so to account for possible memory imperfections and because this information is instrumental in interpreting the perception gap. Second, we inform all participants about the corresponding 2025Q4 consensus (average) forecast from the same wave. We present the information in exactly this order. In the treatment

group, the treatment is shown after these two statistics, see Figure C.4 in Appendix C.[11]

Posterior questions. After presenting the experimental information to the participants, we elicited posterior beliefs. We asked again for their real GDP growth forecasts and corresponding consensus beliefs over the same forecast horizons. We formulated the question identically as for the priors, but motivated the participants by stating: *Please think again about Germany's economic growth prospects and provide your assessment*, see Figure C.5 in Appendix C. Importantly, we do not refer back to the provided information to not trigger experimenter demand effects, which are a common concern when designing RCTs (e.g., Haaland et al., 2023). Our approach contrasts with several papers that instead elicit posteriors with a distributional question that implies a point forecast (e.g., Weber et al., 2025). We adopted our approach to mitigate participant fatigue: Implementing distributional questions for all eight items, encompassing two variables across four forecast horizons, would have significantly increased the cognitive burden, potentially compromising survey completion rates.[12] An additional benefit of our approach may be that this introduces less measurement error compared to a distributional question. Both of these aspects are critical to our empirical strategy. Given that professional forecasting panels are inherently characterized by a small number of participants, maximizing the response rate and limiting measurement error is essential to maintain sufficient statistical power.

4.3 Balancing tests

Participation. The baseline questionnaire of the November wave was completed by 186 professional forecasters, of whom 118 participated in our survey module. This implies a participation rate of 64%. Only 17 did not complete our full survey module, leaving us with 101 professional forecasters who participated in the full experiment. This group is divided

[11] We deliberately did not randomize the order of appearances because this order appears to be natural.

[12] For example, if we had asked for four probabilities for fixed support points, participants would have been required to input 32 numbers as opposed to 8 in our implementation.

into 49 and 52 participants in the treatment and control groups, respectively.

Descriptive comparison. Random assignment to treatment and control groups should ensure that prior beliefs are similarly distributed in both groups. However, since this is not guaranteed in small samples, it is particularly important to verify that treatment and control groups are truly comparable with respect to their prior beliefs.

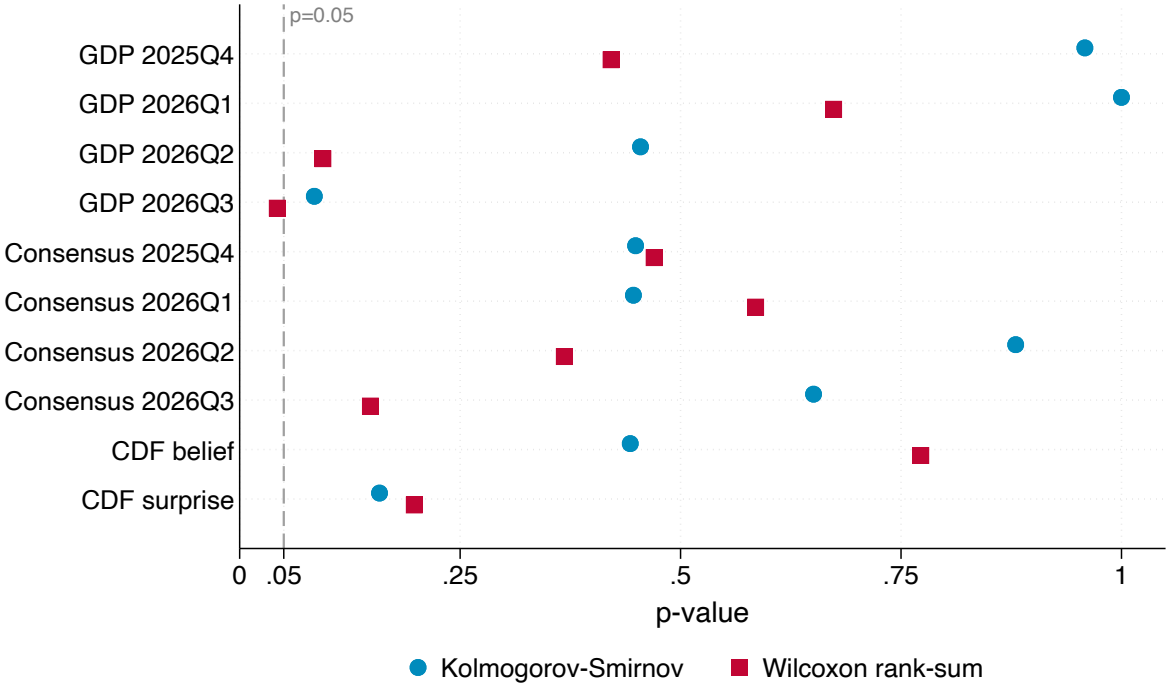
In Appendix A, we compare the prior distributions of both groups via histograms and kernel density estimates. Figure A.5 and Figure A.6 show the prior forecasts and consensus beliefs for each forecast horizon, ranging from 2025Q4 (h=0 quarters ahead) to 2026Q3 (h=3 quarters ahead). These figures reveal a substantial dispersion in forecasts and consensus beliefs ranging from -0.5 to 1.0 percentage points. Yet, for both variables and all four forecast horizons, the distributions are reasonably close to each other, suggesting that treatment and control groups are broadly comparable.

Additionally, Figure A.7 shows the corresponding estimates for the prior belief about their rank, i.e., the forecast CDF evaluated at one's own forecast. Additionally, we show the rank perception gaps, i.e., $|\Delta CDF_i| = |CDF_i - \mathbb{E}_i[CDF_i]|$, where CDF_i denotes the (true) cumulative distribution function for 2025Q4 from previous waves, evaluated at the forecast of participant i.

Formal balancing tests. Lastly, to formally evaluate whether the treatment and control groups are drawn from the same underlying distribution, we perform two non-parametric tests for each prior variable: the Kolmogorov-Smirnov and the Wilcoxon rank-sum test. Figure 5 summarizes the p-values of all 20 tests of the null that both samples are drawn from the same distribution. We can reject the null hypothesis only once at the 5% level. This is the case for the Wilcoxon test for the growth forecast three-quarters ahead. While it may be coincidental, one rejection out of 20 tests is exactly the Type I error rate that one

may expect under the null being true when testing at the 5% level.[13] Thus, we conclude that there is no systematic evidence that the treatment and control groups are not comparable.

Figure 5: Balancing tests



Notes: This figure compares the prior distributions across treatment and control groups. We show the p-values for the Kolmogorov-Smirnov and the Wilcoxon rank-sum test for the null hypothesis that both treatment and control groups are drawn from the same underlying distribution for the corresponding forecast horizon.

[13] Residual concerns may arise since the support occasionally differs between treatment and control groups. While this is not unusual in finite samples, our main results are insensitive to restricting our estimation sample so that the support is similar across both groups.

4.4 Econometric framework

Regression model. To formally test whether peer forecasts matter for our survey participants, we estimate the following regression

$$\begin{aligned}\text{Posterior}_i^h = \beta_0^h + \beta_1^h \,\text{Treat}_i + \beta_2^h \,\text{Prior}_i^h + \beta_3^h \,\text{Prior}_i^h \times \text{Treat}_i \\ + \gamma_1^h \,|\Delta CDF_i| + \gamma_2^h \,|\Delta CDF_i| \times \text{Treat}_i + \gamma_3^h \,|\Delta CDF_i| \times \text{Prior}_i^h \times \text{Treat}_i + \nu_i^h,\end{aligned} \tag{4.1}$$

where Prior_i^h and Posterior_i^h refer to the real GDP growth forecast of participant i, elicited before and after the information provision experiment, respectively. Superscript h indexes the forecast horizon. Variable $\text{Treat}_i \in \{0, 1\}$ is a treatment indicator only activated when i belongs to the treatment group. $|\Delta CDF_i| \geq 0$ is the distance between the prior CDF belief and the truth, that is, the rank perception gap. This measures treatment intensity, which we standardize to ease interpretation. Lastly, ν_i^h is an error term. We estimate (4.1) using OLS and compute standard errors robust to heteroskedasticity.

Interpretation. Note that the first row of (4.1) corresponds to a standard Bayesian updating specification commonly used in the literature (e.g., Weber et al., 2025). Coefficient β_2^h captures the weight put on the prior belief in the control group, and β_3^h captures the differential weight on the prior in the treatment group.[14] The second row of (4.1) augments the regression specification to account for heterogeneous treatment intensities, depending on the distance between the prior CDF belief and the truth, i.e., $|\Delta CDF_i|$. We control for this variable in levels and additionally interact it with the treatment indicator and the treatment indicator times the prior.[15] Thus, γ_3^h measures the differential weight on the prior in the treatment group depending on $|\Delta CDF_i|$. As a result, the total differential weight on the

[14] If there was no information provided to the control group, we should expect $\beta_0^h = 0$ and $\beta_2^h = 1$, i.e., posteriors and priors coincide in the control group. However, in our active control group design, this is not necessarily the case, as we discuss below.

[15] Note that we do not include $|\Delta CDF_i| \times \text{Prior}_i^h$ as an additional regressor because the control group is not informed about $|\Delta CDF_i|$.

prior of participant i is $\beta_3^h + \gamma_3^h |\Delta CDF_i|$.

Model averaging. Throughout, we report estimation results for each of the four forecast horizons. In addition, we also present average estimates across all forecast horizons. When averaging across forecast horizons, we report $\bar{\beta}_j = (\sum_{h=0}^{3} \beta_j^h)/4$ and $\bar{\gamma}_j = (\sum_{h=0}^{3} \gamma_j^h)/4$ for all coefficients j. To obtain valid standard errors, we estimate seemingly unrelated regressions via a stacking approach and compute standard errors robust to heteroskedasticity and clustered at the forecaster level.

4.5 Main results

Our experiment is designed to answer one question: Do treated forecasters, who learn about their rank in the forecast distribution, update differently than the control group? We present the corresponding results for real GDP growth forecasts in Table 1 with standard errors in parentheses. Columns (1) and (2) show the average estimates across forecast horizons as baseline, and columns (3) to (10) show the corresponding estimates for each forecast horizon, which reveal the heterogeneity behind the averages.

