Bank of England

Anchors aweigh? The effect of communicating forecast uncertainty

Staff Working Paper No. 1,196
July 2026

Michael McMahon, Matthew Naylor, Ryan Rholes and Peter Rickards

Bank of England

Staff Working Paper No. 1,196

Anchors aweigh? The effect of communicating forecast uncertainty

Michael McMahon,[1] Matthew Naylor,[2] Ryan Rholes[3] and Peter Rickards[4]

Abstract

We examine how central banks can effectively communicate forecast uncertainty in a two-part experimental study. Part I tests how different visual media – fan charts, dot plots, box-and-whisker plots, speedometers, and ranges – communicate uncertainty to both the general public and expert audiences. We find that fan charts are well understood and perform best at jointly conveying both expectations and uncertainty. Part II implements a novel dynamic information experiment with 1,600 UK participants across four stages, examining the effects of uncertainty communication on expectations and uncertainty perceptions over time. We find that while point forecasts anchor expectations marginally more than fan charts initially, forecast errors significantly de-anchor expectations, particularly for 'unlucky' errors that move inflation away from target. Critically, fan charts materially mitigate this de-anchoring, acting as an 'insurance policy' that helps protect central bank reputation. We also document that the public consistently underestimates the degree of uncertainty, and that communicating uncertainty via fan charts helps the public learn more realistic uncertainty perceptions. Our findings have important implications for central bank communication strategies.

Key words: Central bank communication, forecast uncertainty, fan charts, expectations, anchoring.

JEL classification: C91, D83, E52, E58.

---

(1) University of Oxford, CEPR. Email: michael.mcmahon@economics.ox.ac.uk
(2) Bank of England, University of Oxford. Email: matthew.naylor@bankofengland.co.uk
(3) University of Mississippi. Email: rarholes@olemiss.edu
(4) Reserve Bank of Australia, University of Oxford. Email: peter.rickards@economics.ox.ac.uk

The views expressed in this paper are those of the authors, and not necessarily those of the Bank of England, Reserve Bank of Australia, or their respective committees. We thank Ambrogio Cesa-Bianchi, Petra Geraats, Chris Giles, Anastasios Karantounias, Soumaya Keynes, Simon Lloyd, Clare Lombardelli, Huw Pill, and David Spiegelhalter for helpful comments on the paper. We also thank conference and seminar participants at the Bank of England Chief Economists' Workshop (2025), Banque de France, CEPR Central Bank Communication seminar series (December, 2025), and Swiss National Bank. The views expressed in this paper are ours, and not necessarily those of the Bank of England, Reserve Bank of Australia, or their respective committees. All errors are our own.

The Bank's working paper series can be found at www.bankofengland.co.uk/working-paper/staff-working-papers

Bank of England, Threadneedle Street, London, EC2R 8AH
Email: enquiries@bankofengland.co.uk

ISSN 1749-9135 (on-line)

1 Introduction

Alan Greenspan famously said "Uncertainty is not merely a pervasive feature of the monetary policy landscape, it is the defining characteristic of that landscape." Reflecting this, there is a rich theoretical literature on optimal monetary policy under uncertainty (e.g., Brainard 1967; Söderström 2002; Hansen and Sargent 2001; Levin et al. 2003), a growing empirical literature on how uncertainty influences monetary policy in practice (see e.g., Cieslak and McMahon 2024; Bauer et al. 2023; Fadda et al. 2023), and a literature on how uncertainty affects public expectations and behaviour (Coibion et al., 2024; Kostyshyna and Petersen, 2024; Fischer et al., 2025). However, surprisingly little is known about the effects of *communicating* uncertainty itself.

We conceptualise uncertainty communication as involving a dynamic trade-off. On one hand, communicating uncertainty may weaken the immediate anchoring of expectations by raising the salience of forecast error risk. On the other hand, it may enhance audience's understanding of the presence of uncertainty and protect credibility over time by preparing audiences for inevitable mistakes. While recent evidence documents the immediate anchoring cost (Rholes and Petersen, 2021; Petersen and Rholes, 2022; Kostyshyna and Petersen, 2023), direct evidence on the dynamic benefits remain scarce. This question has gained urgency following the Bernanke Review of the Bank of England's forecasting processes, which recommended discontinuing fan charts while emphasising the continued importance of communicating forecast uncertainty (Bernanke, 2024). In this paper, we address this gap by asking: how should central banks communicate uncertainty, and what are the effects of doing so?

We answer this question using a two-part experimental design. In Part I, we conduct a survey experiment with experts and UK households to measure how well several alternative methods – including point forecasts, fan charts, ranges, box & whisker plots, dot plots, and speedometers – convey central expectations and uncertainty. Using both subjective and objective measures, we find that fan charts perform best overall; combining high perceived usability with superior accuracy in conveying distributional information across audiences.[1] In Part II, informed by this evidence, we conduct a dynamic, incentivised information experiment with 1,600 participants drawn from a representative UK sample. Participants form multi-year inflation expectations and update them sequentially after observing: a central bank forecast (randomised to either a point forecast or fan chart); a forecast error (randomised by the size and direction of the shock); and an updated central bank forecast (of the same format). When we show them

[1] 'Conditional' uncertainty communication, such as scenario analysis, is not explored in this version of the paper.

fan charts, we vary between showing them 70% and 90% of the inflation distribution. We elicit both participants' central beliefs and uncertainty ranges, offering financial incentives for accurate responses, following McMahon and Rholes (2024). The design effectively stacks three Bayesian updating problems back-to-back, allowing us to isolate the impact how uncertainty communication affects both initial belief formation and subsequent updating after credibility shocks, as well as the impact of the credibility shock itself.

Consistent with prior evidence, we find that point forecasts initially anchor expectations marginally more than distributional forecasts (Rholes and Petersen, 2021; Petersen and Rholes, 2022; Kostyshyna and Petersen, 2023). We contribute to the literature with two novel findings. Our first contribution is to show that this anchoring advantage is fragile. Expectations de-anchor sharply following forecast errors, particularly following large errors that move inflation away from its long-run trend. Fan charts materially attenuate this response, acting as an insurance mechanism that could help to preserve forecaster credibility after misses. Revised central bank forecasts partially re-anchor expectations, though to a lower degree than premiss. This suggests there is a persistent reputational cost of forecast misses that fan charts could help to ameliorate.[2]

Our second contribution is that we show that the public consistently underestimates the degree of uncertainty in inflation forecasts (exhibiting substantial overconfidence in their point predictions), but communicating uncertainty through fan charts helps the public "learn" more realistic perceptions of uncertainty, bringing their subjective distributions closer to reasonable benchmarks. This learning effect is persistent, and particularly pronounced when we show them wider fan charts (e.g. 90%) that incorporate almost the entire distribution.

The policy implication is that prioritising short-run anchoring understates the value of uncertainty communication itself, which we argue provides insurance against the risk of expectations de-anchoring in repeated forecasting environments where forecast errors are inevitable. Moreover, helping the public develop more realistic uncertainty perceptions may improve the quality of their economic decision-making.

Our paper relates to a growing experimental and survey literature on how the communication of forecast uncertainty shapes beliefs and expectations. Within macroeconomics, experimental studies (Rholes and Petersen, 2021; Petersen and Rholes, 2022), together with complementary survey evidence from Kostyshyna and Petersen (2023), show that point forecasts generate stronger initial anchoring than distributional forecasts. Across other fields, Spiegelhalter

[2]These findings relate to evidence from a Statistics literature suggesting that communicating uncertainty can promote trust in institutions (Spiegelhalter, 2024; van der Bles et al., 2020).

(2024); van der Bles et al. (2020) study how communicating uncertainty affects public trust. Our contribution is to study uncertainty communication in a controlled Bayesian information environment that allows us to isolate and identify the dynamic trade-offs involved in whether and how to communicate uncertainty. We confirm the previous finding that communicating uncertainty reduces initial anchoring, and extend this literature by demonstrating that distributional forecasts also attenuate the de-anchoring that follows forecast misses and that distributional forecasts lead to substantially more realistic uncertainty beliefs. Finally, our results contribute to the broader literature on central bank communication and expectation management (Blinder et al., 2008; Campbell et al., 2019), by clarifying how alternative uncertainty communication strategies affect the anchoring of inflation expectations.

The remainder of this paper proceeds by analysing the two parts of our experiment. Section 2 presents Part I of our study, examining how different media communicate uncertainty. Section 3 describes Part II's dynamic information experiment and presents results on expectations and uncertainty perceptions. Section 4 concludes with policy implications.

2 Part I: What Medium of Uncertainty Communication Works?

2.1 Experimental Design

We recruited two distinct samples of participants for Part I. Our first sample consists of 300 individuals representative of the UK general public, recruited via the Prolific online survey platform. Our second sample comprises 75 experts, recruited through personal outreach by the authors, including PhD economics students, financial market participants, financial journalists, central bank staff, and individuals working for international organizations.

Each participant completed five evaluation rounds (henceforth, 'runs'). In each run, we showed participants a single forecast for inflation, consisting of a communication format paired with a hypothetical economic environment. To correspond with the five runs, we generated five separate 'worlds' of inflation data to ensure the data paths themselves were not driving preferences for, or understanding of, different uncertainty representations. To construct the data for these five worlds, we randomly selected starting values from a uniform distribution between 1 and 5 to ensure values relatively close to reality. Each series then evolves following a random-walk process, where every new observation is created by adding a random shock to the previous value. These shocks are drawn from mean-zero normal distributions. The historical portion of each series contains eight observations. From the end of this history, three additional periods

are simulated to represent forecasts, allowing the series to continue evolving under smaller random shocks. This process produces five alternative histories and corresponding forecasts.[3]

To mitigate learning or cross-format contamination, each world was randomly paired with one communication format, with the pairing randomized across participants. Moreover, participants viewed the five format–world pairs in random order.

The communication formats we tested, depicted in Figure 1, include: (a) Point forecast only; (b) Fan chart (with 70% and 90% confidence bands); (c) Range (99% confidence interval); (d) Box-and-whisker plot (showing interquartile range); (e) Dot plots, similar to the Federal Reserve's Summary of Economic Projections; and (f) Speedometer. Our choice of unconditional 'media' reflects either communication practices used currently by central banks around the world, common methods of depicting uncertainty, or, in the case of the Speedometer, a novel suggested method.

For each format-world pair, we ask participants a variety of questions designed to assess their understanding of the information conveyed in each format. Questions, specified verbatim in Appendix A, capture both 'subjective' (self-reported) and 'objective' measures of participants' understanding, including of the central forecast as well as of the uncertainty conveyed. Participants are financially incentivised to seek to answer the 'objective' questions accurately with a performance-based bonus payment. Participants received a fixed payment of £3.00 for completing the survey, with the possibility of an additional £2.10 based on performance. The estimated length of the survey was 20 minutes, and the mean time taken to complete it was 18.5 mins. We show in Table B.1 that there are no significant differences in time taken across treatments.

2.2 Descriptive evidence: Subjective understanding and uncertainty

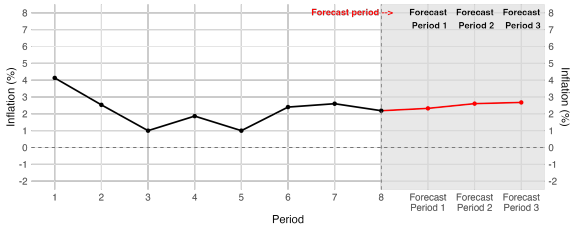
We start by comparing responses across samples and treatments descriptively, before testing these observations statistically in Section 2.4.

Turning first to subjective understanding, Figure 2 presents the mean and standard deviation of responses across communication formats for the public (circles with solid error bars) and experts (triangles with dashed error bars) to Question 1; capturing participants' perceived ability to understand different chart types on a seven-point scale. We see that experts report markedly higher understanding than the general public across nearly all communication for-

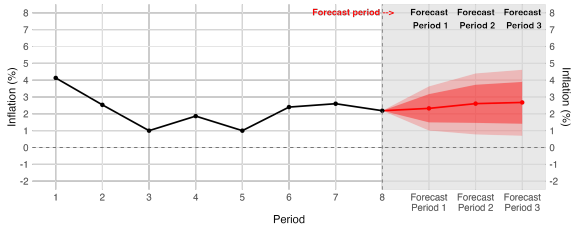
[3]Summary statistics for the inflation series in each world are provided in Table A.1.

