

Duffy, John; Hanaki, Nobuyuki; Yoo, Donghoon

Working Paper

Overreaction in expectations under signal extraction: Experimental evidence

ISER Discussion Paper, No. 1293

Provided in Cooperation with:

The Institute of Social and Economic Research (ISER), Osaka University

Suggested Citation: Duffy, John; Hanaki, Nobuyuki; Yoo, Donghoon (2025) : Overreaction in expectations under signal extraction: Experimental evidence, ISER Discussion Paper, No. 1293, Osaka University, Institute of Social and Economic Research (ISER), Osaka

This Version is available at:

<https://hdl.handle.net/10419/331479>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

**OVERREACTION IN EXPECTATIONS
UNDER SIGNAL EXTRACTION:
EXPERIMENTAL EVIDENCE**

John Duffy
Nobuyuki Hanaki
Donghoon Yoo

October 2025

The Institute of Social and Economic Research
The University of Osaka
6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

Overreaction in Expectations under Signal Extraction: Experimental Evidence

John Duffy*

Nobuyuki Hanaki[†]

Donghoon Yoo[‡]

October 4, 2025

Abstract

We experimentally evaluate three behavioral models of expectation formation that predict overreaction to new information: overconfidence in private signals, misperceptions about the persistence of the data-generating process (DGP), and diagnostic expectations. In our main experiment, participants repeatedly forecast the contemporaneous and one-step-ahead values of a random variable. They are incentivized for accuracy, informed of the exact DGP and its past history, and provided with noisy signals about the unobserved contemporaneous value. One treatment features a persistent AR(1) process, while another has no persistence. We also report on an experiment with no noisy signals. At the individual level, we find systematic overreaction even when the DGP is not persistent and regardless of whether a signal-extraction problem is present. By contrast, consensus (mean) forecasts exhibit underreaction, consistent with evidence from other studies. Overall, our results indicate that misperceptions about persistence provide the most compelling explanation for the observed patterns of expectation formation.

Keywords: Expectation formation; Overreaction; Signal extraction; Diagnostic expectations; Misperception; Overconfidence.

JEL Codes: C91, D03, D84, G41.

*Department of Economics, UC Irvine; Institute of Social and Economic Research, the University of Osaka. duffy@uci.edu

[†]Institute of Social and Economic Research, the University of Osaka; University of Limassol. nobuyuki.hanaki@iser.osaka-u.ac.jp

[‡]Department of Economics, Korea University. donghoonyoo@korea.ac.kr

Acknowledgments: We thank Wing Suen, Robert Waldmann, and participants at Academia Sinica; Sogang Summer Research Workshop; IAER Macroeconomics Workshop; Sungkyunkwan University; the 1st Asian Conference on Expectations in Macroeconomics; and ISER for helpful comments. Financial support: Korea University Grant (K2513021), JSPS KAKENHI (20H05631, 21H04397, 23H00055, 24K00238), and the Joint Usage/Research Center at ISER, Osaka University. The experiment was approved by the IRB of ISER (No. 20231201). We thank Hyunzi Oh for excellent research assistance; Yoo gratefully acknowledges the hospitality of ISER and Eesti Pank. All remaining errors are our own.

1 Introduction

A growing body of research highlights systematic ways in which individuals misinterpret information when forming expectations, posing a challenge to the rational expectations (RE) paradigm. According to RE, individuals process information optimally to form unbiased beliefs about the future, and it is typically assumed that the underlying state of the economy is commonly known. With the increasing availability of survey data on expectations—from households, firm managers, and professional forecasters—the RE hypothesis and alternative behavioral models of belief formation have come under closer empirical scrutiny.

Building on the work of Coibion and Gorodnichenko (2015), one strategy for testing the full information rational expectations (FIRE) hypothesis using survey data is to examine the predictability of forecast errors from forecast revisions. Under the FIRE hypothesis, the forecast error should be unpredictable and uncorrelated with forecast revisions. A positive correlation between forecast errors and revisions suggests that forecasts *underreact* to new information relative to the FIRE benchmark, leading to higher realizations relative to the forecasts. On the other hand, a negative correlation implies that forecasts *overreact* to new information, predicting lower realizations than the forecasts.

Based on this framework, various studies have documented deviations from RE. For example, Coibion and Gorodnichenko (2015) show evidence of underreaction in mean (consensus) forecasts based on the Survey of Professional Forecasters (SPF), while Bordalo et al. (2020) provide evidence of overreaction in *individual* level forecasts using both the SPF and the Blue Chip Survey. Kohlhas and Walther (2021) highlight cases where even aggregate beliefs exhibit overreaction. Angeletos et al. (2020) show that the dynamics of beliefs exhibit underreaction in the short run and overreaction in the medium run. Adam et al. (2024) show that overconfidence in private information can account for SPF forecasters' tendency to overreact to residual information while remaining overly anchored to prior beliefs.