Control group. We first focus on the control group, which provides the benchmark against which the treatment effects are measured. A coefficient of unity on the prior (and a zero intercept) may be expected when the control group receives no information: In this case, there is no updating and full weight on the prior belief, which, therefore, coincides with the posterior. Instead, in our active control setup, some updating away from the prior can be expected. Starting with the average estimates, we obtain a weight on the prior of 0.853 across forecast horizons, together with a small intercept. Importantly, the updating in our control group is not stronger than what is typically found in other survey experiments, even without an active control group. For example, Coibion and Gorodnichenko (2026) report a weight of 0.851; see their Figure 2. Turning to the individual forecast horizons, the weight on the prior ranges between 0.75 and 1.02, which is significantly different from unity for half

of the forecast horizons, see the middle panel of Table 1. The intercept is relatively small throughout, ranging from 0.02 to 0.11, but the estimates are significantly different from zero in most specifications.

Average treatment effects. If treated participants place less weight on their priors than the control group, this indicates peer effects beyond the informational content of peers' forecasts, given the assumption of first-moment sufficiency. Indeed, we find that participants who received additional information about the forecast CDF place significantly less weight on the prior than the control group. On average across forecast horizons, we find a reduction in the weight on the prior of -0.38, which is statistically significant at the 1% level, see our augmented specification in column (2) of Table 1. The simple updating specification from Weber et al. (2025) in column (1) delivers a coefficient of -0.34, which is statistically significant at the 5% level. In the middle panel, we can also see that the weight on prior ranges between 0.47 and 0.52, which is given by $\beta_2^h + \beta_3^h$. The weight on the prior is significantly different from unity at the 1% level in both specifications. Overall, we find strong and highly statistically significant evidence that professional forecasters respond significantly to information about their peers' forecasts.

Heterogeneity across forecast horizons. The average effects could mask heterogeneity across forecast horizons. In particular, our information treatment is based on the nowcast, which may suggest that shorter forecast horizons respond more strongly. To investigate this, we next discuss how the differential weight on the prior depends on the forecast horizon, focusing on our augmented specification in columns (4), (6), (8), and (10).[16] For the nowcast (2025Q4) in column (4), the effect size is -0.31 and statistically significant at the 5% level. For the 1-quarter ahead forecast in column (6), we obtain a larger coefficient of -0.62 which

[16]We suppress the discussion of the simpler specification in the remaining columns since the estimates are fairly similar.

Table 1: Bayesian updating regressions

	Avg. estimate		2025Q4		2026Q1		2026Q2		2026Q3	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
β_0^h: Constant	0.054***	0.053***	0.038**	0.038**	0.019**	0.018**	0.049	0.048	0.111***	0.106***
	(0.014)	(0.012)	(0.017)	(0.017)	(0.008)	(0.009)	(0.037)	(0.035)	(0.029)	(0.027)
β_1^h: Treat_i	0.057*	0.062**	0.001	0.026	0.119***	0.125***	0.070	0.118*	0.040	0.128
	(0.033)	(0.026)	(0.027)	(0.041)	(0.030)	(0.031)	(0.065)	(0.069)	(0.060)	(0.082)
β_2^h: Prior_i^h	0.857***	0.853***	0.777***	0.769***	1.019***	1.013***	0.878***	0.871***	0.754***	0.759***
	(0.050)	(0.047)	(0.104)	(0.103)	(0.032)	(0.032)	(0.133)	(0.124)	(0.070)	(0.064)
β_3^h: $\text{Prior}_i^h \times \text{Treat}_i$	-0.342**	-0.383***	-0.195	-0.307**	-0.652***	-0.617***	-0.301	-0.350	-0.221	-0.257
	(0.149)	(0.098)	(0.193)	(0.154)	(0.146)	(0.077)	(0.247)	(0.228)	(0.201)	(0.190)
γ_1^h: $\lvert\Delta CDF_i\rvert$		0.015		0.007		0.014		0.020		0.018*
		(0.009)		(0.015)		(0.010)		(0.013)		(0.010)
γ_2^h: $\lvert\Delta CDF_i\rvert \times \text{Treat}_i^h$		-0.029		-0.022		-0.027		-0.043		-0.106
		(0.018)		(0.035)		(0.032)		(0.043)		(0.066)
γ_3^h : $\lvert\Delta CDF_i\rvert \times \text{Prior}_i^h \times \text{Treat}_i$		-0.109**		-0.330***		-0.221***		0.043		0.072
		(0.045)		(0.112)		(0.048)		(0.063)		(0.086)
Point estimate: $\beta_2^h + \beta_3^h$	0.515	0.470	0.583	0.462	0.368	0.396	0.577	0.521	0.532	0.502
p-value for H_0: $\beta_2^h + \beta_3^h = 1$	0.001	0.000	0.010	0.000	0.000	0.000	0.042	0.012	0.013	0.005
p-value for H_0: $\beta_2^h = 1$	0.005	0.002	0.033	0.024	0.544	0.694	0.359	0.297	0.000	0.000
R^2	0.876	0.896	0.761	0.812	0.916	0.938	0.918	0.921	0.910	0.913
Observations	101	101	101	101	101	101	101	101	101	101

Notes: This table presents OLS estimates based on equation (4.1), leveraging our RCT among professional forecasters. The outcome is the posterior real GDP growth forecast, and the prior refers to the same variable elicited just before our RCT starts. Treat_i denotes our treatment indicator, and $|\Delta CDF_i| \geq 0$ is the standardized distance between the prior CDF belief and the truth, which is the key information provided to the treatment group, see Section 4.2 for a detailed explanation of the RCT. The CDF belief is the cumulative distribution function of the 2025Q4 forecasts evaluated at one's own forecast, both taken from the previous survey wave. The top panel shows point estimates and standard errors in parentheses, always robust to heteroskedasticity. Columns (1) and (2) report average coefficients across forecast horizons, as described in Section 4.4, with standard errors that are additionally clustered at the forecaster level. Columns (3) to (10) show corresponding estimates for each forecast horizon, where 2025Q4 refers to the nowcast because the survey experiment was fielded in the same quarter. The middle panel shows linear combinations of coefficients and corresponding hypothesis tests that are suggested by a Bayesian updating framework, as discussed in Section 4.4. The bottom panel shows additional regression statistics. Significance levels for regression estimates in the top panel are denoted by *** ($p < 0.01$), ** ($p < 0.05$), and * ($p < 0.10$).

is statistically significant at the 1% level.[17] It implies a weight on the prior of only 0.40 for the treatment group ($\beta_2^h + \beta_3^h$), which is statistically different from unity at all conventional significance levels. This contrasts with the control group, where the weight on the

[17] While this coefficient stands out in size, this is driven by the larger weight on the prior in the control group, as indicated by β_2^h. Indeed, the weight on the prior in the treatment group given by $\beta_2^h + \beta_3^h$ is comparable to the other estimates of the same statistic.

prior is close to unity and not statistically different from one at any level.[18] For forecast horizons further out in columns (8) and (10), we also obtain negative differential weights on the prior for treated participants, ranging from -0.26 to -0.35. While these estimates are not statistically significant, the overall weight on the prior in the treatment group is still significantly different from unity at the 5% level, as shown in the middle panel. Quantitatively, these weights on the prior range between 0.50 and 0.52, being not too different from the corresponding estimates for shorter forecast horizons.

Overall, the horizon-specific estimates suggest that the implied updating in the treatment group is similar across forecast horizons, although the significance of the average effects is primarily driven by shorter forecast horizons. This aligns well with our experimental design since the provided information concerns the nowcast, i.e., a forecast for the quarter in which the survey was fielded.

Accounting for treatment intensity. The treatment does not provide the same amount of information to every participant: those whose prior CDF beliefs were close to the truth received a small surprise, whereas those with large perception gaps received a large one. If treated forecasters incorporate the provided information, their updating should therefore increase in the size of the surprise. To test this, we discuss how our results depend on treatment intensity as captured by the triple interaction $|\Delta CDF_i| \times \text{Prior}_i^h \times \text{Treat}_i$.[19] The associated coefficient, γ_3^h, captures how the differential weight on the prior changes for a participant with a surprise being one standard deviation above the average surprise.

Focusing on the average across forecast horizons, we find that a standard deviation surprise leads to a change in the weight on the prior of -0.11. This reduction in the weight is statistically significant at the 5% level. The average effect is driven by the shorter forecast horizons, i.e., 2025Q4 and 2026Q1 in columns (4) and (6), respectively. These forecast

[18] While statistically significant at the 5% level, the intercept is very small, i.e., 0.02. This suggests that prior and posterior in the control group are very close on average.

[19] We do not discuss γ_1^h and γ_2^h since these estimates are not our coefficients of interest, straightforward to interpret, and largely insignificant.

horizon specific estimates are larger, ranging from -0.22 to -0.33, and both estimates are statistically significant at the 1% level. This contrasts with the estimates for 2026Q2 and 2026Q3 in columns (8) and (10), where we obtain point estimates that are very close to zero and statistically insignificant. The likely explanation for the differences across forecast horizons is our information treatment. As discussed above, the information provided is about 2025Q4, which may suggest that the information content is smaller for larger forecast horizons.

Relation to inaction. Building on our previous results, it is natural to ask how the belief revisions, within our RCT, connect with the lumpiness documented in Section 3. In the experiment, treated forecasters revise their reported forecasts within a few minutes, whereas according to the panel data, there is significant inaction between two distinct survey waves. Thus, to reconcile these two features, we augment the updating specification with the triple interaction $\text{MinRev}_i \times \text{Prior}_i^k \times \text{Treat}_i$, where MinRev_i denotes the standardized inaction threshold elicited in the survey module on lumpiness from Section 3. Intuitively, if fixed adjustment costs also operate within the experiment, treated forecasters with higher inaction thresholds should respond less to the provided rank information, i.e., retain more weight on their priors.