Figure 1: Communication formats

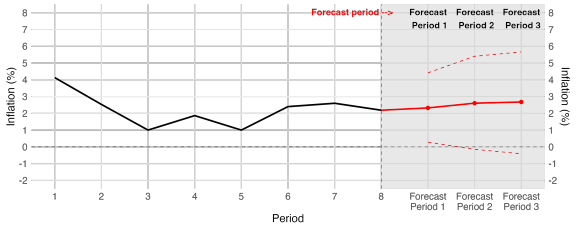
(a) Point Forecast



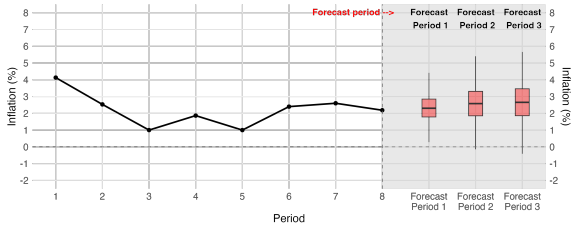
(b) Fan Chart



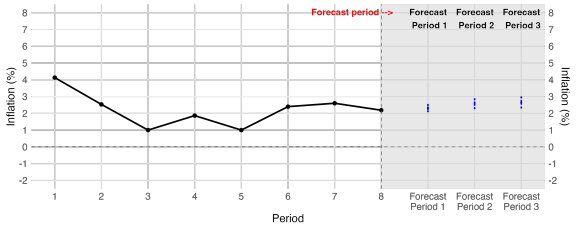
(c) Range



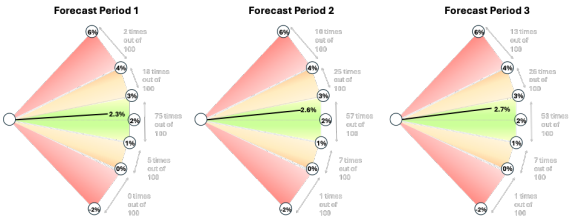
(d) Box & Whisker



(e) Dot Plot



(f) Speedometer

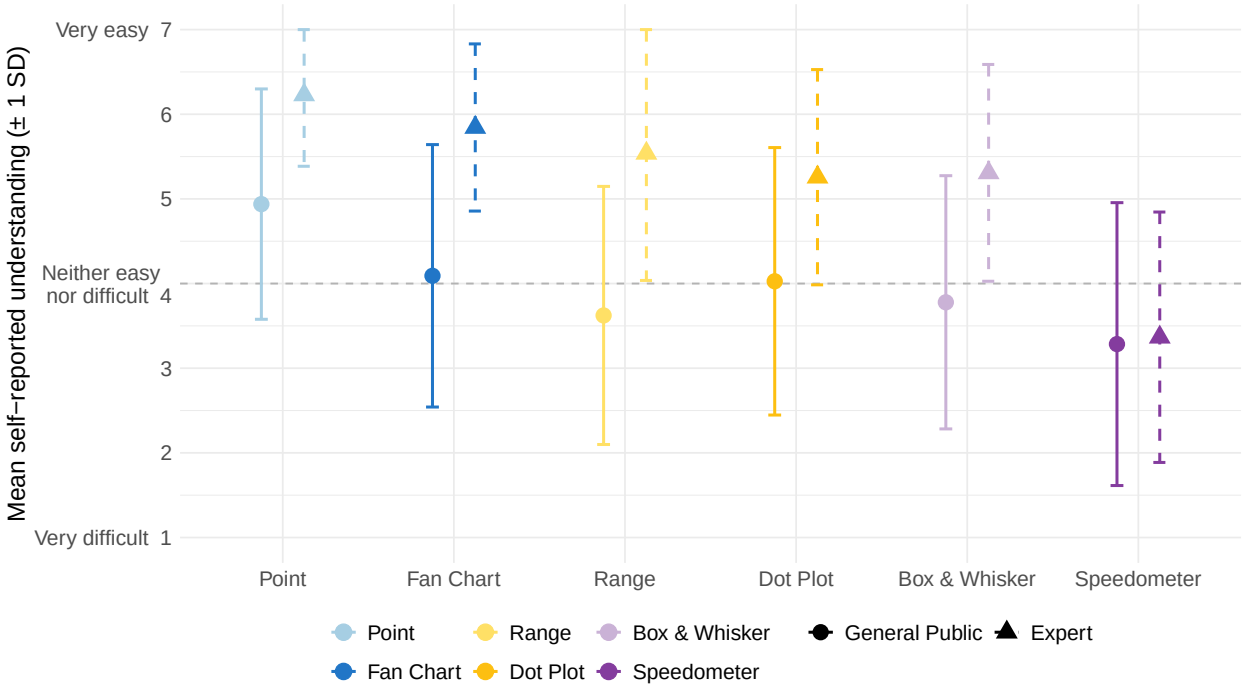


Notes: This figure presents screenshots of each of the six different 'unconditional' media that participants could be randomly assigned to in the Part I experiment. Each of the six media in this figure depict the same 'world'. For each media presented in Figures 1a-1e, the following text is provided to participants: "*The chart below shows the inflation rate in a hypothetical economy between periods 1 and 8 (black line) along with a forecast for inflation in 'Forecast Periods' 1, 2, and 3...*" In addition, for the Point forecast (Figure 1a), participants are told: "*The red line shows the forecasted inflation outcomes with the highest predicted chance of occurring.*" For the Fan chart (Figure 1b), participants are told: "*... in the form of a 'fan chart'. The central red line shows the forecasted inflation outcomes with the highest predicted chance of occurring, while the shaded areas reflect the likelihood of other possible outcomes based on past forecast mistakes. The darkest shaded area covers the set of outcomes with a 70% chance of occurring. The lighter area extends this to 90%.*" For the Range (Figure 1c), participants are told: "*... The central red line shows the forecasted inflation outcomes with the highest predicted chance of occurring, while the area between the dotted lines reflects the set out of outcomes that have a 99% chance based on past forecast mistakes.*" For the Box & Whisker (Figure 1d), participants are told: "*... in the form of a 'box and whisker chart'. The central line within the box shows the forecasted inflation outcomes with the highest predicted chance of occurring, while the box and whiskers reflect the likelihood of other possible outcomes based on past forecast mistakes. The box covers the set of outcomes with 50% chance of occurring. The whiskers extend this to 99%.*" For the Dot Plots (Figure 1e), participants are told: "*... by nine different expert forecasters in the form of a 'dot plot'. Each dot shows the forecasted inflation outcomes that each forecaster predicts has the highest chance of occurring.*" For the Speedometer (Figure 1f), participants are instead told: "*The image below shows a forecast for inflation in a hypothetical economy in 'Forecast Period' 1, 2, and 3 in the form of a 'speedometer'. The black line (or, 'dial') in each speedometer shows the forecasted inflation outcome with the highest predicted chance of occurring, while the grey text reflects the probability of other possible outcomes occurring within certain ranges (distinguished by different coloured shading) based on past forecast mistakes.*"

mats, but the relative ranking of the formats is similar between the two samples. Both the general public (average response of 4.9) and experts (average of 6.2) rate Point Forecasts (light blue) as easiest of all communication formats to understand. This is perhaps unsurprising, given their relative simplicity. Fan charts (dark blue) also perform well, scoring second highest for both samples (averaging 4.1 across the public and 5.8 across the experts). In contrast, the speedometer (dark purple) is notably unpopular with both sample populations (averaging 3.3 and 3.4, respectively). This is a notably lower score amongst experts, relative the other communication formats (where the second lowest score is 5.3 for Dot Plots).

Comparing disagreement across the samples and communication formats, we see considerably less disagreement across experts responses to Point Forecasts and Fan Charts (standard deviation of 0.8 and 1.0), relative to both responses across the general public (minimum standard deviation of 1.4, for Point Forecasts) and to other communication formats (minimum standard deviation of 1.3, for Dot Plots and Box & Whisker).

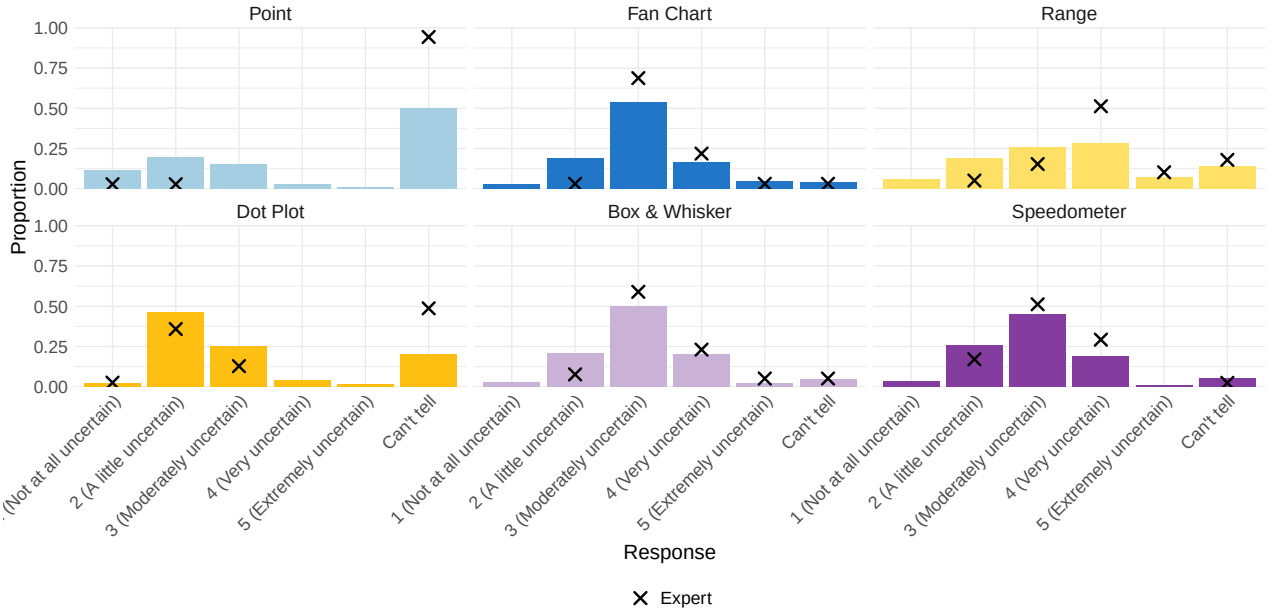
Figure 2: Subjective understanding of information conveyed



Notes: This figure presents the mean (central point, circle for General Public, triangle for Experts) and standard deviation (error bars, solid for General Public, dashed for Experts) of responses to Q1. "*How easy or difficult do you find this graph to understand?*" across 'unconditional' media, with responses on a (1-7) scale, where 1 represents 'Very difficult', 4 represents 'Neither easy nor difficult', and 7 represents 'Very easy'. Light blue represent responses for the Point forecast, dark blue represent responses for the Fan Chart, yellow represent responses for the Range, orange represent responses for the Dot Plot, light purple represent responses for the Box & Whisker, dark purple represent responses for the Speedometer.

Self-reported understanding, however, tells us only that participants *believe* they can read a given chart. Since our objective is to communicate uncertainty, perhaps the more important subjective question is whether people actually perceive uncertainty in a given medium. When asked how much uncertainty each chart conveys (Q2), we observe striking differences across media (Figure 3).

Figure 3: Subjective perception of uncertainty conveyed



Notes: This figure presents the distribution of responses to Q2. "*How much uncertainty does this chart show?*" across 'unconditional' media, with histograms reflecting the proportion of responses across a 1-5 scale, where 1 represents 'Not at all uncertain', 2 represents 'A little uncertain', and 3 represents 'Moderately uncertain', 4 represents 'Very uncertain', 5 represents 'Extremely uncertain'. Participants also have the option to tick 'Can't tell'. The bars represent responses amongst the General Public sample, crosses represent responses amongst the Experts sample. Light blue bars represent responses for the Point forecast, dark blue bars represent responses for the Fan Chart, yellow bars represent responses for the Range, orange bars represent responses for the Dot Plot, light purple bars represent responses for the Box & Whisker, dark purple bars represents responses for the Speedometer.

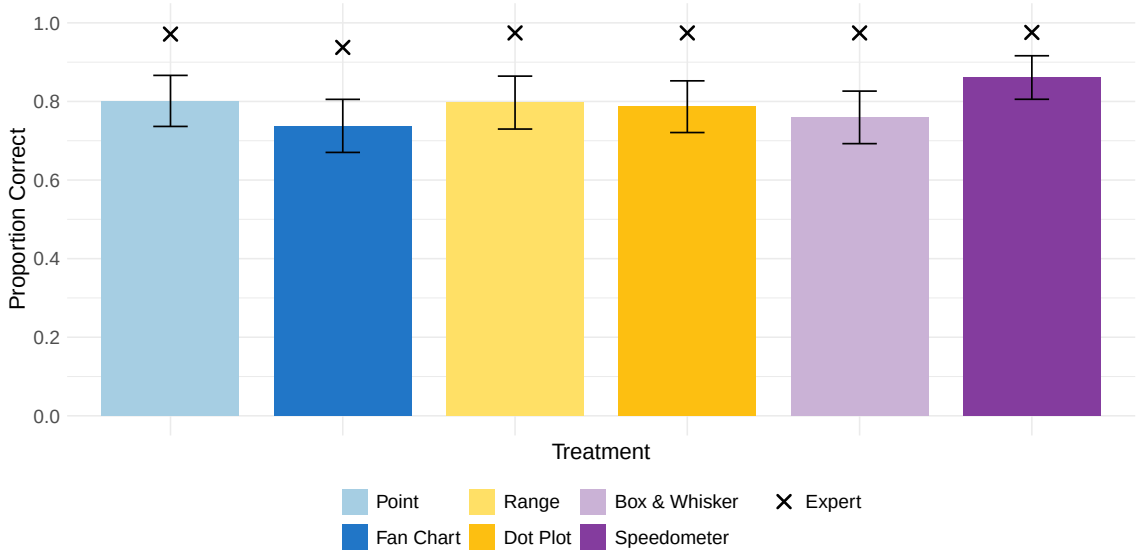
For point forecasts, 50% of the general public correctly indicate that they 'Can't tell' how much uncertainty is present. Disconcertingly, however, 50% perceive there to be little uncertainty despite point forecasts communicating no information about forecast uncertainty. Notably, experts are far more likely to correctly recognise the absence of uncertainty information in point forecasts, whereas the general public tends to mistake precision for certainty — a distinction that may help explain why fan charts provide an insurance benefit against credibility shocks when forecast errors occur. Dot plots generate similar confusion: despite representing disagreement among forecasters rather than uncertainty about outcomes, many respondents interpret them as conveying low uncertainty. This confusion extends, albeit to a lesser degree, to our expert panel. In contrast, fan charts, box-and-whisker plots, and speedometers all elicit responses centred around moderate uncertainty levels, with less pronounced differences

between the general public and experts than we observe for point forecasts and dot plots.

2.3 Descriptive evidence: Objective understanding

The preceding survey questions capture what participants *believe* they understand. However, several of our questions have objectively correct answers, allowing us to measure whether different media actually convey the intended information. We begin by testing whether each medium can successfully convey the central tendency of a distributional outlook. We treat this as a necessary condition for feasible communication media; if a medium cannot communicate the forecaster's view of the most likely outcome, conveying uncertainty around it is moot. On this front (Q3), the media all perform uniformly well (depicted in Figure 4). Approximately 80% of the general public and over 95% of experts correctly identify the most likely value across all media, with no statistically significant differences between treatments within each sample population. Given that all media clearing this hurdle, the primary question becomes which can best communicate additional distributional information.

Figure 4: Objective understanding of central forecast

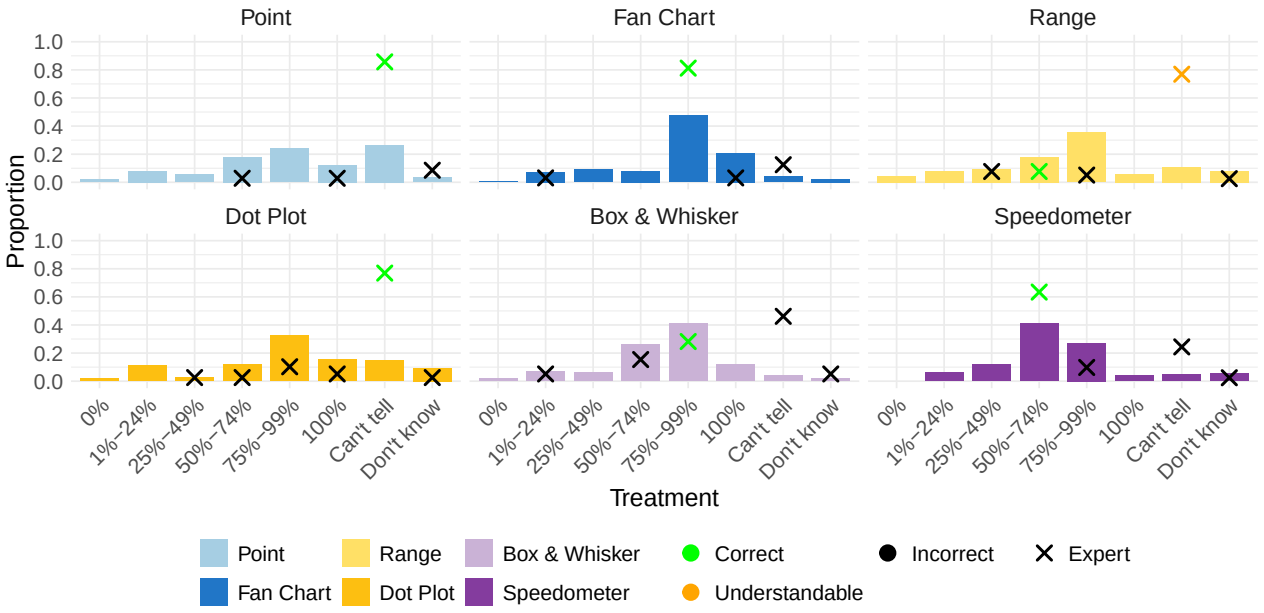


Notes: This figure presents the proportion of correct responses to Q3. "*What is the most likely value for the series in forecast period 1?*" across 'unconditional' media. Responses are rewarded as 'correct' within a ±0.3pp margin for error, with the exact most likely value in forecast period 1 varying across 'worlds'. The bars represent responses amongst the General Public sample with 5% confidence intervals represented by the whiskers. Crosses represent responses amongst the Experts sample. Light blue bars represent responses for the Point forecast, dark blue bars represent responses for the Fan Chart, yellow bars represent responses for the Range, orange bars represent responses for the Dot Plot, light purple bars represent responses for the Box & Whisker, dark purple bars represents responses for the Speedometer.