These empirical studies have generated various models to explain deviations from the FIRE hypothesis—particularly the tendency for beliefs to overreact to new information—with

various behavioral assumptions, including (1) diagnostic expectations (Bordalo et al., 2018; Bordalo et al., 2019; Bianchi et al., 2024a; L’Huillier et al., 2024; Bianchi et al., 2024b; Maxted, 2024; Na and Yoo, 2025); (2) belief misspecification, in particular, mis-perceptions of the persistence of the underlying data generating process (Angeletos et al., 2020; Candian and Leo, 2023; Falato and Xiao, 2024) and (3) over-confidence in private information (Adam et al., 2024).⁰ While each approach can account for departures from FIRE, to the best of our knowledge, a rigorous empirical comparisons across these three frameworks does not yet exist.

Thus, in this paper we pose the following research question: among three behavioral models of expectation formation—(1) diagnostic expectations, (2) over-persistence, defined as the mistaken overestimation of the persistence of the underlying data-generating process (DGP), and (3) over-confidence in private information—which best explains the observed deviations from the FIRE hypothesis, in particular the tendency of individual beliefs to overreact to new information? We address this question by conducting a set of controlled laboratory experiments.

While survey data on expectations have been widely studied, there are several limitations to using such data for rigorous tests of behavioral models. For example, the predictability of forecast errors might simply reflect imperfect knowledge of the true DGP for the macroeconomic and financial variables being forecast. FIRE assumes that agents have a correctly specified model of the DGP (structure and parameters). Since the extent of overreaction or underreaction as characterized by the error-revision coefficient is best understood in relation to the DGP, an incomplete understanding of the true DGP means that forecast predictability could be due to model misspecification. Moreover, econometricians do not fully observe the information sets that are available to forecasters. As a result, distinguishing between the effects of information frictions and purely behavioral or cognitive biases in survey data is challenging. Finally, survey respondents are typically not directly incentivized to submit accurate forecasts; participants may answer in a haphazard manner or in a way that they

⁰In contrast, level-k thinking (Farhi and Werning, 2019; García-Schmidt and Woodford, 2019) and cognitive discounting (Gabaix, 2020; Hajdini, 2023) inherently lead to belief underreaction.

believe is more socially acceptable or favorable, rather than how they truly believe.

Controlled laboratory experiments allow researchers to overcome these limitations of survey-based evidence. In particular, in a controlled experiment, the researcher can specify both the underlying DGP and the forecasters’ information sets, enabling precise measurement of biases in expectation formation. Furthermore, our experimental participants are incentivized in such a way that their payoffs are maximized by being as accurate as possible, thus giving the theory its best chance. Finally, experiments also enable repetition of the same task by the same participants, each of whom faces an identical, randomly drawn sequence of data to forecast. This design yields many well-measured observations in a stable, controlled environment.

In each period t of our experiment, participants provide forecasts for the realization of a random variable in two consecutive periods, the current period t and the next period $t + 1$. This design enables us to observe how participants revise their forecasts over time. In order to minimize the possibility of model misspecification being the cause of departures from FIRE, participants are fully informed of the DGP, which is modeled as an AR(1) process. Specifically, they are told the functional form of the DGP and its associated parameter values. Following the theoretical literature, we use a signal extraction environment in which subjects receive a noisy signal about the actual value of the unobserved random variable before submitting their forecasts.¹

There are two main treatments. In the first, the data-generating process is persistent, with persistence parameter $\rho = 0.6$. In the second, it is non-persistent, corresponding to $\rho = 0$. As discussed in Section 2, models of belief overreaction generate distinct predictions contingent on the persistence of the underlying DGP. Under diagnostic expectations (DE), beliefs overreact only when the DGP is persistent. By contrast, both overestimating the persistence of the DGP (“over-persistence”) and overconfidence in private signals (“overconfidence”) predict

¹The signal extraction environment is widely implemented in studying expectations formations. See, for example, Coibion and Gorodnichenko (2015), Bordalo et al. (2020), Angeletos et al. (2020), Adam et al. (2024), among others. Bao and Duffy (2021) provide experimental evidence on the extent to which individuals can solve signal extraction problems.

overreaction regardless of whether the DGP is persistent or not. Moreover, by estimating the degree of perceived persistence from forecasts, along with the weight placed on signals in the signal extraction model, we can distinguish between over-persistence and overconfidence. Thus, comparing behavior across these two treatments allows us to identify the specific mechanism driving overreaction in the data.

Our paper contributes to a growing literature that uses laboratory experiments to study expectation formation; see Assenza et al. (2014) for a survey. We are most closely related to two recent papers that apply the Coibion and Gorodnichenko (2015) methodology to examine over- and underreaction in forecasts of externally generated time series variables. The first is Afrouzi et al. (2023), who study how people revise their forecasts of a variable following an AR(1) process and show that overreaction is common across a wide range of treatments and parameter values. While their design provides valuable evidence of systematic overreaction, our approach differs from theirs in two important respects. First, participants in our experiment are explicitly informed of the exact DGP, including parameter values, whereas Afrouzi et al. (2023) do not provide this level of detail.² Second, we frame the forecasting problem as one of *signal extraction*: participants observe both past realizations of the variable and noisy signals about its future value, whereas participants in Afrouzi et al. (2023) observe only past realizations. To isolate the role of noisy signals, we also report results from an additional experiment where participants are informed of the exact DGP but do not receive noisy signals, while keeping all other features fixed.