This is what we find. In our baseline specification in Table 2, the coefficient on the triple interaction is 0.163 on average across forecast horizons, and statistically significant at the 1% level. This means that a forecaster with a inaction threshold one standard deviation above the average inaction threshold has a weight on the prior that is 0.163 larger. The horizon-specific estimates are similar, ranging from 0.148 to 0.183, and all statistically significant at the 1% level. Thus, a higher inaction threshold partly offsets the treatment-induced reduction in the weight on the prior. The full specification, which additionally accounts for treatment intensity and its interaction with MinRev_i, delivers similar results, see Table B.7 in the appendix.

Overall, this suggests that the self-reported inaction thresholds capture a friction that remains at work even within our experiment.

Table 2: Bayesian updating regressions and the interaction with inaction thresholds

Dep. variable: Posterior	Avg. estimate	2025Q4	2026Q1	2026Q2	2026Q3
	(1)	(2)	(3)	(4)	(5)
β_0^h: Constant	0.054***	0.038**	0.019**	0.049	0.111***
	(0.014)	(0.017)	(0.008)	(0.037)	(0.029)
β_1^h: Treat_i	0.063***	0.012	0.110***	0.083	0.049
	(0.023)	(0.029)	(0.024)	(0.051)	(0.051)
β_2^h: Prior_i^h	0.857***	0.777***	1.019***	0.878***	0.754***
	(0.050)	(0.104)	(0.032)	(0.133)	(0.070)
β_3^h: $\text{Prior}_i^h \times \text{Treat}_i$	-0.379***	-0.279	-0.608***	-0.363**	-0.266*
	(0.094)	(0.190)	(0.088)	(0.175)	(0.161)
δ_1^h : $\text{MinRev}_i \times \text{Prior}_i^h \times \text{Treat}_i$	0.163***	0.183***	0.154***	0.165***	0.148***
	(0.039)	(0.050)	(0.040)	(0.045)	(0.046)
R^2	0.681	0.639	0.802	0.699	0.583
Observations	404	101	101	101	101

Notes: This table presents OLS estimates based on equation (4.1), leveraging our RCT among professional forecasters. We include $\text{MinRev}_i \times \text{Prior}_i^h \times \text{Treat}_i$ as an additional regressor, where MinRev_i denotes the self-reported minimum revision size from the fixed cost module introduced in Section 3. The outcome is the posterior real GDP growth forecast, and the prior refers to the same variable elicited just before our RCT starts. Treat_i denotes our treatment indicator, and $|\Delta CDF_i| \geq 0$ is the standardized distance between the prior CDF belief and the truth, which is the key information provided to the treatment group; see Section 4.2 for a detailed explanation of the RCT. The CDF belief is the cumulative distribution function of the 2025Q4 forecasts evaluated at one's own forecast, both taken from the previous survey wave. The top panel shows point estimates and standard errors in parentheses, always robust to heteroskedasticity. Column (1) reports average coefficients across forecast horizons, as described in Section 4.4, with standard errors that are additionally clustered at the forecaster level. Columns (2) to (5) show corresponding estimates for each forecast horizon, where 2025Q4 refers to the nowcast because the survey experiment was fielded in the same quarter. Significance levels for regression estimates in the top panel are denoted by *** ($p < 0.01$), ** ($p < 0.05$), and * ($p < 0.10$).

5 The interaction of lumpiness and strategic behavior

The previous sections studied two margins of frictional forecasting. Section 3 showed that forecasts are lumpy, while Section 4 used the survey experiment to show that information about the forecast distribution affects reported forecasts. This section connects both margins in the panel.

The experiment highlights two objects: the distance between a forecaster's report and the consensus, and the distance between a forecaster's position in the distribution and her usual rank. If forecast reporting is shaped by both inaction frictions and positioning motives, these objects should also organize revision behavior outside the experiment. In particular, inherited forecasts that are far from the consensus or far from the forecaster's usual rank position should be more likely to be revised.

The two gaps also have different implications for forecast performance. Deviations from the consensus should predict larger forecast errors if the consensus contains useful information. By contrast, deviations from a forecaster's usual rank may predict smaller forecast errors if moving away from that usual position reflects a stronger accuracy motive and a weaker role for strategic positioning.

We therefore estimate two sets of panel regressions. The first studies the extensive margin of whether a forecast is revised. The second studies absolute forecast errors. Together, the regressions show how the two forces emphasized throughout the paper—consensus information and rank positioning—interact with forecast lumpiness.

5.1 The extensive margin

Let $f^i_{t-1,h+1}$ denote forecaster i's inherited forecast for a fixed event, and let $f^i_{t,h}$ denote the current forecast for the same event. The revision indicator is $Rev^i_{t,h} = \mathbb{1}\left\{\left|f^i_{t,h} - f^i_{t-1,h+1}\right| > 0\right\}$. We relate this revision decision to two lagged gaps. The first is the absolute consensus gap, $\left|C_{t-1,h+1} - f^i_{t-1,h+1}\right|$, where $C_{t-1,h+1}$ is the lagged consensus forecast. The second is the

absolute rank gap, $\left|CDF^i_{t-1,h+1} - \overline{CDF}_i\right|$, where $CDF^i_{t-1,h+1}$ is forecaster i's rank in the cross-sectional forecast distribution, evaluated at the inherited forecast, and $\overline{CDF}_i$ is forecaster i's average rank in the panel. We interpret $\overline{CDF}_i$ as a proxy for the forecaster's usual position in the distribution.

We estimate

$$Rev^i_{t,h} \;=\; \zeta_0 \;+\; \zeta_1 \left|C_{t-1,h+1} - f^i_{t-1,h+1}\right| \;+\; \zeta_2 \left|CDF^i_{t-1,h+1} - \overline{CDF}_i\right| \;+\; u^i_{t,h}. \quad (5.1)$$

The coefficient ζ_1 measures whether forecasters are more likely to revise when their inherited forecast is farther from the consensus. The coefficient ζ_2 measures whether forecasters are more likely to revise when their inherited rank is farther from their usual rank.

Table 3 shows that both gaps predict revision. Forecasters are more likely to revise when their inherited forecast is farther from the consensus, and they are also more likely to revise when their inherited rank is farther from their average rank. The coefficients on both variables are positive and statistically significant in the average specification and across forecast horizons.

Table 3: Consensus and rank gaps predict revision likelihood

Dep. variable:	Avg. estimate		$h=-1$		$h=0$		$h=1$		$h=2$	
Revision indicator										
$\mathbb{1}\{\lvert f^i_{t,h} - f^i_{t-1,h+1}\rvert > 0\}$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Lagged abs. consensus gap:		0.058***		0.046***		0.082***		0.037***		0.066*
$\lvert C_{t-1,h+1} - f^i_{t-1,h+1}\rvert$		(0.013)		(0.013)		(0.013)		(0.012)		(0.036)
Lagged abs. rank gap:	0.108***	0.095***	0.063***	0.050***	0.075***	0.055***	0.049***	0.035**	0.245***	0.238***
$\lvert CDF^i_{t-1,h+1} - \overline{CDF^i_{t-1,h+1}}\rvert$	(0.010)	(0.010)	(0.015)	(0.014)	(0.018)	(0.018)	(0.016)	(0.016)	(0.020)	(0.017)
R^2	0.019	0.025	0.007	0.012	0.009	0.021	0.004	0.010	0.055	0.059
Observations	9797	9797	3520	3520	4039	4039	1880	1880	358	358

Notes: The table reports panel regressions excluding 2020.

5.2 Forecast performance

We next ask whether these two gaps have different implications for forecast accuracy. The outcome is the absolute forecast error, $FE^i_{t,h} = \left|y_{t+h} - f^i_{t,h}\right|$. We estimate

$$\begin{aligned} FE^i_{t,h} &= \xi_0 + \xi_1 \left|C_{t-1,h+1} - f^i_{t-1,h+1}\right| + \xi_2 \left|CDF^i_{t-1,h+1} - \overline{CDF}_i\right| \\ &+ \xi_3 \left|f^i_{t,h} - f^i_{t-1,h+1}\right| + e^i_{t,h}. \end{aligned} \tag{5.2}$$

The coefficient ξ_1 measures whether deviations from the consensus predict larger forecast errors. The coefficient ξ_2 measures whether deviations from the forecaster's usual rank predict forecast performance. The coefficient ξ_3 captures the relation between revision size and subsequent forecast errors.

Table 4 shows that the two gaps have opposite implications for forecast performance. The lagged absolute consensus gap strongly predicts larger absolute forecast errors. This is consistent with the consensus containing useful information: forecasters whose inherited reports are farther from the consensus subsequently make larger errors.

Table 4: Consensus and rank gaps predict absolute forecast errors

Dep. variable: Forecast error $\lvert y_{t+h} - f^i_{t,h}\rvert$	Avg. estimate		$h=-1$		$h=0$		$h=1$		$h=2$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Lagged abs. consensus gap: $\lvert C_{t-1,h+1} - f^i_{t-1,h+1}\rvert$	0.626***	0.505***	0.503***	0.440***	0.572***	0.368***	0.517***	0.397***	0.910***	0.816***
	(0.045)	(0.097)	(0.045)	(0.062)	(0.031)	(0.071)	(0.068)	(0.121)	(0.116)	(0.225)
Lagged abs. rank gap: $\lvert CDF^i_{t-1,h+1} - \overline{CDF^i_{t-1,h+1}}\rvert$	-0.090**	-0.109**	-0.107***	-0.112***	-0.114***	-0.131***	-0.200***	-0.217***	0.062	0.026
	(0.041)	(0.044)	(0.026)	(0.025)	(0.014)	(0.013)	(0.050)	(0.049)	(0.165)	(0.176)
Abs. forecast revision: $\lvert f^i_{t,h} - f^i_{t-1,h+1}\rvert$		0.194***		0.089*		0.292***		0.199**		0.196
		(0.074)		(0.049)		(0.069)		(0.098)		(0.232)
R^2	0.262	0.289	0.266	0.272	0.231	0.295	0.226	0.250	0.325	0.338
Observations	9322	9322	3520	3520	4039	4039	1576	1576	187	187

Notes: The table reports panel regressions excluding 2020.

The lagged absolute rank gap has the opposite sign. Except at the longest horizon, larger deviations from the forecaster's average rank predict smaller absolute forecast errors. This pattern is consistent with the interpretation that a forecaster's usual position in the

distribution partly reflects reporting motives. When a forecaster moves farther away from that usual rank, the reported forecast appears to place more weight on accuracy and less weight on maintaining a preferred relative position.