Substantial differences emerge in participants' ability to extract distributional information across media (Q6). When asked about the probability that inflation falls in a particular

range, Figure 5 shows that fan charts, box-and-whisker plots, and speedometers perform well, with large shares of both public and expert respondents answering correctly. Point forecasts and dot plots perform poorly: while nearly all experts correctly recognise that neither chart type provides enough information to infer probabilities, more than 75% of the general public spreads across the response scale assigning specific probabilities where none exist. Figures B.1 and B.2 depict results demonstrating similar qualitative takeaways based on Q4 and Q5, respectively, which also capture participants' understanding of the uncertainty conveyed by the unconditional media, eliciting views on the presence of uncertainty and the underlying distribution of possible outcomes, respectively.

Figure 5: Objective understanding of uncertainty conveyed (forecast range probabilities)



Notes: This figure presents the distribution of responses to Q6. "*How likely is it that inflation will be between X% and Y% in forecast period 3?*" across 'unconditional' media, across the set of possible responses: ['0%', '1%-24%', '25%-49%', '50%-74%', '75%-99%', '100%', 'Cannot tell - I need more information', 'Do not know']. The bars represent the proportion of responses in each bucket amongst the General Public. Crosses represent the proportion of responses responses in each bucket amongst the Experts sample. The 'correct' bucket is the same across 'worlds' for each media, indicated by the bucket containing the green cross for each media. The bucket in the Range medium containing the orange cross represents responses that are 'understandable': where that bucket is not strictly the correct answer, but it is also not strictly incorrect. Light blue bars represent responses for the Point forecast, dark blue bars represent responses for the Fan Chart, yellow bars represent responses for the Range, orange bars represent responses for the Dot Plot, light purple bars represent responses for the Box & Whisker, dark purple bars represents responses for the Speedometer.

2.4 Regression Analysis

The descriptive evidence above highlights clear visual patterns in how different communication formats perform across our subjective and objective measures. To formally assess whether these differences are statistically significant, and to quantify their magnitudes, we now turn

to a regression framework. Our within-subject design allows us to include participant fixed effects, exploiting within-person variation to estimate treatment effects with greater precision than cross-sectional comparisons alone would permit. The interaction with expert status lets us test whether the apparent differences between our two samples documented in the figures hold up formally, while controls for presentation order allow us to verify that sequencing does not meaningfully affect our estimates. Throughout, the omitted category is the point forecast, so coefficients capture the incremental effect of each alternative format relative to the simplest baseline.

To formally assess treatment effects, we estimate regressions of the form:

$$\mathrm{Y}_{i,r}^{Q} = \beta_1 \mathrm{Treatment}_{i,r} + \beta_2 \mathrm{Treatment}_{i,r} \times \mathrm{Expert}_i + \delta X_{i,r} + \gamma_i + \epsilon_{i,r} \tag{1}$$

where $\mathrm{Y}_{i,r}^{Q}$ is based on participant i's response to question $\mathrm{Q} \in \{\mathrm{Q1}, \mathrm{Q2}, \mathrm{Q3}, \mathrm{Q4}, \mathrm{Q5}, \mathrm{Q6}\}$, in run r. For Q1 and Q2, capturing 'subjective' understanding (on a 1-7 scale) and 'subjective' perceptions of uncertainty conveyed (on a 1-5 scale), respectively, $\mathrm{Y}_{i,r}^{Q}$ is a categorical variable taking discrete values on those respective scales. For Q3-Q6, capturing 'objective' understanding, $\mathrm{Y}_{i,r}^{Q}$ is a binary dummy variable taking the value 1 for 'correct' responses and the value 0 otherwise. $\mathrm{Treatment}_{i,r} \in \{\mathrm{Fan\ Chart}, \mathrm{Range}, \mathrm{Dot\ Plots}, \mathrm{Box\ \&\ Whisker}, \mathrm{Speedometer}\}$, with Point Forecasts as the benchmark omitted treatment. Expert_i is a binary dummy that takes the value 1 if participant i is in the expert sample, and 0 if i is in the general public sample. $X_{i,r}$ is a control for experimental variation regarding the order in which (i.e., in which run r) participant i observes a treatment. γ_i are participant fixed effects, and $\epsilon_{i,r}$ is the error term.[4] We cluster standard errors at the participant level.

Table 1 presents estimates from Equation 1, where the Point Forecasts are the benchmark medium (omitted treatment) relative to which marginal effects of each other media are estimated. For reference, we also report the 'intercept' (representing the average level of response amongst the general public for the Point Forecast) and the 'Expert' constant (representing the average level of response amongst the expert, relative to the general public, for the Point Forecast) from a version of Equation 1 without participant fixed effects (see Table B.2) – which absorb these variables in our baseline specification. Beginning with the general public, we see from Column (1) that the average response to Q1 amongst the general public treated

[4]We rely on a combination of our randomisation across treatments and our participant fixed effects to absorb person-specific characteristics that may drive systematic differences across treatments. We do not in addition include individual controls, and show in Table B.1 that there are no systematic differences across treatments in participants' demographic characteristics, economics experience, financial literacy, or survey experience.

with a point forecast is that participants believe they understand the information conveyed relatively well (5.2 out of 7). Relative to Point Forecasts, Fan Charts score 0.801 points lower on the 7-point scale, and each of the other media score even lower, with Speedometers scoring particularly badly (1.704 points lower than point forecasts). However, the average response to Q2 (Column (2)) amongst participants treated with the point forecast is 2.190 on the 1-5 scale – significantly below the middle option of 3 which reflects 'Moderate' uncertainty; indicating that participants perceive too little uncertainty from the Point Forecast. In contrast, Fan Charts increase perceived uncertainty by 0.533 points, indicating that they successfully convey the existence of uncertainty (Column (2)). As do the Range, Box & Whisker, and Speedometer treatments.

Columns (3)-(6) present estimates of 'objective' understanding. Critically, no media show significant differences in their ability to effectively communicate ('anchor') point estimates (Column (3)), relative to the Point Forecast itself which scores well in absolute terms (on average, 80.2% of participants respond correctly). However, as shown in Columns (5) and (6), each of the other media improve communication of distributional information relative to the Point Forecast. This is most apparent across both the Speedometer and, particularly, the Fan Chart, increasing the proportion of people understanding the underlying distribution of possible outcomes by 11.8 and 14.4 percentage points (Column (5)), respectively, and increasing the proportion of people able to accurately extract forecast range probabilities by 41.9 and 50 percentage points (Column (6)), respectively.

Table 1: Part I: Baseline regression

	Dependent variable: Responses to Part I questions					
	Subjective		Anchoring	Uncertainty		
	Q1.Liked? (1)	Q2.How uncertain? (2)	Q3.Point estimate (3)	Q4.Presence (4)	Q5.Distribution (5)	Q6.Probabilities (6)
Intercept (from no FE model)	**5.181*****	**2.190*****	**0.802*****	**0.699*****	**0.536*****	0.002
β_1 **('Unconditional' Treatments)**						
Fan	**−0.801*****	**0.533*****	−0.010	−0.017	**0.144*****	**0.500*****
	(0.131)	(0.131)	(0.040)	(0.051)	(0.042)	(0.047)
Range	**−1.199*****	**0.516*****	0.006	−0.021	0.059	**0.167*****
	(0.133)	(0.142)	(0.041)	(0.051)	(0.044)	(0.045)
Dot Plots	**−0.977*****	0.018	−0.022	**−0.087***	**0.067***	0.003
	(0.132)	(0.132)	(0.040)	(0.049)	(0.040)	(0.037)
Box & Whisker	**−1.253*****	**0.454*****	−0.064	0.044	0.058	**0.416*****
	(0.131)	(0.130)	(0.041)	(0.046)	(0.041)	(0.046)
Speedometer	**−1.704*****	**0.340****	0.057	0.047	**0.118*****	**0.419*****
	(0.143)	(0.136)	(0.038)	(0.050)	(0.042)	(0.048)
Expert (from no FE model)	**1.312*****	**−0.741***	**0.170*****	**0.121***	−0.082	0.0001
β_2 **('Unconditional' Treatments)**						
Fan*Expert	**0.764*****	**1.977****	−0.027	0.094	**0.270*****	**0.297*****
	(0.284)	(0.851)	(0.069)	(0.087)	(0.101)	(0.100)
Range*Expert	**0.652****	**2.482*****	−0.025	−0.051	0.077	−0.111
	(0.265)	(0.835)	(0.060)	(0.088)	(0.098)	(0.082)
Dot Plots*Expert	0.189	1.181	0.012	**0.146***	0.088	−0.007
	(0.273)	(0.849)	(0.063)	(0.081)	(0.099)	(0.071)
Box & Whisker*Expert	0.465	**1.940****	0.063	−0.075	**0.361*****	−0.129
	(0.293)	(0.848)	(0.065)	(0.093)	(0.099)	(0.096)
Speedometer*Expert	**−0.829*****	**1.937****	−0.057	0.005	**0.244****	**0.233****
	(0.314)	(0.845)	(0.059)	(0.082)	(0.109)	(0.097)
Participant FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls: Experimental variation	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,132	951	1,132	1,132	1,132	1,132
Adjusted R^2	0.625	0.323	0.414	0.322	0.576	0.277

*p<0.1; **p<0.05; ***p<0.01. Cluster-robust standard errors in parentheses, clustered at participant level.

Notes: This table reports results from Equation 1. Columns (1)–(6) correspond to Questions Q1–Q6, respectively. In columns (1) and (2), the dependent variable is the participant's categorical response to Q1 (self-reported understanding, 1–7 scale) and Q2 (perceived uncertainty, 1–5 scale). In columns (3)–(6), the dependent variable is a binary indicator for a correct response: identifying the most likely forecast value within ±0.3pp (Q3), recognising the trend in uncertainty across forecast horizons (Q4), recognising that outcomes nearer the forecast are more likely (Q5), and identifying the correct probability range (Q6). Full question wording is provided in Appendix A. The rows labelled Intercept' and Expert' report the constant and expert dummy coefficient from a version of the regression without participant fixed effects, to provide baseline levels and sample differences; bold indicates significance at the 10% level.

Experts score better across the board. Relative to the general public, they report higher subjective understanding of each of the media (summing the 1.312 points out of 7 higher 'Expert' benchmark with each marginal coefficient), and are less prone to under-interpreting the degree of uncertainty across all treatments with the exception of the Point Forecast. Experts also score better on objective measures across both 'anchoring' (by 17 percentage points on average, Column (3)) and understanding the presence of uncertainty (by 12.1 percentage points on average in Column (4). Understandably, there is no significant difference in their understanding of the underlying distribution of outcomes (Column (5)) or ability to extract forecast

range probabilities (Column (6)) for the Point Forecast, which provides no such information. However, there is a significant improvement in their ability to understand this information as communicated by the Fan Chart (by 27 and 30 percentage points, respectively), the Box & Whisker (by 36.1 percentage points for the former), and the Speedometer (by 24.4 and 23.3 percentage points, respectively).[5]

We summarise the regression results from Table 1 in Table 2 for the general public and in Table B.3 for the expert sample, scaling the estimated coefficients from the 'subjective' measures on a (-1 to 1) scale for interpretability, and developing composite 'accuracy scores' for the estimated coefficients on the 'objective' measures to weight together performance on anchoring and uncertainty dimensions. Positive values in Column (1) indicate that media are perceived to be relatively well understood (or, 'liked'). Values closer to 0 in Column (2) indicate a relatively more neutral perception of the degree of uncertainty, while negative values reflect an interpretation of relatively little uncertainty. Columns (3)-(6) report the values of the Intercept (no FE) + β_1 coefficients from Table 1, while Columns (7)-(9) report weighted averages of these coefficients (composite 'accuracy scores'), across different weightings $\alpha \in \{0.25, 0.5, 0.75\}$ assigned to coefficients for 'anchoring' relative to coefficients (averaged across) the uncertainty questions. We highlight in blue the top three performing media across the subjective measures and the accuracy scores. We see that fan charts emerge as the most consistent top performer for the general public (Table 2). This reflects their ability to successfully convey both expectations and uncertainty while simultaneously being well-liked and understood. In contrast, Point Forecasts are well-liked but do not convey uncertainty, while Speedometers convey expectations and uncertainty well, but are not well-liked. We draw similar conclusions for the expert sample in Table B.3.

[5]Table B.2 presents results without participant fixed effects, from which we report the Intercept and Expert constants in Table 1, for reference.

Table 2: Part I: Composite 'scores' across regression results (General Public)

Public	**Subjective** Scaled: $[-1, 1]$		**Objective** Unscaled: β = 'proportion correct'						
			Anchoring	**Uncertainty**			**Accuracy 'Score'**		
	Q1.Liked?	Q2.How uncertain?	Q3.Point Estimate	Q4.Presence	Q5.Distribution	Q6.Probabilities	$\alpha = 0.25$	$\alpha = 0.5$	$\alpha = 0.75$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Point	**0.39**	-0.41	0.802	0.699	0.536	0.002	0.51	0.61	0.70
Fan	**0.13**	**-0.11**	0.792	0.682	0.680	0.502	**0.66**	**0.71**	**0.75**
Range	-0.01	**-0.11**	0.808	0.678	0.595	0.169	0.56	0.64	**0.73**
Dot Plots	**0.07**	-0.39	0.780	0.612	0.603	0.005	0.50	0.59	0.69
Box & Whisker	-0.02	**-0.14**	0.738	0.743	0.594	0.418	**0.62**	**0.66**	0.70
Speedometer	-0.17	-0.20	0.859	0.746	0.654	0.421	**0.67**	**0.73**	**0.80**

Scaling for 'Subjective' coefficients: $\left[\frac{\text{Intercept}+\beta_1-1}{\gamma-1} - 0.5\right] \times 2$ where $\gamma = [7, 5]$ for [Liked?, How uncertain?]

Accuracy 'Score' $= \alpha \times \beta^{\text{Anchoring}} + (1-\alpha) \times \bar{\beta}^{\text{Uncertainty}}$

Notes: This table reports summarised versions of the estimated coefficients reported in Table 1 for the general public based on Equation 1. Columns (1) and (2) report Intercept (no FE) $+ \beta_1$ values from Columns (1) and (2), respectively, of Table 1, scaled to a -1 to 1 scale. Columns (3)-(6) report Intercept (no FE) $+ \beta_1$ values from Columns (3)-(6), respectively, of Table 1. Columns (7)-(9) report weighted averages of the coefficients across Columns (3)-(6) (referred to as composite 'accuracy scores'), using different weightings $\alpha \in \{0.25, 0.5, 0.75\}$ assigned to coefficients for 'anchoring' (Column (3)) relative to coefficients (averaged across) the uncertainty questions (Columns (4)-(6)). Blue highlighted coefficients reflect the top three performing media for each of the 'subjective' measures and the 'accuracy scores'.

Taken together, several clear patterns emerge from Part I. First, Point Forecasts are most "liked" but misleadingly suggest less uncertainty than exists, while Fan Charts are next most liked and mislead the least. Speedometers are liked the least. Second, while all media successfully communicate point expectations, Fan Charts and Speedometers are best at jointly conveying expectations and uncertainty. Third, while experts understand all media better than the public, relative rankings are similar across the two samples, with Fan Charts consistently performing best across the board for both.