Finally, larger absolute revisions predict larger forecast errors. This is consistent with large revisions occurring in periods of substantial news or overreaction.

The interaction between lumpiness and strategic behavior is therefore visible in both outcomes. On the extensive margin, both the consensus gap and the rank gap predict whether forecasters revise. For forecast performance, however, the two gaps point in opposite directions: distance from the consensus predicts worse performance, while distance from the usual rank position predicts better performance. This contrast supports the paper's main interpretation that consensus information and rank positioning are distinct forces in professional forecast reporting.

6 Conclusion

We argue that reported predictions by professional forecasters are not pure measures of beliefs. Reporting frictions create a wedge between what forecasters believe and what they report, so professional forecasts are frictional forecasts. We document this wedge along several dimensions. First, forecasts are lumpy: zero revisions are frequent, while non-zero revisions are large. Second, direct survey evidence shows that this lumpiness is not only due to the absence of new information. Forecasters themselves report adjustment costs, credibility concerns, decision relevance, and the consensus forecast as reasons for leaving forecasts unchanged. Third, using a novel randomized experiment among professional forecasters, we show that information about the forecast distribution affects reported forecasts, providing causal evidence consistent with strategic reporting. Finally, we show that the extensive and intensive margins interact: the same consensus and rank-related gaps that matter in the experiment also predict revision decisions and forecast performance in the panel.

These findings have two broader implications. First, reporting frictions matter beyond the measurement of expectations. Reported professional forecasts may influence the expectations of households, firms, financial markets, and policymakers (e.g., Carroll, 2003). If these reports are frictional, then reporting behavior itself can shape the transmission of macroeconomic information. Second, reporting frictions affect how researchers should interpret survey forecasts. Many empirical tests of expectation formation treat reported forecasts as direct observations of beliefs (e.g., Bordalo et al., 2020). Our evidence suggests that such tests may confound belief formation with reporting behavior. Accounting for frictional forecast reporting is therefore important for understanding both professional expectations and the macroeconomic role of public forecast surveys.

References

Adam, K., P. Kuang, and S. Xie (2025): "Overconfidence in Private Information Explains Biases in Professional Forecasts," *Journal of Monetary Economics*, 103839.

Afrouzi, H., S. Y. Kwon, A. Landier, Y. Ma, and D. Thesmar (2023): "Overreaction in Expectations: Evidence and Theory," *Quarterly Journal of Economics*, 138, 1713–1764.

Andrade, P. and H. Le Bihan (2013): "Inattentive Professional Forecasters," *Journal of Monetary Economics*, 60, 967–982.

Assenza, T., T. Bao, C. Hommes, and D. Massaro (2014): "Experiments on Expectations in Macroeconomics and Finance," in *Experiments in Macroeconomics*, Emerald Group Publishing Limited, vol. 17, 11–70.

Baley, I. and J. Turén (2025a): "Lumpy Forecasts," *Available at SSRN 4988612*.

——— (2025b): "State-Dependent Forecasting in Volatile Times," Mimeo.

Bordalo, P., N. Gennaioli, Y. Ma, and A. Shleifer (2020): "Overreaction in Macroeconomic Expectations," *American Economic Review*, 110, 2748–2782.

Broer, T. and A. N. Kohlhas (2024): "Forecaster (Mis-) Behavior," *Review of Economics and Statistics*, 106, 1334–1351.

Brückbauer, F. and M. Schröder (2023): "The ZEW Financial Market Survey Panel," *Journal of Economics and Statistics*, 243, 451–469.

Carroll, C. D. (2003): "Macroeconomic Expectations of Households and Professional Forecasters," *Quarterly Journal of economics*, 118, 269–298.

Coibion, O. and Y. Gorodnichenko (2015): "Information rigidity and the expectations formation process: A simple framework and new facts," *American Economic Review*, 105, 2644–2678.

——— (2026): "Schumpeter Lecture 2025: The New Causal Macroeconomics of Surveys and Experiments," *Journal of the European Economic Association*, 24, 58–81.

Coibion, O., Y. Gorodnichenko, and S. Kumar (2018): "How Do Firms Form Their Expectations? New Survey Evidence," *American Economic Review*, 108, 2671–2713.

Coibion, O., Y. Gorodnichenko, and T. Ropele (2020): "Inflation Expectations and Firm Decisions: New Causal Evidence," *Quarterly Journal of Economics*, 135, 165–219.

Ehrbeck, T. and R. Waldmann (1996): "Why Are Professional Forecasters Biased? Agency Versus Behavioral Explanations," *Quarterly Journal of Economics*, 111, 21–40.

Farmer, L. E., E. Nakamura, and J. Steinsson (2024): "Learning about the Long Run," *Journal of Political Economy*, 132, 3334–3377.

Gemmi, L. and R. Valchev (2025): "Biased Surveys," *Journal of Monetary Economics*,

103868.

Haaland, I., C. Roth, and J. Wohlfart (2023): "Designing Information Provision Experiments," *Journal of Economic Literature*, 61, 3–40.

Juodis, A. and S. Kučinskas (2023): "Quantifying Noise in Survey Expectations," *Quantitative Economics*, 14, 609–650.

Kerssenfischer, M. and M. Schmeling (2024): "What Moves Markets?" *Journal of Monetary Economics*, 145, 103560.

Ottaviani, M. and P. N. Sørensen (2006): "The Strategy of Professional Forecasting," *Journal of Financial Economics*, 81, 441–466.

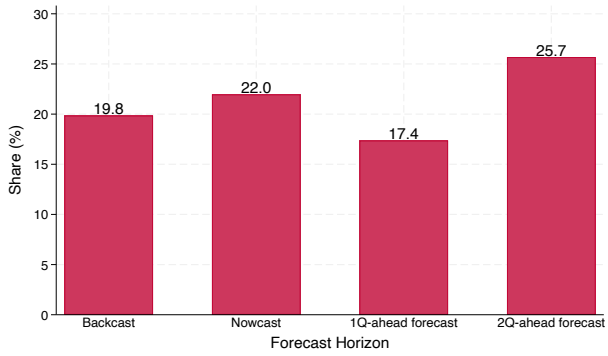
Weber, M., B. Candia, H. Afrouzi, T. Ropele, R. Lluberas, S. Frache, B. Meyer, S. Kumar, Y. Gorodnichenko, D. Georgarakos, et al. (2025): "Tell Me Something I Don't Already Know: Learning in Low-and High-Inflation Settings," *Econometrica*, 93, 229–264.