Given these results, we focus our dynamic experiment in Part II on comparing Point Forecasts and Fan Charts, as these represent the central bank communication methods of greatest policy relevance.

3 Part II: Dynamic Information Experiment

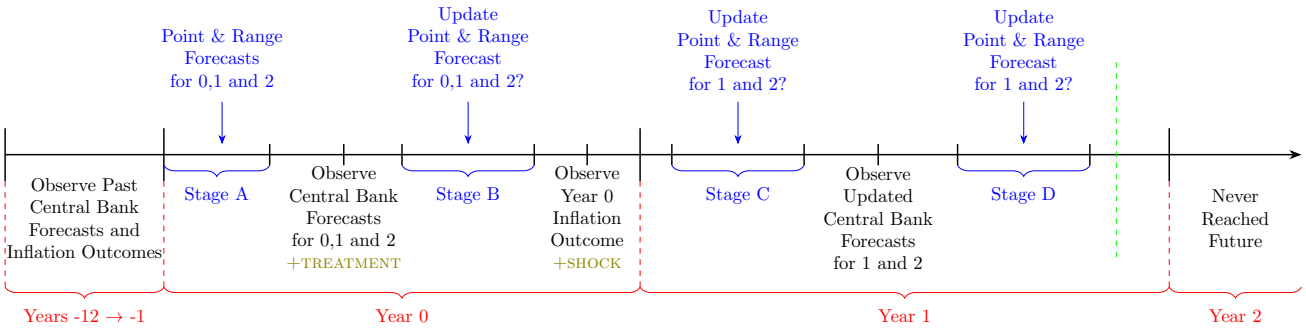
3.1 Experimental Design and Timeline

Part II implements a novel dynamic information experiment designed to trace how uncertainty communication affects expectations and uncertainty perceptions over time as new information is revealed. Our survey consists of 1,600 observations made up of a UK-representative sample of the general public via Prolific.

Figure 6 illustrates the experimental timeline. Time periods are indexed from -12 to $+2$,

where negative periods represent historical data and periods 0, 1, and 2 are forecast horizons. Participants never actually reach period 2 in the experiment. The experiment proceeds through four stages:

Figure 6: Experimental Timeline for Part II



Notes: Figure shows the four-stage experimental design with information sets expanding from Stage A through Stage D.

Stage A (Priors): Participants observe a 12-year history of annual inflation data $\{\Pi_{-12}, \ldots, \Pi_{-1}\}$ along with associated central bank year-ahead forecasts $\{\pi_{cb,-12}, \ldots, \pi_{cb,-1}\}$. Based on this information set Ω^A, participants provide their expected inflation for years 0, 1, and 2, denoted $\{\pi^A_{i,0}, \pi^A_{i,1}, \pi^A_{i,2}\}$. They also provide confidence ranges $\{r^A_{i,0}, r^A_{i,1}, r^A_{i,2}\}$ around these forecasts.

Stage B (Central Bank Forecast): The information set expands to $\Omega^B = \Omega^A \cup \{\pi^0_{cb,0}, \pi^0_{cb,1}, \pi^0_{cb,2}\}$, where participants observe the central bank's latest forecast. Crucially, participants are randomly assigned to one of three treatment groups:

- **Point forecast:** Receives only point forecasts $\pi^0_{cb,t}$ for $t \in \{0, 1, 2\}$
- **Fan (Narrow):** Receives point forecasts plus narrow fan charts representing 70% confidence intervals, with the interval communicated to participants
- **Fan (Wide):** Receives point forecasts plus wide fan charts representing 90% confidence intervals, with the interval communicated to participants

Participants then update to posterior forecasts $\{\pi^B_{i,0}, \pi^B_{i,1}, \pi^B_{i,2}\}$ and ranges $\{r^B_{i,0}, r^B_{i,1}, r^B_{i,2}\}$.

Stage C (Forecast Error): Year 0 inflation is revealed: $\Omega^C = \Omega^B \cup \{\Pi_0\}$. The realized value Π_0 generates a forecast error relative to both the participant's forecast and the central bank's forecast. Critically, the realized inflation is randomly assigned across participants, varying in both size and direction. This generates heterogeneous "shocks" leading to: small positive,

small negative, large negative, or large positive forecast errors. Each of these errors take the inflation series either toward or away from the CB's Year 2 forecast, depending on the inflation series (history and forecast) that participants are randomly assigned to. We refer to errors that take the series *toward* the CB's Year 2 forecast as "lucky" errors, while those that take the series away are "unlucky".

Participants update their forecasts for Years 1 and 2: $\{\pi^C_{i,1}, \pi^C_{i,2}\}$ and $\{r^C_{i,1}, r^C_{i,2}\}$.

Stage D (Updated Central Bank Forecast): Finally, $\Omega^D = \Omega^C \cup \{\pi^1_{cb,1}, \pi^1_{cb,2}\}$, as participants observe the central bank's updated forecast (in the same format as Stage B). They make final updates: $\{\pi^D_{i,1}, \pi^D_{i,2}\}$ and $\{r^D_{i,1}, r^D_{i,2}\}$.

Throughout the experiment, participants are incentivized for accuracy through financial payments based on both the proximity of their point forecasts to realized values and whether the realized values fall within their stated confidence ranges (with smaller ranges receiving higher rewards, preventing participants from stating infinitely wide ranges).

Each participant completes two 'runs' of the experiment, consisting of two different economic 'worlds' (details below). Participants are randomly assigned a Point or a Fan in the first run. They are then assigned to the other treatment in the second run. Moreover, they are randomly assigned an error type in the first run, and subsequently receive an error of the opposite direction and magnitude. Worlds are assigned randomly across treatments and error types and in random orders across runs.

3.2 Data-Generating Process

The underlying environment is a standard three-equation New Keynesian model:

$$\text{Phillips curve:} \quad \hat{\pi}_t = \beta \mathbb{E}_t \hat{\pi}_{t+1} + \kappa \hat{y}_t \tag{2}$$

$$\text{IS curve:} \quad \hat{y}_t = \mathbb{E}_t \hat{y}_{t+1} - \sigma(\hat{r}_t - \hat{\bar{r}}_t) \tag{3}$$

$$\text{Monetary policy rule:} \quad i_t = \chi_\pi \pi_t + \chi_y y_t + u_t \tag{4}$$

The historical inflation data shown to participants represents simulations of this model with specific shock realizations. Critically, all central bank forecasts are optimal forecasts conditional on the model and available information—we do not introduce bias or strategic behavior. The fan charts shown to participants in the Fan treatment groups are derived from stochas-

tic simulations of the model and represent genuine confidence intervals (i.e., a 90% fan truly contains 90% of the simulated outcomes).

3.3 Results: Expectations and Anchoring

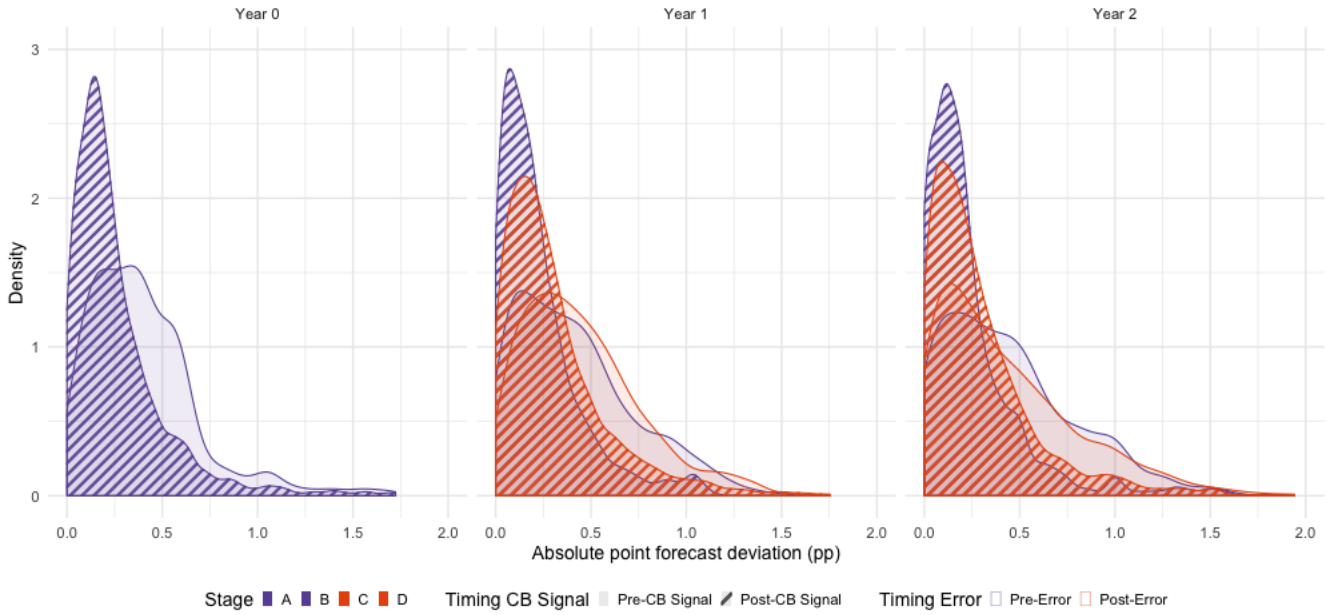
Throughout this section, we measure anchoring by the absolute deviation of participant forecasts from the central bank's (optimal) forecast, at Stage Z, where $Z \in \{A, B, C, D\}$:

$$\text{Anchoring}_{i,t}^{Z} = |\pi_{i,t}^{Z} - \pi_{cb,t}^{Z}| \tag{5}$$

3.3.1 Dynamics across treatments and error types

We begin by examining the degree of anchoring across participants at each Stage in the dynamic experiment, across all treatments and types of error observed by participants. Figure 7 plots the distribution of absolute deviations from the optimal forecast for each of Forecast Years 1-3.

Figure 7: 'Anchoring' by Stage: distribution of absolute forecast deviations from optimal



Notes: This figure presents the distribution of absolute point forecast deviation (in pp) across participants across Stages A-D for each of Years 0, 1, and 2. Purple density plots represent responses in Stages A and B – before participants have observed the forecast error – while orange density plots represent responses in Stages C and D – after participants have observed the forecast error. The translucent filled density plots represent responses in Stages A and C – before participants have observed the latest central bank forecast – while striped density plots represents responses in Stages B and D – after participants have observed the latest central bank forecast.

The purple distributions represent responses in Stages A and B, before the forecast error in Year 0 is revealed. The orange distributions represent responses in Stages C and D, post-error. The distributions without patterns represent responses in Stages A and C, before the (updated) central bank forecast, while those with striped patterns represent responses in Stages B and D, post-updated CB signal. The greater the right tail of the distribution, the greater the proportion of participants' forecasts with large absolute deviations from the optimal, the less 'anchored' are expectations.

Focusing on the patternless purple distributions first (Stage A), we see that priors are initially relatively dispersed. And this dispersion increases the further out the forecast horizon (the right-tail sequentially increases at Years 1 and 2, relative to Year 0). At the Year 2 horizon, for instance, the proportion of participants within 50bps of the optimal forecast is 65%, compared with 71% at the Year 0 horizon (statistics are provided in Table C.1).

Once participants observe the CB's forecast (Stage B, striped purple distribution), expectations tighten around the optimal forecast each forecast horizon. Specifically, the proportion of participants within 50bps of the optimal forecast at Year 2 increases from 65% to 90%. This represents an 'anchoring' around the CB forecast.

However, once the forecast error at Year 0 is revealed (Stage C, patternless orange distribution), we see that the distribution disperses once again with a large mass of respondents shifting right-wards at Forecast Years 1 and 2. This represents a 'de-anchoring' of expectations post-error. Indeed the distributions at Stage C look relatively similar to the dispersed priors at Stage A, with the proportion of participants within 50bps of the optimal forecast at Year 2 returning to 65% – the same as in Stage A. This suggests that observing the forecast error undoes most, if not all, of the initial anchoring from the CB forecast in Stage B.

Once participants observe the updated CB signal (Stage D, striped orange distributions), the mass tightens once again around the optimal forecast. However, it seems to do so to a lesser extent than pre-error, with 84% of participants' forecasts within 50bps of the optimal forecast (compared with 90% at Stage B). This indicates a 'persistent' loss of anchoring by the CB as a result of the forecast error, which could be indicative of persistent damage to the CB's reputation or credibility.

We confirm in our regression analysis below that these descriptive observations hold statistically.

3.3.2 Dynamics by error type

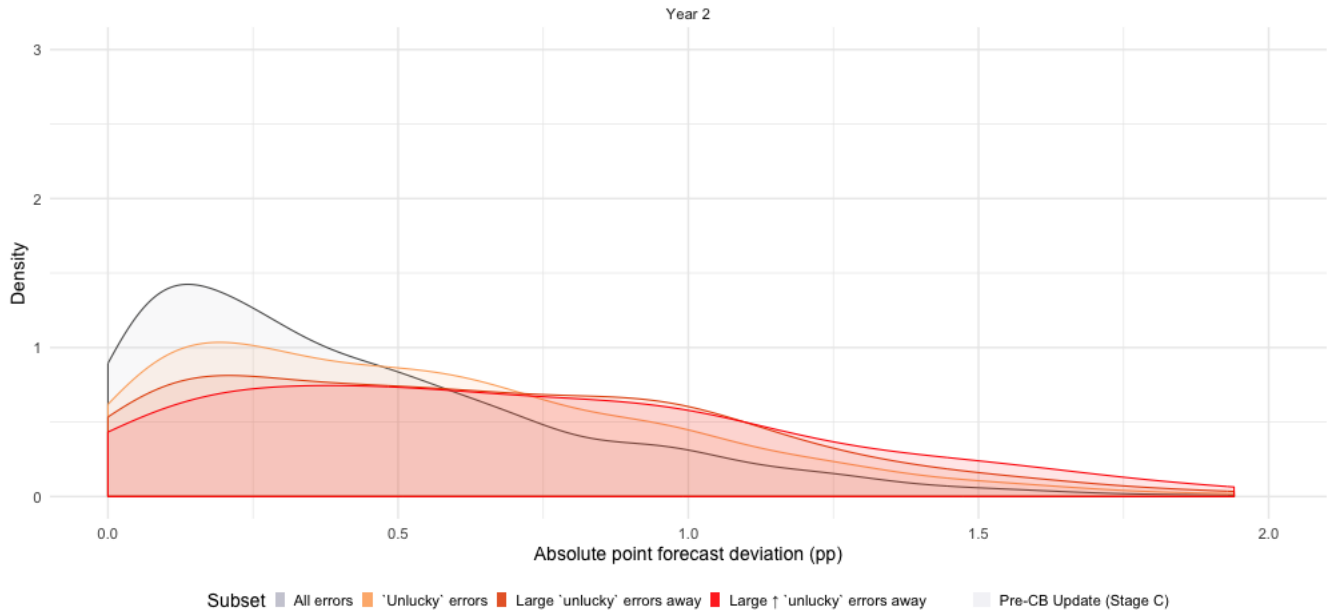
The key innovation of our experiment is examining what happens when the central bank's forecast inevitably misses. In Stage C, participants observe realized inflation that differs from the CB's forecast (and typically from their own forecast). We now examine differences in responses across participants who were exposed to different types of forecast error. In particular, we distinguish between errors that take the inflation series *toward* ("lucky") or *away from* ("unlucky") the CB's Year 2 forecast.

Figure 8 depicts the distribution of absolute forecast errors in Stage C, across the full sample of participants (in grey), across a sub-sample of participants who saw "unlucky" errors (in yellow), across the sub-sample who received *large* "unlucky" errors (in orange), and the sub-sample who received *large positive* "unlucky" errors. The latter two sub-samples allow us to observe whether participants' responses are influenced by the magnitude of the forecast miss, or whether there exists an asymmetry between shocks that take inflation further above or further below the forecast.[6]

We see that for each respective subset, the distribution of absolute forecast deviations shifts sequentially to the right. This implies an increasing degree of 'de-anchoring' of participants' expectations following each respective type of error. Regression analysis below confirms that these descriptive observations hold statistically.

[6]The potential asymmetry is motivated by existing empirical evidence that households are more sensitive, in updating their inflation expectations, to increases in inflation than they are to decreases (Anesti, Esady and Naylor, 2025).

Figure 8: De-anchoring post-forecast error (Stage C), by error type: distribution of absolute forecast deviations from optimal



Notes: This figure presents the distribution of absolute point forecast deviation (in pp) across participants in Stage C for Year 2. The grey density plot represents responses across all participants. The orange density plot represents responses across the subset of participants who observed an 'unlucky' forecast error. The dark red density plot represents responses across the subset of participants who observed a large 'unlucky' forecast error. The bright red density plot represents responses across the subset of participants who observed a large positive 'unlucky' forecast error.

3.3.3 Dynamics by treatment: Point versus Fans

We now compare dynamics by treatment. We do so by running the following regression specification:

$$\text{Anchoring}_{i,t}^{Z} = \beta_1 \text{Fan}_{i,r} + \beta_2 \text{Error Type}_{i,r} + \beta_3 \text{Fan}_{i,r} \times \text{Error Type}_{i,r} + \delta X_i + \gamma_i + \epsilon_{i,r} \quad (6)$$

where $\text{Fan}_{i,r}$ is a dummy variable that takes value 1 if participant i was treated in run r with a Fan forecast from the CB, and 0 if they were treated with a Point forecast. $\text{Error Type}_{i,r}$ is a categorical variable that represents which forecast error participant i in run r observed, where Error Type $\in \{\text{All}, \text{Unlucky}, \text{Unlucky large}, \text{Unlucky large positive}\}$. γ_i are participant fixed effects and $\epsilon_{i,r}$ is the error term. We also explicitly control for experimental factors X_i, which should be randomly assigned across treatments, including the worlds, run, and error each participant experienced alongside each treatment. We report results from this regression

in Table 3.[7]

Focusing first on Stage B in Column 1, we find evidence that point forecasts initially anchor expectations marginally more than fan charts (β_1). The effect is modest but statistically significant. Participants receiving Fan Chart forecasts have absolute deviations approximately 3 basis points larger than those receiving point forecasts. This initial anchoring advantage for point forecasts aligns with the findings of Rholes and Petersen (2021), Petersen and Rholes (2022), and Kostyshyna and Petersen (2023). In a one-shot setting, the apparent precision of a point forecast leads participants to place more weight on it.

Across Columns 2-5, we report results from Stage C – after participants have observed a forecast error. We report results across different error types observed by participants. Consistent with the observational evidence reported in Figure 7, all participants de-anchor to some degree following forecast errors, with absolute deviations increasing relative to Stage B (represented by the increase in the magnitude of the intercept). Further, consistent with the observations in Figure 8, we report that this de-anchoring is particularly pronounced for (particularly large and positive) "unlucky" errors, increasing the magnitude of the absolute deviation by up to 29 basis points on average (β_2).

Most importantly, examining β_3 and $\beta_1+\beta_3$ (the linear combination of the Fan Chart treatment effect), we see that fan charts materially mitigate this de-anchoring effect. Among participants who observed "unlucky" errors, those who had received Fan Charts in Stage B show significantly less de-anchoring than those who received Point forecasts. Economically, this effect is substantial: the Fan Chart treatment reduces the de-anchoring, relative to the point forecast by up to 25 basis points for the most damaging errors ($\beta_1 + \beta_3$). We interpret this result as indicating that Fan Chart treatment acts as an "insurance policy": by acknowledging upfront that errors are possible and even likely, it reduces the damage to credibility when errors occur. Indeed, the insurance value of Fan Charts is largest precisely when it matters most – when unlucky errors have the potential to be most damaging to the CB's credibility.

[7]We report the intercept from a version of this regression where participant fixed effects are replaced with demographic controls including participants' age, sex, English fluency, financial literacy, economic engagement and pre-treatment quiz performance – shown in Table D.1, for reference. We also note that we cannot report the value of β_2 for 'Unlucky Error' alone in the fixed effects model, as all participants who experienced an 'unlucky error' in run 1 will also have experienced an 'unlucky error' in run 2, by design. We show in Table D.1 that the narrative is consistent across each of these error types.

Table 3: Absolute Deviation by Stage, Horizon, and Error Type - Year 2

	Dependent variable: Absolute deviation of point forecast from optimal: Year 2								
	Stage B	Stage C				Stage D			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept (no FE model)	**0.278**	**0.522**	**0.375**	**0.365**	**0.406**	**0.366**	**0.326**	**0.329**	**0.342**
β_1									
Fan	**0.031*****	−0.012	**0.027***	**0.027***	**0.029***	0.011	0.016	0.016	0.019
	(0.011)	(0.015)	(0.016)	(0.016)	(0.017)	(0.012)	(0.014)	(0.014)	(0.013)
β_2									
Unlucky Large Error				**0.148*****				0.011	
				(0.042)				(0.033)	
Unlucky Large Positive Error					**0.290*****				0.030
					(0.072)				(0.053)
β_3									
Fan*Unlucky Error			**−0.076****				−0.009		
			(0.030)				(0.023)		
Fan*Unlucky Large Error				**−0.146****				−0.019	
				(0.061)				(0.046)	
Fan*Unlucky Large Positive Error					**−0.282*****				−0.056
					(0.074)				(0.056)
$\beta_1 + \beta_3$									
Fan + Fan*Error Type (linear comb.)			**−0.049****	**−0.120****	**−0.253*****		0.006	−0.003	−0.037
			(0.022)	(0.048)	(0.060)		(0.016)	(0.037)	(0.046)
Participant FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls: Experimental Variation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,598	1,598	1,598	1,598	1,598	1,598	1,598	1,598	1,598
Adjusted R^2	0.271	0.299	0.304	0.308	0.318	0.317	0.316	0.315	0.316

Notes: *p<0.1; **p<0.05; ***p<0.01. Cluster-robust standard errors in parentheses, clustered at participant level.

In Stage D, participants receive an updated CB forecast. We see that the intercept across Columns 6-9 reduces in size relative Columns 2-5 (Stage C), indicating that the CB is able to re-anchor expectations to some extent. However, consistent with the observations in Figure 7, the intercept remains larger than that in Column 1 (Stage B), suggesting persistent damage from the forecast error.

Focusing on the coefficient of key interest, β_3, in Stage D we see that coefficients are negative – consistent with a persistent insurance effect of the Fan Charts, following damaging errors. However, these coefficients are not statistically significant, indicating that the CB is able – in our financially incentivised experimental setting – to re-anchor expectations to some extent across both Point and Fan Chart treatments. In practice, there is a risk that people may 'disengage' with the CB following (especially damaging) forecast errors, which Fan Charts may protect against. Further work could benefit from investigating this further.

A potential concern with our within-subject design is that carry-over effects between runs

may differentially impact Point and Fan treatments. To address this concern directly, we re-estimate our main specifications (Tables 3 and D.2) using only Run 1 data, restricting the sample to a pure between-subjects comparison. We confirm in Table E.1 of Appendix E that each of our results hold.

3.3.4 Narrow versus Wide Fans

We also examine whether the width of the Fan Chart matters. Our design includes both narrow fans (70%) and wide fans (90%). We report results distinguishing across fan types in Table D.2. We find no significant difference between narrow and wide fans in their ability to mitigate de-anchoring. Both types of fans provide similar insurance effects against unlucky errors. This suggests that simply acknowledging uncertainty – rather than precisely calibrating its extent – is what matters for maintaining credibility.

3.4 Results: Uncertainty Perceptions

3.4.1 Measuring Uncertainty

In addition to eliciting point forecasts, we asked participants to provide confidence ranges around their forecasts. Specifically, we instructed them to: "Choose a lower and upper bound for inflation in each forecast year – pick bounds that you think will **almost certainly** include the true value."

We measure participants' perceived uncertainty as the width of their stated range relative to a sensible benchmark. Our benchmark is the interquartile range (IQR) from the historical data shown to participants (backhistory 2), which is 1.25 percentage points. We also compare to the model-based stochastic simulation IQR of 1.4 percentage points. The dependent variable in our analysis is:

$$\text{Range Ratio}_{i,t}^{Z} = \frac{r_{i,t}^{Z}}{\text{IQR}_{-12:-1}} \tag{7}$$

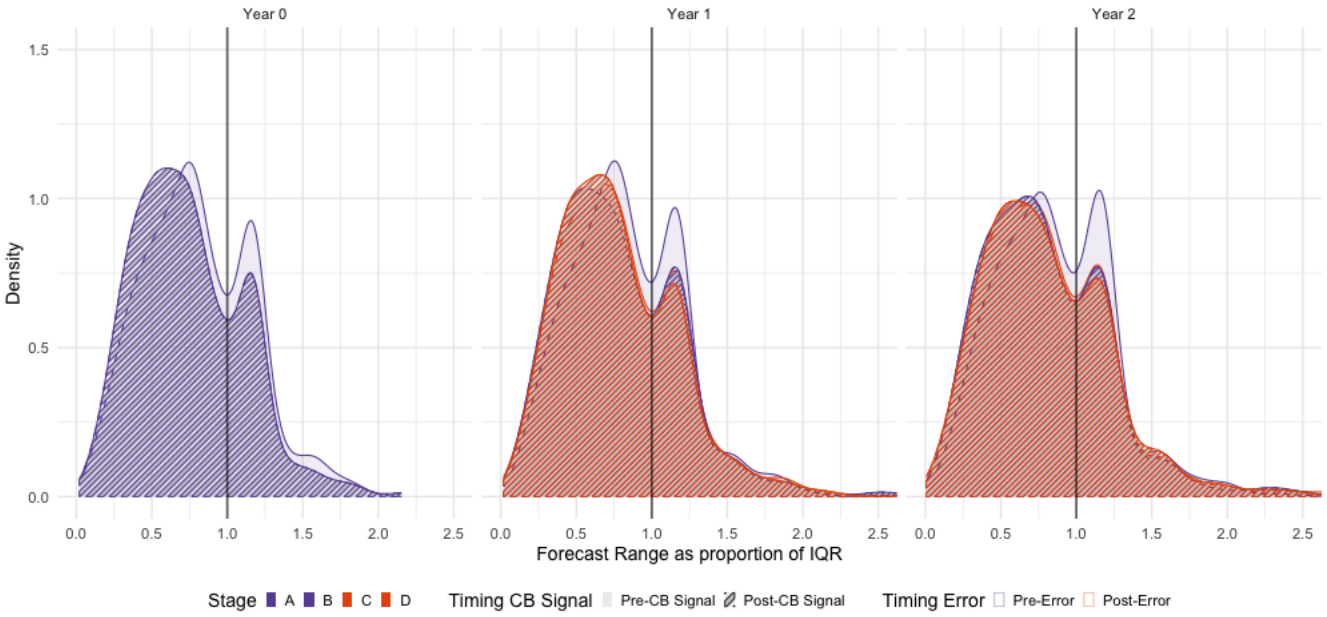
A ratio below 1 indicates "overconfidence" (ranges narrower than the historical IQR), while a ratio above 1 indicates "appropriate" or conservative uncertainty perceptions.

3.4.2 Overconfidence in the General Public

Figure 9 presents the distribution of range ratios by Stage A-D, and Forecast Years 0-2, across all treatments and error types. The legend key is the same as that in Figure 7. A striking pattern emerges: the public is dramatically overconfident at each stage of the experiment. The majority of participants (between 65% and 75% across stages and years) choose a range that is materially smaller than 1; implying ranges narrower than the historical IQR. While there is little change in the distributions across stages, the size of the range seems to becomes even smaller after participants see the CB forecast, and remain largely unchanged following the forecast error.

One notable difference between Stage A and B is that the second 'hump' in the distribution, just above 1, present in Stage A reduces in size in Stage B. This hump corresponds with the default range of 1.4pp (ratio of approximately 1.1), suggesting that many participants initially simply accepted the pre-filled value. On seeing the CB forecast, many of these participants opted to change (specifically, reduce) the range.

Figure 9: Uncertainty perceptions by Stage: distribution of forecast range ratio as a proportion of IQR



Notes: This figure presents the distribution of forecast ranges as a proportion of the inter-quartile range (IQR) of the inflation backhistory that participants observed, across participants across Stages A-D for each of Years 0, 1, and 2. Purple density plots represent responses in Stages A and B – before participants have observed the forecast error – while orange density plots represent responses in Stages C and D – after participants have observed the forecast error. The translucent filled density plots represent responses in Stages A and C – before participants have observed the latest central bank forecast – while striped density plots represents responses in Stages B and D – after participants have observed the latest central bank forecast. The solid vertical black line represents a forecast range that is equal to the IQR.

3.4.3 Uncertainty across Treatments: Fans vs Points

We test differences across treatments by employing the following regression specification:

$$\text{Range Ratio}_{i,r}^{Z} = \beta_1 \text{Fan}_{i,r} + \beta_2 \text{Error Type}_{i,r} + \beta_3 \text{Fan}_{i,r} \times \text{Error Type}_{i,r} + \delta X_{i,r} + \gamma_i + \epsilon_{i,r} \quad (8)$$

where the RHS variables are defined as in Equation 6. Table 4 reports the results.[8]

Consistent with observations from Figure 9, we see that the magnitude of the intercept – across all Stages – is smaller than 1. This confirms that, on average, the general public are over-confident in their perception of uncertainty.

However, across all Stages, we see that the coefficient β_1 is positive and significant. This indicates that Fan Charts increase range widths toward more realistic levels. In Stage B (Column 1), Fan Charts increase ranges by 0.055 (5.5% of the IQR), significant at the 1% level. This effect persists and even grows in Stages C and D, with evidence that this effect is more pronounced for "unlucky errors".[9]

Table 4: Part II: Regression results: Uncertainty

	Dependent variable: Range as proportion of IQR by Stage, Horizon, and Error Type: Year 2								
	Stage B	Stage C				Stage D			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept (no FE)	**0.790**	**0.832**	**0.823**	**0.818**	**0.824**	**0.845**	**0.815**	**0.819**	**0.834**
β_1									
Fan	**0.055*****	**0.035*****	0.024	0.023	**0.029***	**0.041*****	**0.052*****	**0.052*****	**0.038*****
	(0.014)	(0.013)	(0.017)	(0.017)	(0.015)	(0.013)	(0.016)	(0.016)	(0.014)
$\beta_1 + \beta_3$									
Fan + Fan*Error Type			**0.044****	**0.072***	0.076		0.029	0.007	0.062
			(0.018)	(0.040)	(0.051)		(0.018)	(0.040)	(0.050)
Error Type	All	All	Unlucky	Unlucky Big	Unlucky Big ↑	All	Unlucky	Unlucky Big	Unlucky Big ↑
Participant FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls: Experimental Variation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,598	1,598	1,598	1,598	1,598	1,598	1,598	1,598	1,598
Adjusted R^2	0.553	0.640	0.635	0.640	0.639	0.655	0.654	0.654	0.654

Notes: *p<0.1; **p<0.05; ***p<0.01. Cluster-robust standard errors in parentheses, clustered at participant level.