Appendix

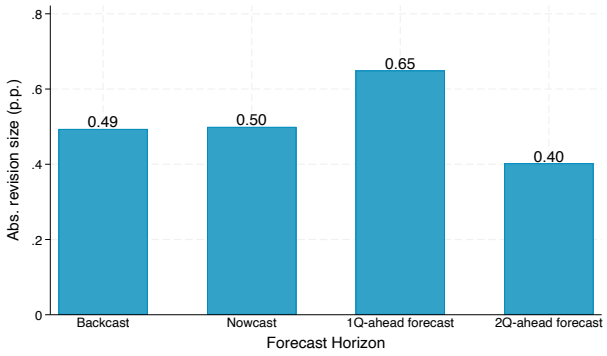
Appendix A Supplementary Figures

Figure A.1: Lumpiness in the panel, include 2020

(a) Share of zero revisions

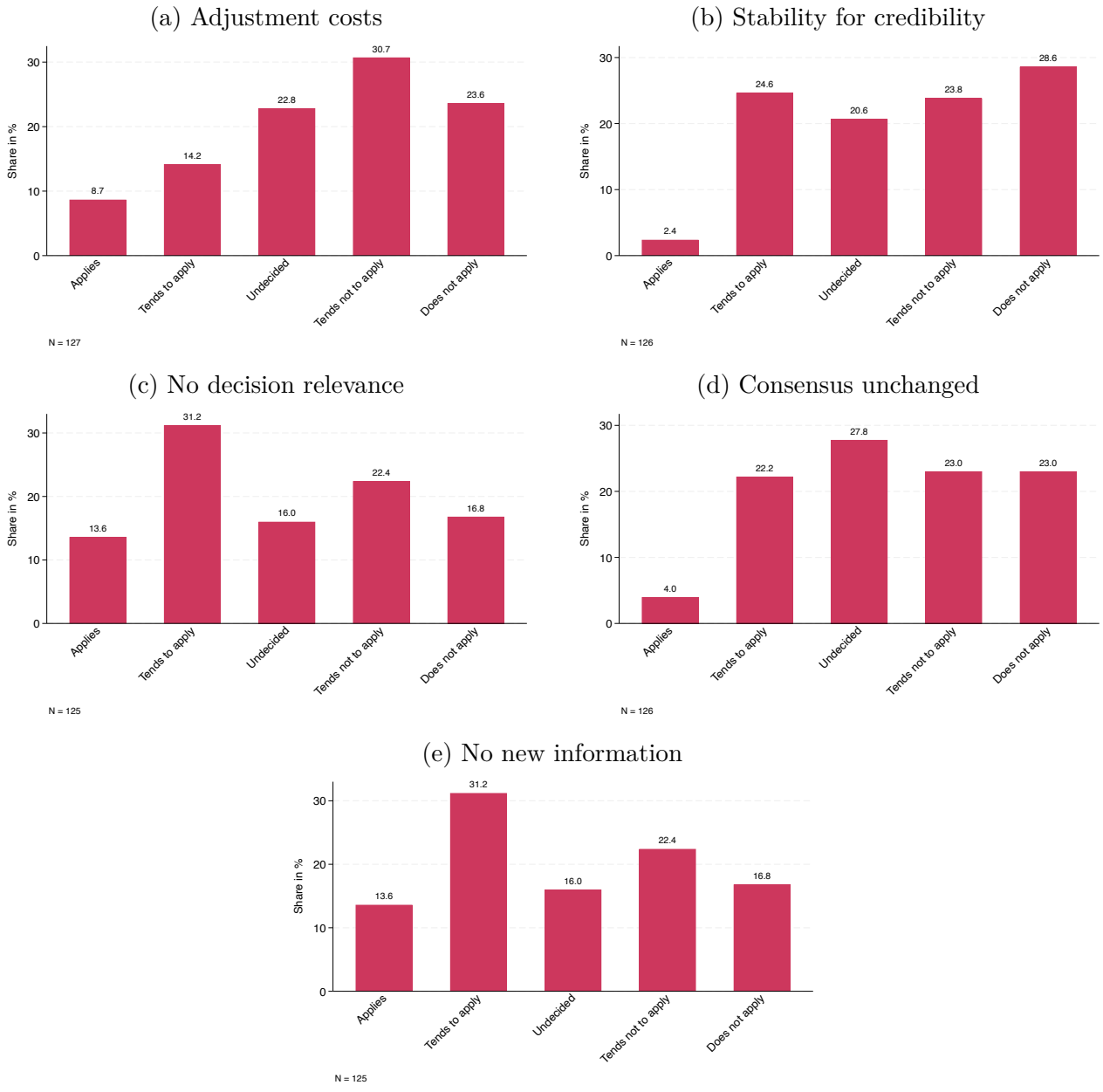


(b) Average non-zero revision size



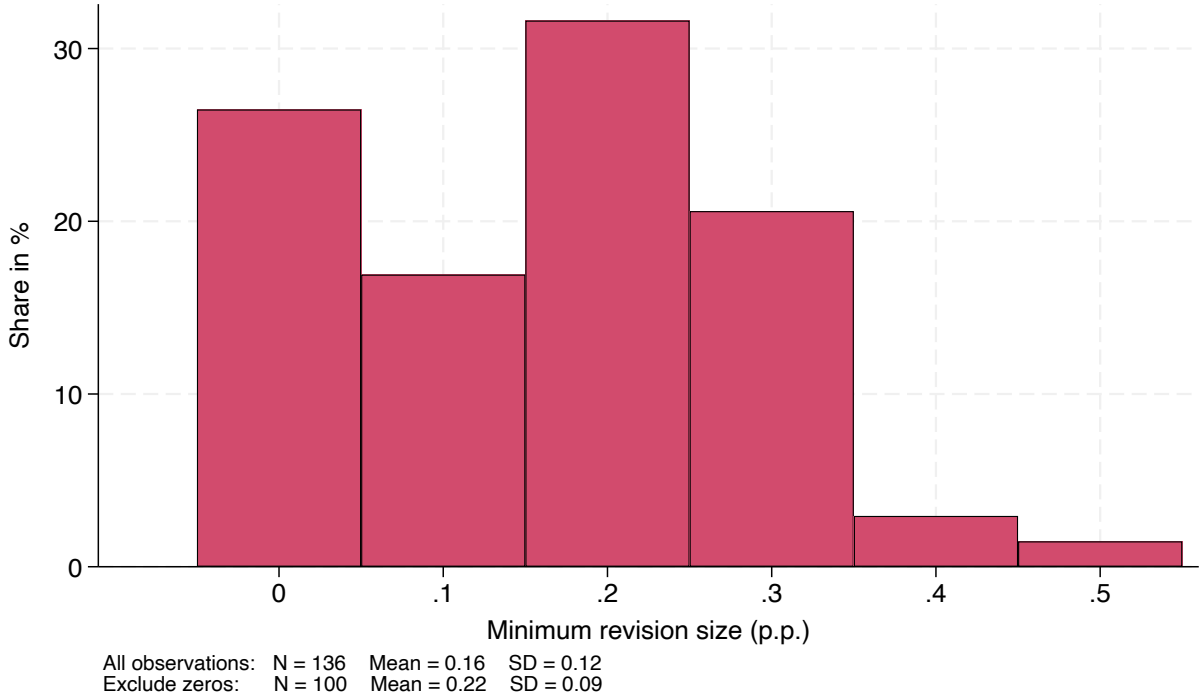
Notes: This figure shows statistics computed from the panel data by forecast horizon, including 2020. Backcast refers to the forecast for the previous quarter (before the official estimate is released). Nowcast refers to the prediction for the current quarter (in which the survey is conducted. Panel (a) shows the share of revisions between two consecutive survey waves for a fixed event. Panel (b) shows the average absolute revision size, conditional on a non-zero revision.

Figure A.2: Motives for lumpiness, full Likert scale



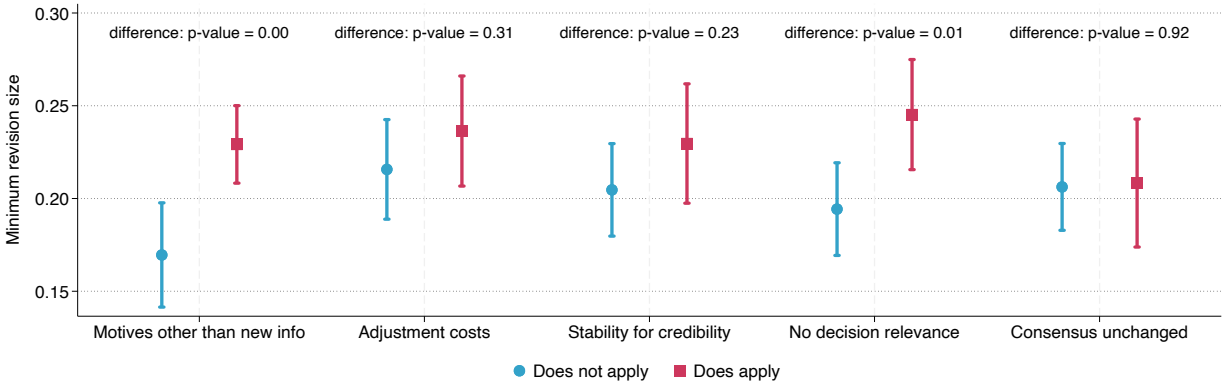
Notes: This figure shows reported motives for not revising a forecast based on a five-point Likert scale, for details see Appendix C. In the bottom left, we indicate the number of respondents N.

Figure A.3: Histogram of reported inaction thresholds



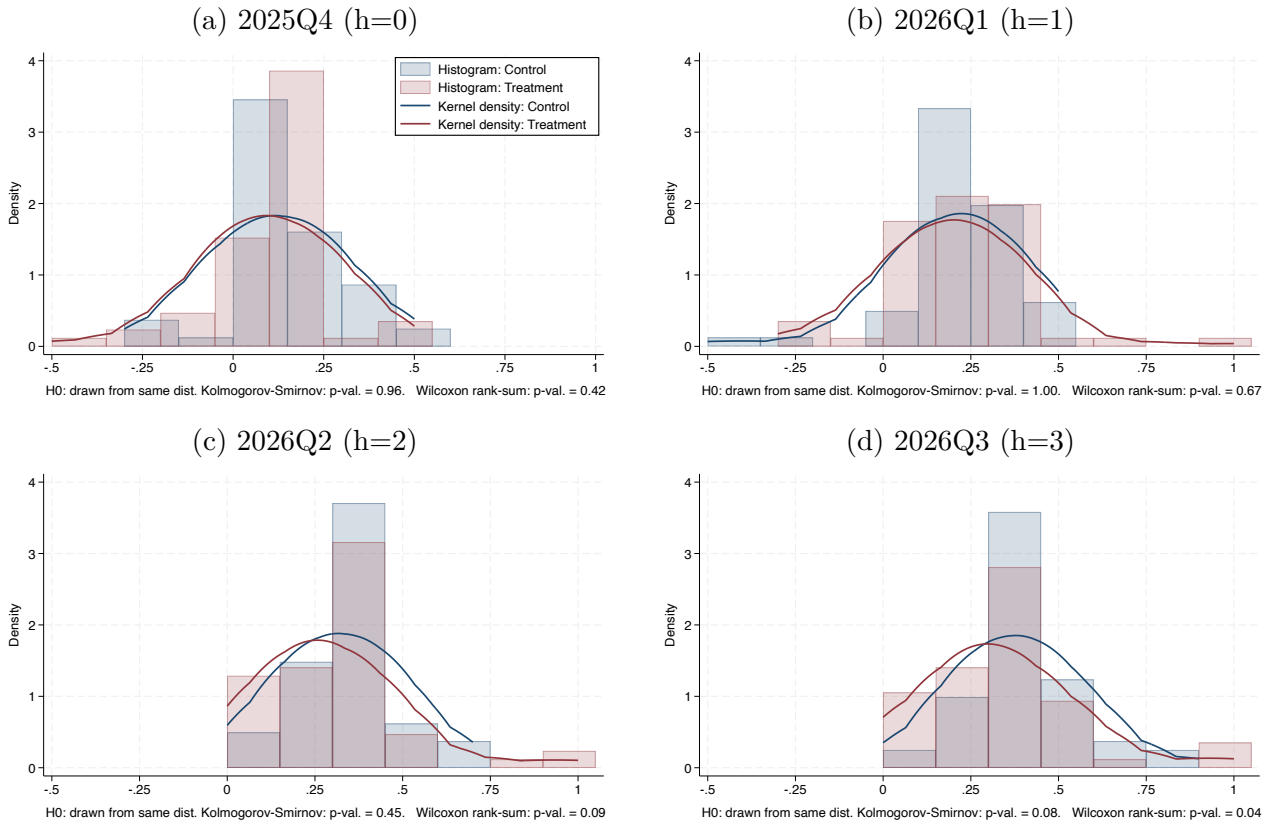
Notes: ...