[8] Again, the Intercept is taken from a version of the regression without fixed effects, reported in Table D.3, for reference.

[9] Though, perhaps surprisingly, we do not see significant effects on participants' ranges of observing an unlucky error in isolation (β_2) – see Table D.3.

3.4.4 Narrow versus Wide Fans

We again compare the relative effects of narrow versus wide fans. In contrast to our results on anchoring in Section 3.3, we find in Table 5 significant differences across fan widths.

In particular, we see that wide fans have substantially larger effects than narrow fans. Wide fans increase ranges by approximately 0.10 relative to point forecasts (10% of IQR), while narrow fans show smaller and less consistent effects. This effect is especially large for (particularly large and positive) unlucky errors, where Fan Charts increase ranges by up to 14% of IQR relative to point forecasts.

We interpret these results as suggesting that Fan Charts – particularly wide fans – help the public "learn" more realistic uncertainty perceptions. By visualizing the range of possible outcomes, fans correct the systematic overconfidence we observe in prior beliefs. This learning appears to persist: even after multiple stages, participants who received fan charts maintain wider, more realistic ranges.[10]

Table 5: Part II: Regression results: Uncertainty (by fan width)

	Dependent variable: Range as proportion of IQR by Stage, Horizon, and Error Type: Year 2								
	Stage B	Stage C				Stage D			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept (no FE)	**0.790**	**0.833**	**0.823**	**0.817**	**0.825**	**0.846**	**0.815**	**0.818**	**0.834**
β_1^{Narrow}									
Fan Narrow	0.013	0.019	−0.001	0.018	0.019	0.025	**0.038***	**0.048****	0.029
	(0.020)	(0.017)	(0.023)	(0.022)	(0.020)	(0.017)	(0.022)	(0.020)	(0.019)
$\beta_1^{Narrow} + \beta_3^{Narrow}$									
Fan Narrow + Fan Narrow*Error Type			0.040	0.025	0.025		0.011	-0.045	0.003
			(0.025)	(0.046)	(0.059)		(0.025)	(0.046)	(0.059)
β_1^{Wide}									
Fan Wide	**0.097*****	**0.051*****	**0.049****	0.029	**0.040***	**0.057*****	**0.067*****	**0.057****	**0.047****
	(0.020)	(0.019)	(0.025)	(0.023)	(0.020)	(0.019)	(0.025)	(0.023)	(0.020)
$\beta_1^{Wide} + \beta_3^{Wide}$									
Fan Wide + Fan Wide*Error Type			**0.054****	**0.128*****	**0.142****		**0.047***	0.069	**0.138****
			(0.025)	(0.049)	(0.065)		(0.025)	(0.049)	(0.065)
Error Type	All	All	Unlucky	Unlucky Big	Unlucky Big ↑	All	Unlucky	Unlucky Big	Unlucky Big ↑
Participant FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls: Experimental Variation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,598	1,598	1,598	1,598	1,598	1,598	1,598	1,598	1,598
Adjusted R^2	0.558	0.640	0.640	0.640	0.640	0.655	0.654	0.655	0.655

Notes: *p<0.1; **p<0.05; ***p<0.01. Cluster-robust standard errors in parentheses, clustered at participant level.

[10] In turn, this could have implications for people's ability to make sensible economic decisions (see e.g., Coibion et al. 2024; Kostyshyna and Petersen 2024).

4 Conclusion and Policy Implications

We examine how central banks can effectively communicate forecast uncertainty through a two-part experimental study. Part I establishes that fan charts are well understood and best at jointly conveying expectations and uncertainty across various media. Part II implements a novel dynamic information experiment, revealing that while point forecasts anchor marginally better initially, this advantage disappears when inevitable forecast errors occur. Importantly, fan charts help the public learn more realistic perceptions of economic uncertainty, correcting systematic overconfidence.

Our findings have direct implications for central bank communication policy. The current debate, sparked by the Bernanke Review, centres on whether fan charts should be abandoned. Bernanke (2024) argued that fan charts "*conveyed little useful information over and above what could communicated in other, more direct ways*". One option would be to replace them with 'simpler' communication of uncertainty. Our evidence suggests caution before rushing to replace fan charts.

Our experiment reveals a subtle but important dynamic trade-off. In a static, one-shot setting (Stage B), point forecasts anchor expectations marginally more than fan charts. If this were the end of the story, one might conclude that simpler is better—why complicate communication with distributional forecasts if they reduce anchoring?However, this perspective ignores the inevitable reality of forecast errors. No forecaster, however skilled, can consistently hit precise point forecasts for macroeconomic variables. In our experiment, even though the central bank forecasts are optimal and unbiased, they miss—often substantially. When this happens, the initial anchoring advantage of point forecasts evaporates.

Fan charts provide insurance against this de-anchoring. By acknowledging upfront that forecasts are uncertain and errors are likely, fan charts reduce the credibility damage when errors occur. In Stage C, following large unlucky errors, participants who received fan charts remain substantially better anchored than those who received point forecasts. This insurance effect persists into Stage D when the central bank attempts to re-establish its influence.

Economically, the insurance value substantially outweighs the initial anchoring cost. It is around five times larger. The initial anchoring disadvantage of fans is approximately 5 basis points. The insurance benefit for large unlucky errors is approximately 26 basis points. Even if large unlucky errors occur only occasionally, the expected value of the insurance effect likely exceeds the certain cost of reduced initial anchoring.

Beyond the anchoring dimension, we find that fan charts serve an important uncertainty communication function. The public systematically underestimates the degree of uncertainty surrounding economic forecasts, exhibiting substantial overconfidence. Fan charts help correct this misperception, bringing subjective uncertainty assessments closer to objective benchmarks. This is important because if individuals underestimate economic uncertainty, they may make suboptimal decisions—saving too little for precautionary motives, investing too aggressively, or responding too strongly to each new data point. By helping the public develop more realistic uncertainty perceptions, central banks may improve the quality of economic decision-making throughout the economy. Our results suggest that this benefit is particularly strong for "wide" fans that emphasize the substantial range of possible outcomes.

While our Part II experiment examined the effects of fan charts as the communication medium, the dynamic trade-off we emphasise is likely present for other communication media. And the Part I results suggest that fan charts are well understood and best at jointly conveying expectations and uncertainty. This suggests that in countries that already have a long-track record with fan charts, it is likely there is some benefit to keeping with them. But elsewhere, other communication media could also work.

Our Part I results reveal that box-and-whisker plots and speedometers can convey uncertainty, but the speedometer is particularly disliked (especially by experts), and box-and-whisker plots may be unfamiliar to general audiences. Dot plots, while popular in some central banks (notably the Federal Reserve), represent disagreement among forecasters rather than uncertainty about outcomes. Our results show that general audiences often misinterpret dot plots as representing uncertainty, hallucinating distributions from what is fundamentally a different concept.

Verbal descriptions of uncertainty ("inflation will likely be between X and Y") or skewed risk distributions (e.g., "inflation risks are tilted to the upside") could potentially work – Hansen et al. (2019) suggest that the text of the Bank of England's Inflation Report communicates additional uncertainty to financial markets than is conveyed in the numerical forecast information. However, the lack the visual salience of fan charts and may not achieve the same educational benefits for uncertainty perceptions. This is an area for future research.

Similarly for the use of "scenarios", or "conditional" uncertainty outlooks. Understanding how to communicate asymmetric or scenario-dependent risks remains an important area for future research.

Finally, for central banks that use fan charts, there is an important issue of exactly how the

fan charts are constructed. In our model environment, we use stochastic simulations from the data generating process to create the true distribution of possible future outcomes. In practice, there is considerable uncertainty about the correct underlying model and so other options typically considered include using a window of recent forecast errors, or using policymaker adjustments. This challenge applies to any visual representation of uncertainty. Individual central banks, with different forecasting infrastructures and different decision-making bodies, have to assess the best way to calculate the uncertainty to communicate.

But the key decision is whether to communicate uncertainty. The dynamic trade-off between influence today and credibility tomorrow, alongside the desire to adequately convey how much uncertainty households and firms should perceive, favours transparency about the central bank's thinking on uncertainty. This means that the bar for abandoning a reasonably well-understood medium of communication should be high. So unless a central bank considering reforming its communication strategy has a suitable alternative, our results caution against abandoning uncertainty visualization. And those without such visualization should consider adopting one.

References

Anesti, Nikoleta, Vania Esady, and Matthew Naylor, "Food Prices Matter Most: Sensitive Household Inflation Expectations," Staff Working Paper 1125, Bank of England April 2025.

Bauer, Michael D, Ben S Bernanke, and Eric Milstein, "Risk Appetite and the Risk-Taking Channel of Monetary Policy," *Journal of Economic Perspectives*, 2023, *37* (1), 77–100.

Bernanke, Ben, "Forecasting for monetary policy making and communication at the Bank of England: A review," *Bank of England*, 2024.

Blinder, Alan S., Michael Ehrmann, Marcel Fratzscher, Jakob De Haan, and David-Jan Jansen, "Central Bank Communication and Monetary Policy: A Survey of Theory and Evidence," *Journal of Economic Literature*, December 2008, *46* (4), 910–45.

Brainard, William C., "Uncertainty and the Effectiveness of Policy," *American Economic Review*, 1967, *57* (2), 411–425.

Campbell, Jeffrey R., Filippo Ferroni, Jonas D. M. Fisher, and Leonardo Melosi, "The Limits of Forward Guidance," *Journal of Monetary Economics*, December 2019, *108*, 118–134.

Cieslak, Anna and Michael McMahon, "Tough Talk: The Fed and the Risk Premium," CEPR Discussion Paper 19313 2024.

Coibion, Olivier, Dimitris Georgarakos, Yuriy Gorodnichenko, Geoff Kenny, and Michael Weber, "The Effect of Macroeconomic Uncertainty on Household Spending," *American Economic Review*, March 2024, *114* (3), 645–77.

Fadda, Pietro, Rayane Hanifi, Klodiana Istrefi, and Adrian Penalver, "Central Bank Communication of Uncertainty," CEPR Discussion Paper 17728, CEPR 2023.

Fischer, Johannes J., Christoph Herler, and Philip Schnattinger, "When the Fog Clears: The Effect of Reduced Inflation Uncertainty on Households' Financial Behaviour," Staff Working Paper 1133, Bank of England June 2025. Published 27 June 2025.

Hansen, LarsPeter and Thomas J. Sargent, "Robust Control and Model Uncertainty," *American Economic Review*, May 2001, *91* (2), 60–66.

Hansen, Stephen, Michael McMahon, and Matthew Tong, "The long-run information effect of central bank communication," *Journal of Monetary Economics*, 2019, *108* (C), 185–202.

Kostyshyna, Olena and Luba Petersen, "Communicating Inflation Uncertainty and Household Expectations," Staff Working Paper 63, Bank of Canada 2023.

_ and _, "The Effect of Inflation Uncertainty on Household Expectations and Spending," Technical Report, National Bureau of Economic Research 2024.

Levin, Andrew, Volker Wieland, and John C. Williams, "The Performance of Forecast-Based Monetary Policy Rules Under Model Uncertainty," *American Economic Review*, June 2003, *93* (3), 622–645.

Lusardi, Annamaria and Olivia S. Mitchell, "Financial Literacy and Planning: Implications for Retirement Wellbeing," in Annamaria Lusardi and Olivia S. Mitchell, eds., *Financial Literacy: Implications for Retirement Security and the Financial Marketplace*, Oxford: Oxford University Press, 2011, pp. 17–39.

McMahon, Michael and Ryan Rholes, "Building central bank credibility: The role of forecast performance," Technical Report, Centre for Economic Policy Research 2024.

Petersen, Luba and Ryan Rholes, "Macroeconomic expectations, central bank communication, and background uncertainty: A COVID-19 laboratory experiment," *Journal of Economic Dynamics and Control*, 2022, *143*, 104460.

Rholes, Ryan and Luba Petersen, "Should central banks communicate uncertainty in their projections?," *Journal of Economic Behavior & Organization*, 2021, *183*, 320–341.

Söderström, Ulf, "Monetary Policy with Uncertain Parameters," *The Scandinavian Journal of Economics*, March 2002, *104* (1), 125–145.

Spiegelhalter, David, *The Art of Uncertainty: How to Navigate Chance, Ignorance, Risk and Luck*, London: Pelican Books, 2024.

van der Bles, Anne Marthe, Sander van der Linden, Alexandra L. J. Freeman, and David J. Spiegelhalter, "The Effects of Communicating Uncertainty on Public Trust in Facts and Numbers," *Proceedings of the National Academy of Sciences of the United States of America*, 2020, *117* (14), 7672–7683.

Appendix A Part I: Experimental Design details

A.1 Questions

Subjective questions

1. "To what extent do you think you understand this chart?" (1-7 scale, with 1 = 'Very difficult', 4 = 'Neither easy nor difficult', 7 = 'Very easy')
2. "How much uncertainty does this chart show?" (1-5 scale, with 1 = 'Not at all uncertain', 2 = 'A little uncertain', 3 = 'Moderately uncertain', 4 = 'Very uncertain', 5 = 'Extremely uncertain', or 'Can't tell')

Objective questions

Questions testing understanding of central forecast ('anchoring'):

3. "*What do you think is the most likely value for the series in forecast period 1?*" (Responses rewarded as 'correct' within ±0.3pp)

Questions testing understanding of uncertainty communicated:

4. "*Complete the sentence by replacing the [?] with one of the options below: The figure shows that, between Forecast Period 1 and Forecast Period 3, inflation [?].*" ['Will definitely increase', 'Is most likely to increase', 'Will definitely stay the same', 'Is most likely to stay the same', 'Is most likely to decrease', 'Will definitely decrease', 'Do not know', 'Cannot tell - I need more information']
   (Possible correct responses underlined, with direction depending on the 'world' observed.)

   For 'unconditional' media only:

5. "Complete the sentence by replacing the [?] with one of the options below: Inflation outcomes closer towards the forecasts are [?] than inflation outcomes further from the forecasts." ['More likely', 'Equally likely', 'Less likely', 'I do not know']
   (Correct response underlined.)

6. "*How likely is it that inflation will be between X% and Y% in forecast period 3?*" ['0%', '1%-24%', '25%-49%', '50%-74%', '75%-99%', '100%', 'Cannot tell - I need more information', 'Do not know']
   (Correct responses vary across media, as specified in Figure 5.)

   For 'conditional' media (all) only:

7. "*What is this chart trying to convey? Tick all that apply*" ['If the 'event' materialises, the 'Scenario' outcomes will definitely occur', 'If the 'event' materialises, the 'Scenario' outcomes are most likely to occur', 'If the 'event' materialises, the 'Baseline' outcomes could still occur', 'If the 'event' does not materialise, the 'Baseline' outcomes will definitely occur', 'If the 'event' does not materialise, the 'Baseline' outcomes are most likely to occur', 'If the 'event' does not materialise, the 'Scenario' outcomes could still occur', 'Both 'Scenario' and 'Baseline' outcomes will definitely occur', 'It is possible that neither outcome occurs', 'It is certain that neither outcome will occur', 'None of the above', 'I do not know']
   (Correct responses underlined.)