Figure A.4: Relation of inaction threshold to inaction reasons



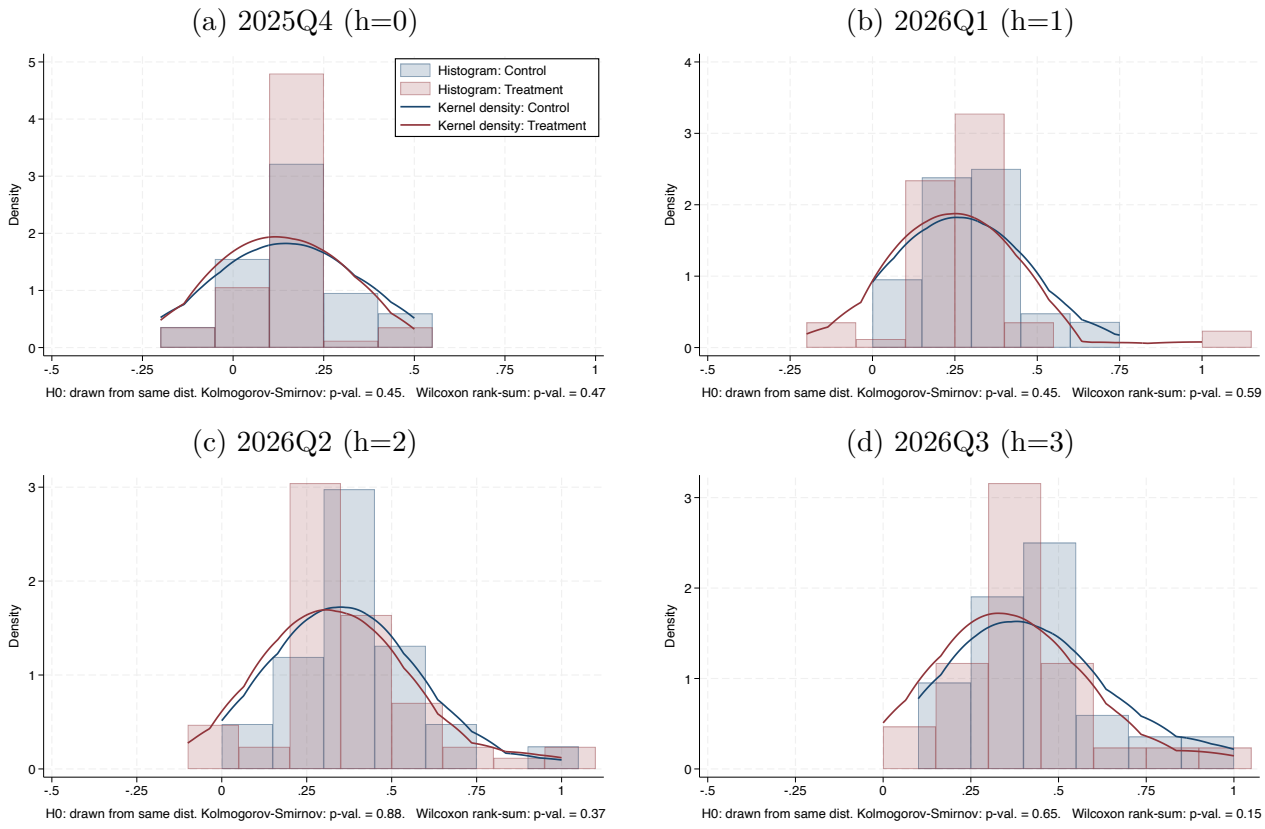
Notes: This figure shows the histogram of self-reported inaction thresholds in percentage points. Respondents that indicate to always revise are subsumed in the zero bin.

Figure A.5: Comparing treatment and control groups based on the real GDP growth forecasts



Notes: This figure compares the prior distributions across treatment and control groups. We show histograms of real GDP growth forecasts across forecast horizons , as well as the corresponding kernel density estimate using the Epanechnikov kernel with a fixed bandwidth of 0.15. Lastly, the notes below each panel provide the p-values for the Kolmogorov-Smirnov and the Wilcoxon rank-sum test for the null hypothesis that both treatment and control groups are drawn from the same underlying distribution for the corresponding forecast horizon.

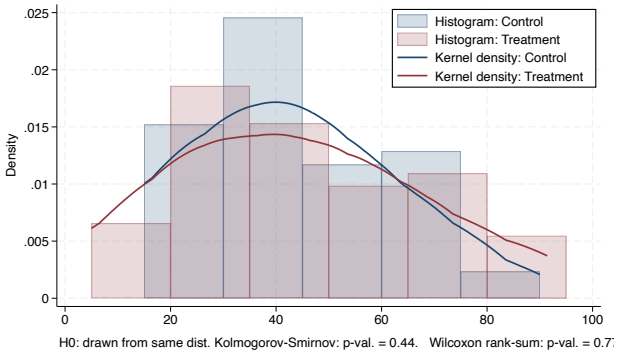
Figure A.6: Comparing treatment and control groups based on the consensus beliefs



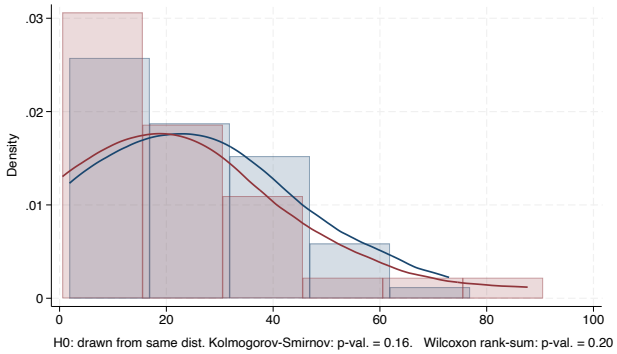
Notes: This figure compares the prior distributions across treatment and control groups. We show histograms of consensus beliefs over real GDP growth forecasts across forecast horizons, as well as the corresponding kernel density estimate using the Epanechnikov kernel with a fixed bandwidth of 0.15. Lastly, the notes below each panel provide the p-values for the Kolmogorov-Smirnov and the Wilcoxon rank-sum test for the null hypothesis that both treatment and control groups are drawn from the same underlying distribution for the corresponding forecast horizon.

Figure A.7: Comparing treatment and control groups based on the CDF beliefs

(a) Prior CDF belief



(b) Absolute perception gap



Notes: This figure compares the prior distributions across treatment and control groups. We show histograms of CDF belief and the implied absolute perception gap, as well as the corresponding kernel density estimate using the Epanechnikov kernel with a fixed bandwidth of 15. Lastly, the notes below each panel provide the p-values for the Kolmogorov-Smirnov and the Wilcoxon rank-sum test for the null hypothesis that both treatment and control groups are drawn from the same underlying distribution for the corresponding forecast horizon.

Appendix B Supplementary Tables

Table B.1: Predicting revision sizes by self-reported inaction thresholds

Dep. variable: Abs. forecast revision $\lvert f^i_{t,h} - f^i_{t-1,h+1}\rvert \times \mathbb{1}\{\lvert f^i_{t,h} - f^i_{t-1,h+1}\rvert > 0\}$	Avg. estimate	$h=-1$	$h=0$	$h=1$	$h=2$
	(1)	(2)	(3)	(4)	(5)
Min. revision size$_i$	0.422**	0.306	0.459**	0.574**	0.349
(self-reported in February 2026)	(0.164)	(0.264)	(0.203)	(0.239)	(0.247)
R^2	0.004	0.001	0.004	0.006	0.006
Observations	2987	925	1116	742	204

Notes: This table presents OLS estimates based on $\lvert f^i_{t,h} - f^i_{t-1,h+1}\rvert = \alpha^h + \beta^h$ Min. revision size$_i + \varepsilon^i_{t,h}$, which is estimated on all non-zero revisions in the panel, i.e., $\lvert f^i_{t,h} - f^i_{t-1,h+1}\rvert \neq 0$, but excluding 2020. Standard errors in parentheses are clustered at the forecaster and time level. Significance levels are denoted by *** ($p < 0.01$), ** ($p < 0.05$), and * ($p < 0.10$).

Table B.2: Predicting revision sizes by self-reported inaction thresholds, include 2020

Dep. variable: Abs. forecast revision $\lvert f^i_{t,h} - f^i_{t-1,h+1}\rvert \times \mathbb{1}\{\lvert f^i_{t,h} - f^i_{t-1,h+1}\rvert > 0\}$	Avg. estimate	$h=-1$	$h=0$	$h=1$	$h=2$
	(1)	(2)	(3)	(4)	(5)
Min. revision size$_i$	0.390*	0.310	0.296	0.604***	0.349
(self-reported in February 2026)	(0.209)	(0.284)	(0.364)	(0.191)	(0.247)
R^2	0.003	0.001	0.001	0.004	0.006
Observations	3213	1015	1204	790	204

Notes: This table presents OLS estimates based on $\lvert f^i_{t,h} - f^i_{t-1,h+1}\rvert = \alpha^h + \beta^h$ Min. revision size$_i + \varepsilon^i_{t,h}$, which is estimated on all non-zero revisions in the panel, i.e., $\lvert f^i_{t,h} - f^i_{t-1,h+1}\rvert \neq 0$, and including 2020. Standard errors in parentheses are clustered at the forecaster and time level. Significance levels are denoted by *** ($p < 0.01$), ** ($p < 0.05$), and * ($p < 0.10$).

Table B.3: Imperfect pass-through from short- to long-horizon revisions

Dep. variable: Revision indcator $\mathbb{1}\{\lvert f^i_{t,h} - f^i_{t-1,h+1}\rvert > 0\}$	Avg. estimate		$h=-1$		$h=0$	
	(1)	(2)	(3)	(4)	(5)	(6)
Imperfect pass through from $j=1$	0.125***		0.121***		0.128***	
Testing: $1-\alpha^h-\beta^h$	(0.010)		(0.012)		(0.014)	
Imperfect pass through from $j=1,2$		0.128***		0.121***		0.135***
Testing: $1-\alpha^h-\beta^h$		(0.011)		(0.012)		(0.016)
R^2	0.031	0.027	0.007	0.007	0.056	0.047
Observations	3313	3313	1367	1367	1946	1946

Notes: This table presents OLS estimates based on $\mathbb{1}\{\lvert f^i_{t,h} - f^i_{t-1,h+1}\rvert > 0\} = \alpha^h + \beta^h\ \mathbb{1}\{\lvert f^i_{t,j} - f^i_{t-1,j+1}\rvert > 0\} + \varepsilon^i_{t,h}$, estimated on a sample excluding 2020. The reported coefficients are given by $1-\alpha^h-\beta^h$, so that positive values indicate a pass-through from short to long revisions below unity. Standard errors in parentheses are clustered at the forecaster and time level. Significance levels are denoted by *** ($p < 0.01$), ** ($p < 0.05$), and * ($p < 0.10$).