   For 'conditional' media (Tree Diagram, Point and Scenario) only:

8. "*Complete the sentence by replacing the [?] with one of the options below: It is [?] likely that inflation outcomes are above the baseline forecasts than below.*" ['More likely', 'Equally likely', 'Less likely', 'I do not know', 'Can't tell - I need more information']
   (Correct response underlined.)

   For 'conditional' media (Point and Scenario with Fan, Short and Long) only:

9. "*Complete the sentence by replacing the [?] with one of the options below: Inflation outcomes in Forecast Period 3 are [?] than those in Forecast Period 1.*" ['Less uncertain', 'Equally uncertain', 'More uncertain', 'I do not know', 'Cannot tell']
   (Correct response underlined.)

A.2 Inflation Series across Worlds

Table A.1: Summary statistics of inflation series by World

World	Mean	Median	SD	Min	Max
World 1	2.30	2.40	0.86	1.00	4.13
World 2	2.91	2.80	1.38	1.00	4.98
World 3	4.17	4.15	1.26	2.44	6.12
World 4	2.93	3.03	1.26	1.00	4.68
World 5	2.33	1.91	1.59	0.48	4.87

Appendix B Part I: Auxiliary Results

B.1 Summary Statistics

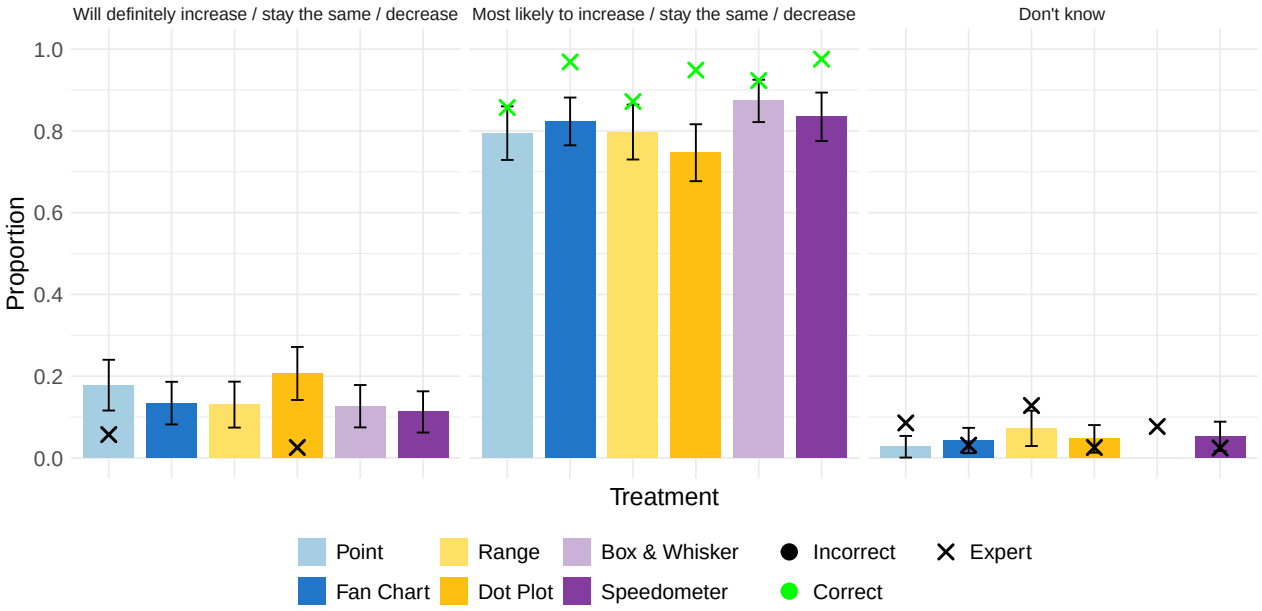
Table B.1: Summary Statistics by Treatment

		Demographic characteristics						Pre-Treatment Qs, Experience, and Time		
Treatment	N	Age	Sex	Born in UK	Student	Employed	Econ. Exp.	Fin. Lit.	Past surveys	Time taken
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Point Forecast	146	47.80	0.44	0.85	0.17	0.63	0.32	0.73	1856.77	18.99
Fan	164	46.24	0.48	0.90	0.19	0.67	0.29	0.72	1692.19	18.04
Range	138	45.16	0.50	0.86	0.21	0.66	0.34	0.67	1828.70	17.97
Dot	150	47.70	0.45	0.85	0.19	0.66	0.33	0.71	1889.50	18.22
Box & Whisker	158	44.92	0.53	0.85	0.21	0.69	0.33	0.75	1721.11	19.31
Speedometer	151	47.11	0.48	0.89	0.11	0.67	0.28	0.70	1806.28	18.61
ANOVA p-value		0.51	0.74	0.75	0.76	1.00	0.99	0.80	0.96	0.86

Notes: This table reports, by unconditional 'media', mean values across a range of participant characteristics, including demographics, responses to pre-treatment questions, past survey experience, and time taken to complete the survey. Column (1) reports the number of participants treated by each media. Column (2) reports the mean Age. Column (3) reports the proportion of participants that are Male. Column (4) reports the proportion of participants born in the UK. Column (5) reports the proportion of participants that are students. Column (6) reports the proportion of participants in either full- or part-time employment. Column (7) reports the proportion of participants that have some economics or finance experience – including having studied either at school, university, or doing at least one of them as part of their job. Column (8) reports the proportion of participants who correctly responded to a set of three Financial Literacy questions, following Lusardi and Mitchell (2011). These questions are as follow. Q1. "*Suppose you had £100 in a savings account and the interest rate was 2% per year. After 5 years, how much do you think you would have in the account if you left the money to grow?*" with options: ['More than £102', 'Exactly £102', 'Less than £102', 'Do not know', 'Refuse to answer']. The correct answer is 'More than £102'. Q2. "*Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After 1 year, how much would you be able to buy with the money in this account?*" with options ['More than today', 'Exactly the same', 'Less than today', 'Do not know', 'Refuse to answer']. The correct answer is 'Less than today'. Q3. "*Please tell me whether this statement is true or false. "Buying a share in a single company usually provides a safer return than the same value of a diversified share fund.*"" with options ['True', 'False', 'Do not know', 'Refuse to answer']. The correct answer is 'False'. Column (9) reports the mean number of previous surveys participants have completed with Prolific. Column (10) reports the mean time taken to complete the survey across participants who were treated with media. The bottom row reports p-values from ANOVA tests of whether any mean outcomes are statistically different across groups, across the set of pre-treatment variables. p-values greater than 0.1 indicate that we cannot reject that the mean outcomes are statistically equal. We also confirm that treatments do not predict covariates in joint tests using OLS regressions, for each variable listed in the table.

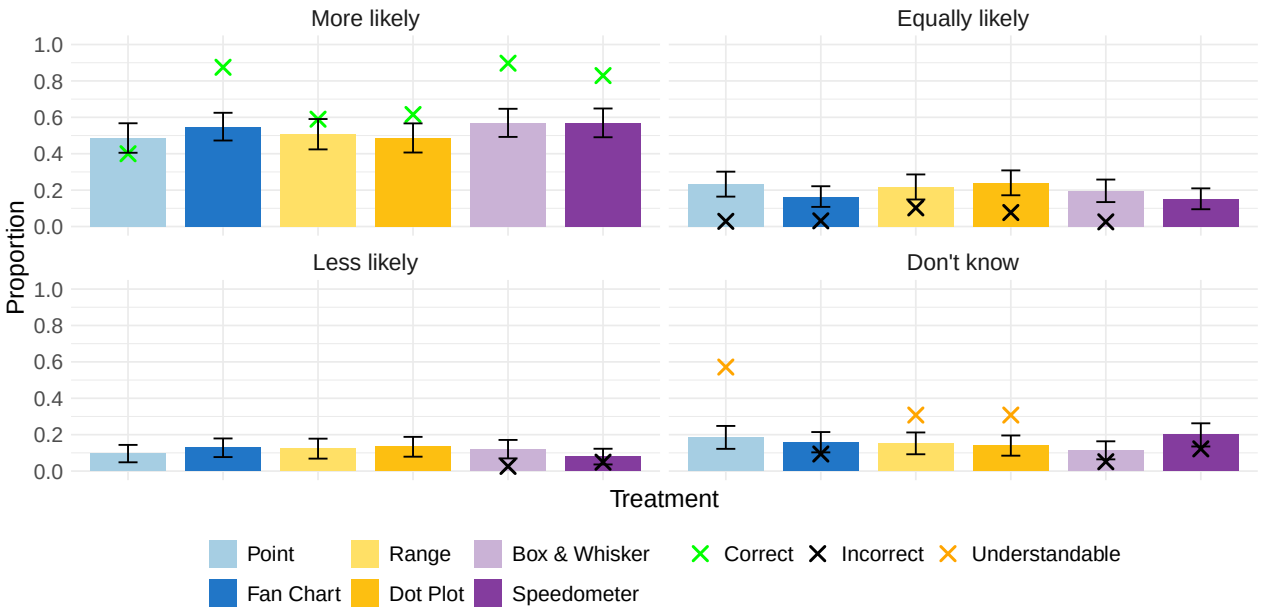
B.2 Descriptive results

Figure B.1: Objective understanding of uncertainty conveyed (presence of uncertainty): Q4. "The figure shows that between Forecast Periods 1 and 3 inflation ..."



Notes: This figure presents the distribution of responses to Q4. "*Complete the sentence by replacing the [?] with one of the options below: The figure shows that, between Forecast Period 1 and Forecast Period 3, inflation [?].*" across 'unconditional' media, with histograms reflecting the proportion of respondents selecting one of: ['Will definitely increase', 'Will definitely stay the same', 'Will definitely decrease'] in the left hand panel; ['Is most likely to increase', 'Is most likely to stay the same', 'Is most likely to decrease'] in the central panel; ['Do not know', 'Cannot tell - I need more information'] in the right hand panel. The bars represent responses amongst the General Public sample with 5% confidence intervals represented by the whiskers. Crosses represent responses amongst the Experts sample. Responses in the central panel are 'correct', reflected by the green colour of the crosses. Light blue bars represent responses for the Point forecast, dark blue bars represent responses for the Fan Chart, yellow bars represent responses for the Range, orange bars represent responses for the Dot Plot, light purple bars represent responses for the Box & Whisker, dark purple bars represents responses for the Speedometer.

Figure B.2: Objective understanding of uncertainty conveyed (distribution of possible outcomes): Q6. "Inflation outcomes closer towards the forecasts are [?] than inflation outcomes further from the forecasts."



Notes: This figure presents the distribution of responses to Q6. "*Complete the sentence by replacing the [?] with one of the options below: Inflation outcomes closer towards the forecasts are [?] than inflation outcomes further from the forecasts.*." across 'unconditional' media, with histograms reflecting the proportion of respondents selecting one of: ['More likely', 'Equally likely', 'Less likely', 'Do not know' or 'Cannot tell - I need more information']. The bars represent responses amongst the General Public sample with 5% confidence intervals represented by the whiskers. Crosses represent responses amongst the Experts sample. Responses of 'More likely' are 'correct', reflected by the green colour of the crosses. Responses of 'Do not know' or 'Cannot tell - I need more information' are labelled as 'understandable', reflected by the orange colour of the crosses, as while these responses are not strictly correct they are also not strictly incorrect. Light blue bars represent responses for the Point forecast, dark blue bars represent responses for the Fan Chart, yellow bars represent responses for the Range, orange bars represent responses for the Dot Plot, light purple bars represent responses for the Box & Whisker, dark purple bars represents responses for the Speedometer.

B.3 Regression results

B.3.1 No participant fixed effects

$$Y_{i,r} = \beta_0 + \beta_1 \text{Treatment}_{i,r} + \beta_2 \text{Expert}_i + \beta_3 \text{Treatment}_{i,r} \times \text{Expert}_i + \epsilon_{i,r} \tag{9}$$

Table B.2: Part I: Baseline results (no participant fixed effects)

	Dependent variable: Responses					
	Subjective		Anchoring	Uncertainty		
	Understanding	Degree of Uncertainty	Point forecast	Presence of uncertainty	Uncertainty distribution	Forecast range probability
	(1)	(2)	(3)	(4)	(5)	(6)
$\underline{\beta_0}$						
Constant	**5.181*****	**2.190*****	**0.802*****	**0.699*****	**0.536*****	0.002
	(0.144)	(0.112)	(0.040)	(0.046)	(0.051)	(0.023)
$\underline{\beta_1}$						
Fan	**−0.839*****	**0.702*****	−0.064	−0.063	0.064	**0.476*****
	(0.165)	(0.122)	(0.048)	(0.052)	(0.057)	(0.039)
Range	**−1.300*****	**0.706*****	−0.004	−0.061	0.024	**0.181*****
	(0.171)	(0.141)	(0.048)	(0.055)	(0.060)	(0.033)
Dots Plots	**−0.902*****	**0.229***	−0.015	**−0.094***	0.002	0.0001
	(0.171)	(0.121)	(0.047)	(0.054)	(0.059)	(0.001)
Box & Whisker	**−1.152*****	**0.647*****	−0.042	0.044	0.085	**0.411*****
	(0.166)	(0.121)	(0.048)	(0.050)	(0.057)	(0.039)
Speedometer	**−1.641*****	**0.530*****	0.060	0.014	0.086	**0.411*****
	(0.176)	(0.124)	(0.044)	(0.051)	(0.058)	(0.040)
$\underline{\beta_2}$						
Expert	**1.312*****	**−0.741***	**0.170*****	**0.121***	−0.082	0.0001
	(0.180)	(0.388)	(0.044)	(0.070)	(0.094)	(0.002)
$\underline{\beta_3}$						
Fan*Expert	**0.471***	**1.023****	0.030	0.110	**0.414*****	**0.337*****
	(0.277)	(0.407)	(0.070)	(0.094)	(0.118)	(0.080)
Range*Expert	**0.574***	**1.593*****	0.007	0.004	0.158	**−0.105***
	(0.324)	(0.421)	(0.061)	(0.103)	(0.130)	(0.054)
Dot Plots*Expert	−0.080	0.451	0.018	**0.187****	0.211	−0.0001
	(0.298)	(0.410)	(0.061)	(0.088)	(0.129)	(0.002)
Box & Whisker*Expert	0.223	**1.095*****	0.045	−0.080	**0.411*****	−0.129
	(0.297)	(0.409)	(0.061)	(0.099)	(0.113)	(0.082)
Speedometer*Expert	**−1.226*****	**1.067*****	−0.055	0.081	**0.343*****	**0.223*****
	(0.322)	(0.408)	(0.057)	(0.085)	(0.118)	(0.086)
$\underline{\beta_2 + \beta_3}$						
Expert + Fan*Expert	**1.783*****	**0.282****	**0.200*****	**0.231*****	**0.332*****	**0.337*****
	(0.210)	(0.123)	(0.055)	(0.063)	(0.071)	(0.079)
Expert + Range*Expert	**1.886*****	**0.852*****	**0.177*****	**0.125***	0.077	**-0.104***
	(0.267)	(0.163)	(0.043)	(0.076)	(0.090)	(0.054)
Expert + Dot Plots*Expert	**1.232*****	**-0.290****	**0.188*****	**0.308*****	0.129	0.000
	(0.236)	(0.133)	(0.042)	(0.053)	(0.088)	(0.000)
Expert + Box & Whisker*Expert	**1.535*****	**0.354*****	**0.215*****	0.041	**0.329*****	-0.129
	(0.236)	(0.129)	(0.042)	(0.070)	(0.063)	(0.082)
Expert + Speedometer*Expert	0.087	**0.326****	**0.115*****	**0.202*****	**0.261*****	**0.224*****
	(0.266)	(0.127)	(0.037)	(0.049)	(0.071)	(0.085)
Participant FE	No	No	No	No	No	No
Controls: Experimental Variation	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,132	951	1,132	1,132	1,132	1,132
Adjusted R^2	0.219	0.131	0.034	0.030	0.041	0.255

Note: *p<0.1; **p<0.05; ***p<0.01. Table reports results from Equation 9. Cluster-robust standard errors in parentheses, clustered at participant level.

B.3.2 Composite 'scores'

Table B.3: Part I: Composite 'scores' across regression results (Expert)

Experts	**Subjective** Scaled: $[-1, 1]$		**Objective** Unscaled: β = 'proportion correct'						
			Anchoring	**Uncertainty**			**Accuracy 'Score'**		
	Q1.Liked?	Q2.How uncertain?	Q3.Point Estimate	Q4.Presence	Q5.Distribution	Q6.Probabilities	$\alpha = 0.25$	$\alpha = 0.5$	$\alpha = 0.75$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Point	**0.83**	'Can't tell'	0.97	0.82	0.45	0.00	0.56	0.70	0.84
Fan	**0.82**	-	0.94	0.90	0.87	0.80	**0.87**	**0.89**	**0.91**
Range	**0.65**	-	0.95	0.75	0.59	0.06	0.59	0.71	0.83
Dot Plots	0.57	-	0.96	0.88	0.61	0.00	0.61	0.73	0.85
Box & Whisker	0.57	-	0.97	0.79	0.87	0.29	**0.73**	**0.81**	**0.89**
Speedometer	-0.01	-	0.97	0.87	0.82	0.65	**0.83**	**0.88**	**0.92**

Scaling for 'Subjective' coefficients: $\left[\frac{\text{Intercept}+\beta_1+\text{Expert}+\beta_2-1}{\gamma-1} - 0.5\right] \times 2$ where $\gamma = [7, 5]$ for [Liked?, How much uncertainty?]

Accuracy 'Score' $= \alpha \times \beta^{Anchoring} + \alpha \times \bar{\beta}^{Uncertainty}$

Notes: This table reports summarised versions of the estimated coefficients reported in Table 1 for the expert sample based on Equation 1. Columns (1) and (2) report Intercept (no FE) + β_1+Expert+β_2 values from Columns (1) and (2), respectively, of Table 1, scaled to a -1 to 1 scale. Columns (3)-(6) report Intercept (no FE) + β_1+Expert+β_2 values from Columns (3)-(6), respectively, of Table 1. Columns (7)-(9) report weighted averages of the coefficients across Columns (3)-(6) (referred to as composite 'accuracy scores'), using different weightings $\alpha \in \{0.25, 0.5, 0.75\}$ assigned to coefficients for 'anchoring' (Column (3)) relative to coefficients (averaged across) the uncertainty questions (Columns (4)-(6)). Blue highlighted coefficients reflect the top three performing media for each of the 'subjective' measures and the 'accuracy scores'.

Appendix C Part II: Experimental Design details

Table C.1: Share of forecasts close to optimal, by Stage and Year

Stage	**Year 0**		**Year 1**		**Year 2**	
	$\leq$ 25bp	$\leq$ 50bp	$\leq$ 25bp	$\leq$ 50bp	$\leq$ 25bp	$\leq$ 50bp
A	35%	71%	37%	67%	36%	65%
B	64%	88%	68%	90%	70%	90%
C	-	-	30%	64%	41%	65%
D	-	-	55%	84%	58%	84%

Appendix D Part II: Auxiliary Results

Table D.1: Absolute Deviation by Stage, Horizon, and Error Type - Year 2 (no Fixed Effects)

	Dependent variable: Absolute deviation of point forecast from optimal: Year 2								
	Stage B	Stage C				Stage D			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Constant	**0.278*****	**0.522*****	**0.375*****	**0.365*****	**0.406*****	**0.366*****	**0.326*****	**0.329*****	**0.342*****
	(0.046)	(0.064)	(0.061)	(0.063)	(0.064)	(0.048)	(0.049)	(0.049)	(0.050)
$\underline{\beta_1}$									
Fan	**0.031****	−0.012	0.027	0.015	0.014	0.011	0.016	0.011	0.012
	(0.013)	(0.018)	(0.020)	(0.018)	(0.018)	(0.014)	(0.018)	(0.015)	(0.014)
$\underline{\beta_2}$									
Unlucky Error			**0.255*****				**0.071*****		
			(0.024)				(0.020)		
Unlucky Large Error				**0.299*****				**0.065****	
				(0.035)				(0.027)	
Unlucky Large Positive Error					**0.358*****				**0.068***
					(0.053)				(0.040)
$\underline{\beta_3}$									
Fan*Unlucky Error			**−0.076****				−0.009		
			(0.033)				(0.027)		
Fan*Unlucky Large Error				**−0.091****				0.004	
				(0.046)				(0.036)	
Fan*Unlucky Large Positive Error					**−0.169*****				−0.005
					(0.063)				(0.049)
$\underline{\beta_1 + \beta_3}$									
Fan + Fan*Error Type (linear comb.)			**-0.049****	**-0.075****	**-0.155*****		0.006	0.015	0.007
			(0.024)	(0.034)	(0.047)		(0.019)	(0.027)	(0.037)
Participant FE	No	No	No	No	No	No	No	No	No
Controls: Demog. and pre-treatment quiz	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls: Experimental Variation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,598	1,598	1,598	1,598	1,598	1,598	1,598	1,598	1,598
Adjusted R^2	0.021	0.071	0.160	0.133	0.103	0.040	0.053	0.046	0.041

Note: *p<0.1; **p<0.05; ***p<0.01. Cluster-robust standard errors in parentheses, clustered at participant level.

Table D.2: Absolute Deviation by Stage, Horizon, and Error Type - Year 2 - By Fan Width

	Dependent variable: Absolute deviation of point forecast from optimal: Year 2								
	Stage B	Stage C				Stage D			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept (no FE)	**0.277**	**0.522**	**0.375**	**0.365**	**0.406**	**0.365**	**0.324**	**0.327**	**0.341**
β_1^{Narrow}									
Fan Narrow	**0.030***	0.016	0.019	**0.059*****	**0.057****	0.013	−0.008	0.028	0.025
	(0.016)	(0.021)	(0.022)	(0.022)	(0.023)	(0.015)	(0.018)	(0.018)	(0.017)
$\beta_1^{Narrow} + \beta_3^{Narrow}$									
Fan Narrow + Fan Narrow*Error Type			0.012	**−0.115****	**−0.240*****		0.034	−0.031	−0.056
			(0.030)	(0.056)	(0.070)		(0.023)	(0.043)	(0.054)
β_1^{Wide}									
Fan Wide	**0.033****	**−0.039***	0.035	−0.007	0.0001	0.008	**0.040***	0.002	0.012
	(0.016)	(0.022)	(0.024)	(0.024)	(0.023)	(0.017)	(0.022)	(0.021)	(0.018)
$\beta_1^{Wide} + \beta_3^{Wide}$									
Fan Wide + Fan Wide*Error Type			**−0.111*****	**−0.117****	**−0.263*****		−0.022	0.033	−0.011
			(0.030)	(0.059)	(0.077)		(0.023)	(0.045)	(0.059)
Error Type	All	All	Unlucky	Unlucky Big	Unlucky Big ↑	All	Unlucky	Unlucky Big	Unlucky Big ↑
Participant FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls: Experimental Variation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,598	1,598	1,598	1,598	1,598	1,598	1,598	1,598	1,598
Adjusted R^2	0.271	0.301	0.309	0.309	0.319	0.316	0.319	0.316	0.315

Note: *p<0.1; **p<0.05; ***p<0.01. Cluster-robust standard errors in parentheses, clustered at participant level.

Table D.3: Part II: Regression results: Uncertainty - Year 2 (no Fixed Effects)

	Dependent variable:								
	Stage B	Stage C				Stage D			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Constant	**0.790*****	**0.832*****	**0.823*****	**0.818*****	**0.824*****	**0.845*****	**0.815*****	**0.819*****	**0.834*****
	(0.069)	(0.065)	(0.067)	(0.069)	(0.068)	(0.064)	(0.065)	(0.066)	(0.066)
$\underline{\beta_1}$									
Fan	**0.055*****	**0.035***	0.023	0.023	0.012	**0.041***	**0.052***	**0.040***	0.018
	(0.021)	(0.021)	(0.027)	(0.023)	(0.022)	(0.021)	(0.027)	(0.023)	(0.022)
$\underline{\beta_2}$									
Unlucky Error			0.015				**0.053***		
			(0.028)				(0.029)		
Unlucky Large Error				0.014				0.046	
				(0.036)				(0.039)	
Unlucky Large Positive Error					−0.020				−0.010
					(0.050)				(0.053)
$\underline{\beta_3}$									
Fan*Unlucky Error			0.024				−0.023		
			(0.041)				(0.042)		
Fan*Unlucky Large Error				0.049				0.006	
				(0.051)				(0.054)	
Fan*Unlucky Large Positive Error					**0.173*****				**0.166****
					(0.065)				(0.072)
$\underline{\beta_1 + \beta_3}$									
Fan + Fan*Error Type (linear comb.)			0.047	**0.072***	**0.185*****		0.029	0.046	**0.185*****
			(0.029)	(0.041)	(0.056)		(0.030)	(0.042)	(0.057)
Participant FE	No	No	No	No	No	No	No	No	No
Controls: Demog. and pre-treat quiz	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls: Experimental Variation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,598	1,598	1,598	1,598	1,598	1,598	1,598	1,598	1,598
Adjusted R^2	0.028	0.037	0.037	0.038	0.042	0.041	0.042	0.041	0.046

Note: *p<0.1; **p<0.05; ***p<0.01. Cluster-robust standard errors in parentheses, clustered at participant level.
Demographic controls: age, sex, English fluency, financial literacy, economic engagement, quiz performance.

Appendix E Between-Subjects Robustness Check

A potential concern with our within-subject design is that carry-over effects between runs may differentially impact Point and Fan treatments. Each participant sees one treatment in Run 1 and the other in Run 2. A participant who sees a Fan Chart first learns that forecasts are inherently uncertain. When the participant then encounters a Point Forecast in Run 2, he or she may treat it differently than someone seeing a Point Forecast for the first time by effectively "self-imposing" uncertainty around the point estimate. The reverse carry-over (Point first, Fan second) is less problematic, since the Fan Chart is self-contained and informative regardless of prior experience. That is, the respondent who sees a Point first is not likely to self-impose more certainty around the forecast than a person seeing a density forecast first. If this asymmetry *is* present, participants who see Point Forecasts in Run 2 would behave more like Fan-treated participants, compressing the estimated treatment difference in the pooled sample.

To address this concern directly, we re-estimate our main specification in Table 3 using only Run 1 data, restricting the sample to a pure between-subjects comparison. In Run 1, no participant has been exposed to the other treatment, so carry-over is impossible by construction. Since each participant contributes a single observation, we replace participant fixed effects with demographic controls (age, sex, English fluency, financial literacy, economic engagement, and quiz performance) and retain controls for experimental variation (data outturn, backhistory). The between-subjects sample contains 799 observations.

Table E.1 reports the results. At Stage B (Column 1), the initial anchoring advantage of point forecasts is confirmed: the Fan coefficient is positive and significant ($\hat{\beta}_1 = 0.051$, $p < 0.01$).

At Stage C (Columns 2–5), the insurance effect of fan charts holds. The interaction $\hat{\beta}_3$ is negative for all unlucky error types, and is statistically significant for unlucky large errors ($\hat{\beta}_3 = -0.111$, $p < 0.10$) and unlucky large positive errors ($\hat{\beta}_3 = -0.201$, $p < 0.05$). The linear combination $\hat{\beta}_1 + \hat{\beta}_3$ for the most damaging errors is -0.151 ($p < 0.05$), confirming that Fan-treated participants remain substantially better anchored than Point-treated participants following large unlucky errors, even with zero prior treatment exposure.

At Stage D (Columns 6–9), the pattern is again consistent with the main results: $\hat{\beta}_3$ is negative throughout but not statistically significant, indicating that the updated central bank forecast re-anchors both groups.

As expected, standard errors are larger in the between-subjects sample, and some coefficients that are significant in the pooled analysis lose significance here. This reflects the halving of

the sample and the replacement of participant fixed effects with observable controls. The goal of this exercise is not to achieve identical p-values, but to confirm that the qualitative story holds in a setting where carry-over is ruled out by construction. It does: the insurance effect of fan charts is present even when no participant has been previously exposed to the other treatment.

Table E.1: Between-Subjects Robustness: Absolute Deviation by Stage and Error Type — Year 2 (Run 1 Only)

	Dependent variable: Absolute deviation of point forecast from optimal: Year 2								
	Stage B	Stage C				Stage D			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Constant	**0.310*****	**0.574*****	**0.424*****	**0.398*****	**0.436*****	**0.401*****	**0.342*****	**0.346*****	**0.363*****
	(0.076)	(0.096)	(0.091)	(0.097)	(0.099)	(0.071)	(0.070)	(0.073)	(0.075)
β_1									
Fan	**0.051*****	0.021	**0.044**	**0.055****	**0.050***	**0.037***	**0.063****	**0.044****	**0.046****
	(0.019)	(0.025)	(0.028)	(0.026)	(0.026)	(0.020)	(0.026)	(0.021)	(0.021)
β_2									
Unlucky Error			**0.242*****				**0.097*****		
			(0.034)				(0.028)		
Unlucky Large Error				**0.311*****				**0.094****	
				(0.049)				(0.040)	
Unlucky Large Positive Error					**0.389*****				**0.108***
					(0.076)				(0.059)
β_3									
Fan*Unlucky Error			−0.047				−0.051		
			(0.049)				(0.041)		
Fan*Unlucky Large Error				**−0.111***				−0.021	
				(0.065)				(0.054)	
Fan*Unlucky Large Positive Error					**−0.201****				−0.063
					(0.088)				(0.074)
$\beta_1 + \beta_3$									
Fan + Fan*Error Type (linear comb.)			−0.003	−0.056	**−0.151****		0.012	0.023	−0.017
			(0.034)	(0.048)	(0.066)		(0.029)	(0.040)	(0.054)
Participant FE	No	No	No	No	No	No	No	No	No
Controls: Demog. and pre-treatment quiz	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls: Experimental Variation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	799	799	799	799	799	799	799	799	799
Adjusted R^2	0.026	0.063	0.149	0.125	0.099	0.047	0.061	0.055	0.050

Notes: *p<0.1; **p<0.05; ***p<0.01. Cluster-robust standard errors in parentheses, clustered at participant level.
Run 1 only (between-subjects). Demographic controls: age, sex, English fluency, financial literacy, economic engagement, quiz performance.