Table B.4: Imperfect pass-through from short- to long-horizon revisions, include 2020

Dep. variable: Revision indcator $\mathbb{1}\{\lvert f^i_{t,h} - f^i_{t-1,h+1}\rvert > 0\}$	Avg. estimate		$h=-1$		$h=0$	
	(1)	(2)	(3)	(4)	(5)	(6)
Imperfect pass through from $j=1$	0.122***		0.116***		0.127***	
Testing: $1-\alpha^h-\beta^h$	(0.009)		(0.011)		(0.012)	
Imperfect pass through from $j=1,2$		0.125***		0.116***		0.134***
Testing: $1-\alpha^h-\beta^h$		(0.010)		(0.011)		(0.015)
R^2	0.031	0.027	0.010	0.010	0.053	0.045
Observations	3619	3619	1506	1506	2113	2113

Notes: This table presents OLS estimates based on $\mathbb{1}\{\lvert f^i_{t,h} - f^i_{t-1,h+1}\rvert > 0\} = \alpha^h + \beta^h\ \mathbb{1}\{\lvert f^i_{t,j} - f^i_{t-1,j+1}\rvert > 0\} + \varepsilon^i_{t,h}$, estimated on a sample including 2020. The reported coefficients are given by $1-\alpha^h-\beta^h$, so that positive values indicate a pass-through from short to long revisions below unity. Standard errors in parentheses are clustered at the forecaster and time level. Significance levels are denoted by *** ($p<0.01$), ** ($p<0.05$), and * ($p<0.10$).

Table B.5: Explaining revision indicators with fixed effects

Dep. variable: Revision indicator $\mathbb{1}\{\lvert f^i_{t,h} - f^i_{t-1,h+1}\rvert > 0\}$	Avg. estimate		$h=-1$		$h=0$		$h=1$		$h=2$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Forecaster FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Time FE		✓		✓		✓		✓		✓
R^2	0.184	0.210	0.126	0.149	0.141	0.173	0.144	0.175	0.326	0.342
Observations	9947	9947	3528	3528	4081	4081	1914	1914	424	424

Notes: This table presents OLS estimates based on $\mathbb{1}\{\lvert f^i_{t,h} - f^i_{t-1,h+1}\rvert > 0\} = \alpha^h_i + \alpha^h_t + \varepsilon^i_{t,h}$, estimated on a sample excluding 2020. We report the R^2 showing how much variation can be explained by forecaster and time fixed effects.

Table B.6: Explaining revision indicators with fixed effects, include 2020

Dep. variable: Revision indcator $\mathbb{1}\{\lvert f^i_{t,h} - f^i_{t-1,h+1}\rvert > 0\}$	Avg. estimate		$h=-1$		$h=0$		$h=1$		$h=2$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Forecaster FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Time FE		✓		✓		✓		✓		✓
R^2	0.179	0.207	0.120	0.147	0.133	0.169	0.138	0.169	0.326	0.342
Observations	10780	10780	3854	3854	4418	4418	2084	2084	424	424

Notes: This table presents OLS estimates based on $\mathbb{1}\{\lvert f^i_{t,h} - f^i_{t-1,h+1}\rvert > 0\} = \alpha^h_i + \alpha^h_t + \varepsilon^i_{t,h}$, estimated on a sample including 2020. We report the R^2 showing how much variation can be explained by forecaster and time fixed effects.

Table B.7: Bayesian updating regressions and the interaction with inaction thresholds, full specification

Dep. variable: Posterior	Avg. estimate	2025Q4	2026Q1	2026Q2	2026Q3
	(1)	(2)	(3)	(4)	(5)
β_0^h: Constant	0.053***	0.038**	0.018**	0.048	0.106***
	(0.012)	(0.017)	(0.009)	(0.035)	(0.027)
β_1^h: Treat_i	0.051***	0.009	0.104***	0.062	0.029
	(0.019)	(0.027)	(0.027)	(0.051)	(0.045)
β_2^h: Prior_i^h	0.853***	0.769***	1.013***	0.871***	0.759***
	(0.047)	(0.103)	(0.032)	(0.124)	(0.064)
β_3^h: $\text{Prior}_i^h \times \text{Treat}_i$	-0.339***	-0.285	-0.603***	-0.261	-0.209
	(0.072)	(0.215)	(0.122)	(0.168)	(0.131)
γ_1^h: $\lvert\Delta CDF_i\rvert$	0.015	0.007	0.014	0.020	0.018*
	(0.009)	(0.015)	(0.010)	(0.013)	(0.010)
γ_2^h: $\lvert\Delta CDF_i\rvert \times \text{Treat}_i$	-0.003	-0.013	-0.019	0.024	-0.005
	(0.015)	(0.021)	(0.019)	(0.024)	(0.037)
γ_3^h : $\lvert\Delta CDF_i\rvert \times \text{Prior}_i^h \times \text{Treat}_i$	-0.221***	-0.313*	-0.154**	-0.209***	-0.208***
	(0.038)	(0.160)	(0.075)	(0.068)	(0.076)
δ_1^h : $\text{MinRev}_i \times \text{Prior}_i^h \times \text{Treat}_i$	0.068**	0.025	0.070	0.095***	0.081***
	(0.032)	(0.091)	(0.051)	(0.035)	(0.025)
δ_2^h : $\text{MinRev}_i \times \lvert\Delta CDF_i\rvert \times \text{Prior}_i^h \times \text{Treat}_i$	0.065**	0.034	0.004	0.101***	0.122***
	(0.033)	(0.129)	(0.064)	(0.029)	(0.024)
R^2	0.729	0.689	0.825	0.747	0.653
Observations	404	101	101	101	101

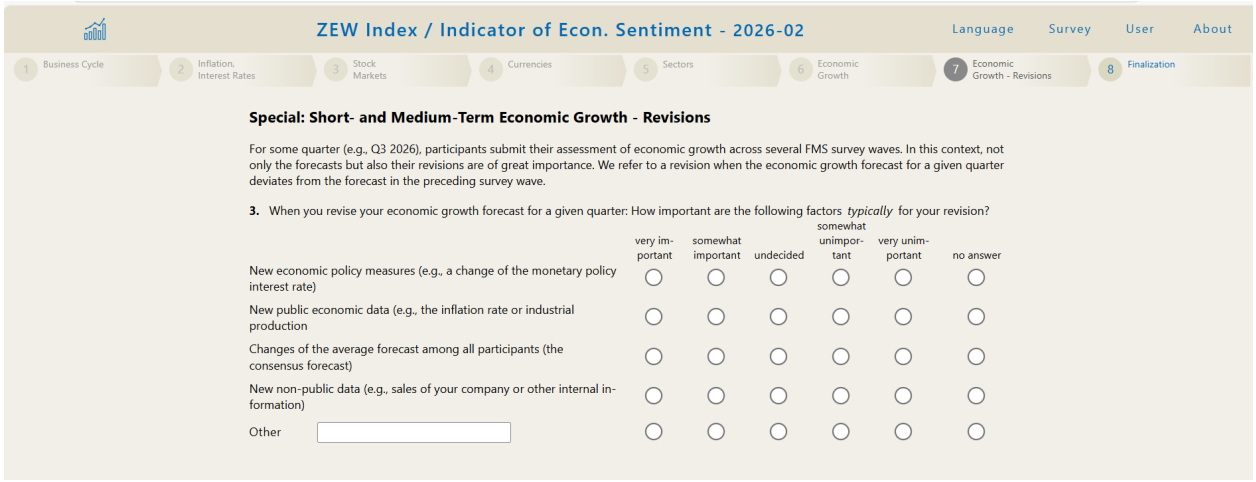
Notes: This table presents OLS estimates based on equation (4.1), leveraging our RCT among professional forecasters. We include $\text{MinRev}_i \times \text{Prior}_i^k \times \text{Treat}_i$ and $\text{MinRev}_i \times \lvert\Delta CDF_i\rvert \times \text{Prior}_i^k \times \text{Treat}_i$ as additional regressors, where MinRev_i denotes the self-reported minimum revision size from the fixed cost module introduced in Section 3. The outcome is the posterior real GDP growth forecast, and the prior refers to the same variable elicited just before our RCT starts. Treat_i denotes our treatment indicator, and $\lvert\Delta CDF_i\rvert \geq 0$ is the standardized distance between the prior CDF belief and the truth, which is the key information provided to the treatment group, see Section 4.2 for a detailed explanation of the RCT. The CDF belief is the cumulative distribution function of the 2025Q4 forecasts evaluated at one's own forecast, both taken from the previous survey wave. The top panel shows point estimates and standard errors in parentheses, always robust to heteroskedasticity. Column (1) reports average coefficients across forecast horizons, as described in Section 4.4, with standard errors that are additionally clustered at the forecaster level. Columns (2) to (5) show corresponding estimates for each forecast horizon, where 2025Q4 refers to the nowcast because the survey experiment was fielded in the same quarter. Significance levels for regression estimates in the top panel are denoted by *** ($p < 0.01$), ** ($p < 0.05$), and * ($p < 0.10$).

Appendix C Questionnaire

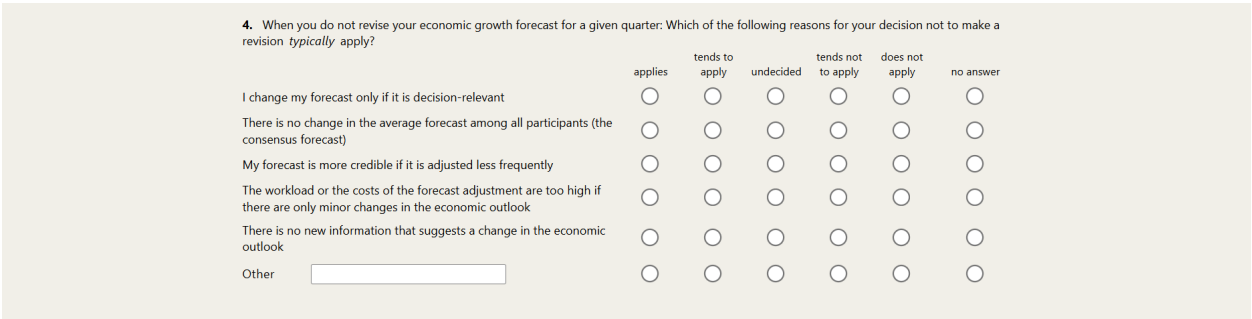
C.1 Fixed cost question module (February 2026)

Figure C.1: Motives

(a) Motives for revising



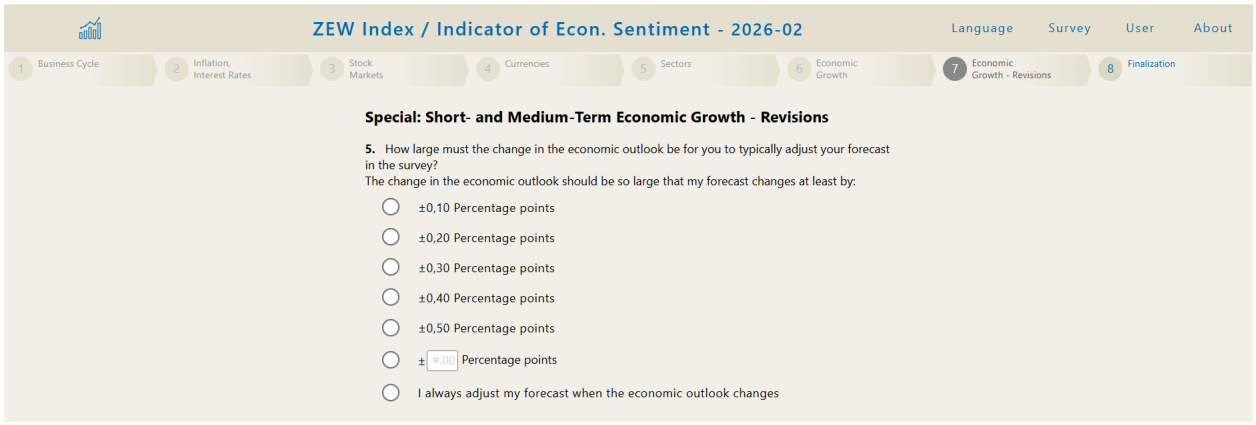
(b) Motives for not revising



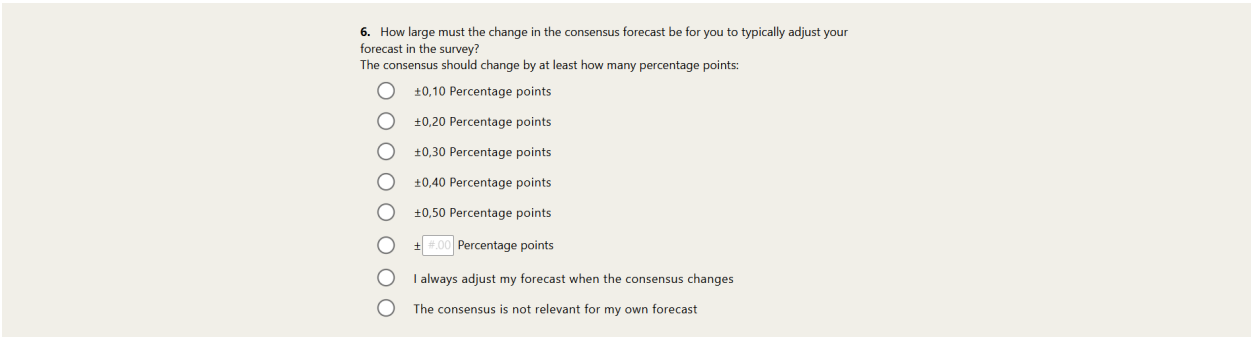
Notes: This figure shows screenshots displaying elements of our survey module.

Figure C.2: Inaction thresholds

(a) Threshold in the outlook change



(b) Threshold in the consensus change

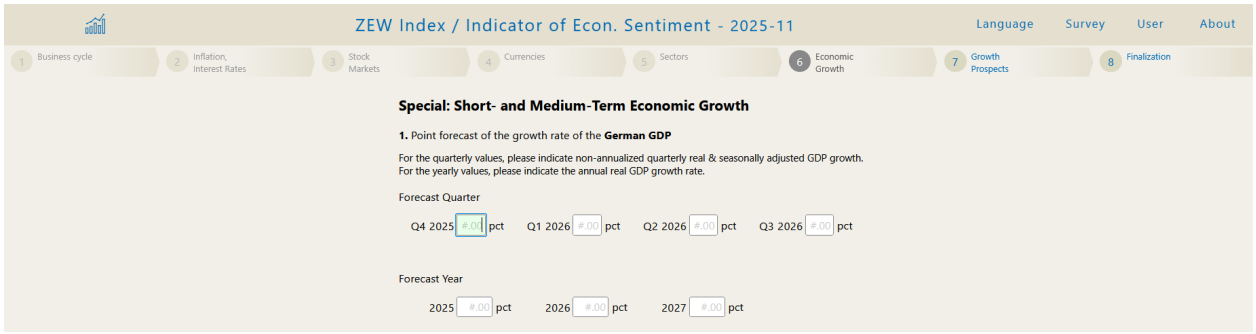


Notes: This figure shows screenshots displaying elements of our survey module.

C.2 RCT module (November 2025)

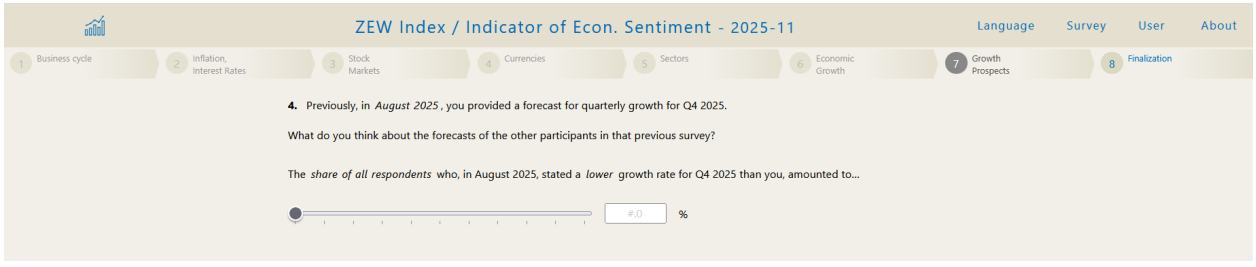
Figure C.3: Prior questions

(a) Real GDP growth



(b) Consensus beliefs

(c) Cumulative distribution function beliefs

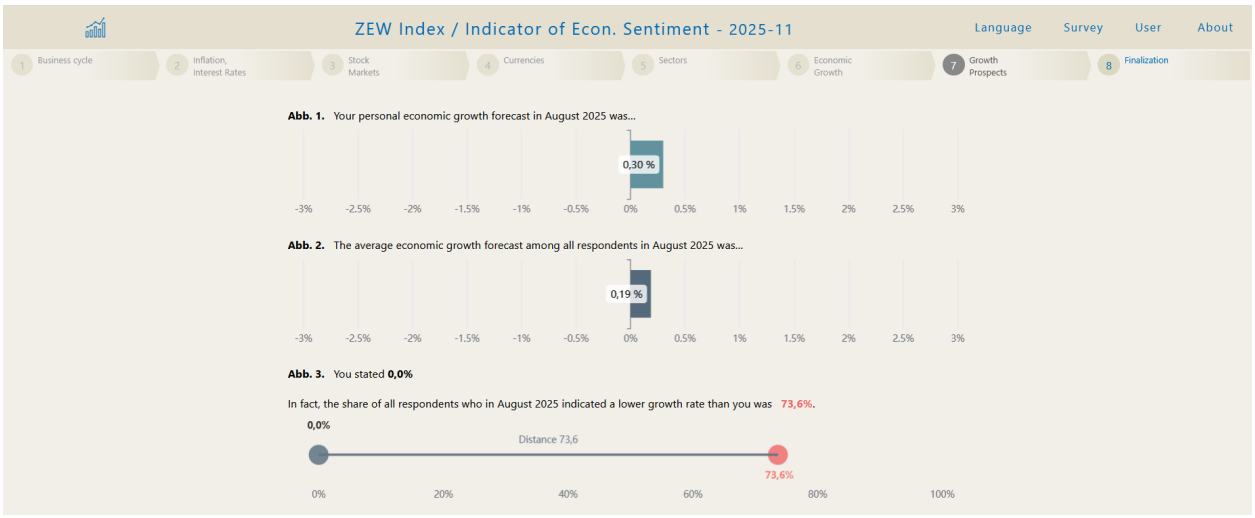


Notes: This figure shows screenshots displaying elements of our survey module.

Figure C.4: Information intervention

(a) Control group

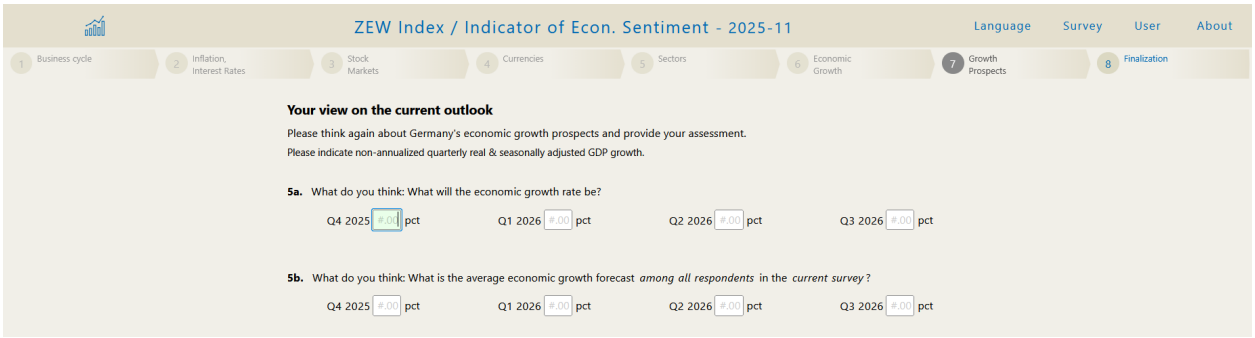
(b) Treatment group



Notes: This figure shows screenshots displaying elements of our survey module.

Figure C.5: Posterior questions

(a) Real GDP growth and consensus beliefs



(b) Additional questions

Notes: This figure shows screenshots displaying elements of our survey module.