The second related study is He and Kučinskas (2023), who report online experiments on forecasting behavior, showing that when information is split across two correlated time series, subjects neglect the correlation and therefore underreact. By contrast, our study focuses on a univariate signal-extraction environment, where we seek to identify the behavioral forces—diagnostic expectations, overpersistence, or overconfidence—that drive the systematic overreaction that we observe in our setting.

²In one treatment of Afrouzi et al. (2023), participants are told that the DGP is an AR(1) process, but parameter values remain undisclosed.

Our results shed new light on the process of expectation formation. First, we find systematic deviations from FIRE, regardless of whether participants face a signal extraction problem or not. Second, overreaction occurs even when the DGP is non-persistent, contrary to the prediction of DE. Third, we estimate the (subjective) persistence and mean implied by subjects' forecasts. Most subjects over-estimate the persistence of the underlying DGP, although because they also misperceive the mean of the process, their estimated long-run mean, as determined by both parameters, closely aligns with the actual value. Fourth, the degree of over-confidence, computed from the estimated weight on the signal based on a signal extraction model and the estimated subjective persistence of the process, shows that over-persistence, and not over-confidence, is the model that best explains the observed belief overreaction of most of the participants. (Among those exhibiting belief overreaction, 80% are driven by over-persistence, either alone or in combination with over-confidence, whereas only 11% are driven by over-confidence alone.) The result of the additional experiment without signal extraction is consistent with the conclusion we drew from the main experiment. Note that by construction, over-confidence does not play a role in the absence of a private signal. Yet, we continue to observe belief overreaction.

Furthermore, comparison of our experiments with and without signal extraction brings us two additional insights. First, when subjects form expectations with signal extraction, forecasts exhibit a similar degree of overreaction regardless of whether the underlying process is i.i.d. or persistent. However, when subjects do not receive signals, overreaction is greater when the underlying process is less persistent. The latter result aligns with the findings of Afrouzi et al. (2023). Second, we observe a divergence between individual and consensus forecast responses to new information only in the experiment with private signals. Specifically, with signal extraction, individual forecasts tend to overreact, while the consensus forecast (mean or median) underreacts. This finding supports the hypothesis of Coibion and Gorodnichenko (2015), whereby noisy private signals generate information frictions that suppress the responsiveness of the consensus forecast.

The rest of the paper is organized as follows. Section 2 reviews theories of belief overreaction and provides hypotheses that we test in our experiment. Section 3 describes the design and procedures followed in our signal extraction experiment while section 4 presents the main findings from the signal extraction experiment. Section 5 reports results from an additional experiment without signal extraction and examines the role of information frictions in generating underreaction in consensus forecasts. Finally, section 6 concludes.

2 Theory and Hypotheses

In this section, we present a simple forecasting model following Coibion and Gorodnichenko (2015), Angeletos et al. (2020) and Bordalo et al. (2020), in which forecasters are fully informed about the structure and parameters of the model and continuously update their information sets, but do not observe the underlying state directly.³ We begin with the benchmark case in which forecasters form rational expectations and then introduce three behavioral deviations from this benchmark: *overconfidence*, *overpersistence*, and *diagnostic expectations*. We conclude by outlining the distinct predictions of the rational and behavioral models, which we evaluate experimentally to determine which, if any, best explains the observed forecasting behavior.

Suppose each forecaster, indexed by $i \in [0, 1]$, forms expectations about a random variable x_t , which is known to evolve according to the data-generating process (DGP)

$$x_t = \mu + \rho x_{t-1} + \varepsilon_t, \tag{1}$$

where μ denotes a constant, $\rho \in [0, 1)$ is the persistence parameter, and ε_t is an i.i.d. Normal shock with mean zero and variance σ_ε^2 .

Before forming forecasts of x_t , each agent i receives a noisy private signal s_t^i about the

³Throughout the paper, we use the terms *forecasters*, *individuals* and *agents* interchangeably.

realized but unobserved value of x_t :

$$s_t^i = x_t + \omega_t^i, \quad (2)$$

where ω_t^i denotes an idiosyncratic, mean-zero, i.i.d. Normal shock with variance σ_ω^2 . Let $F_t^i x_t$ be agent i 's belief (forecast) about the unobserved fundamental x_t based on information available at time t . This information set includes the noisy signal, s_t^i , all past realizations of the random variable $\{x_k\}_{k=1}^{t-1}$ and the agents' own prior belief about x_t , $F_{t-1}^i x_t$.

Rational Bayesian learning. Suppose that agent i updates their beliefs rationally. Using the Kalman filter technique, their forecast $F_t^i x_t$, evolves according to

$$F_t^i x_t = F_{t-1}^i x_t + \lambda (s_t^i - F_{t-1}^i x_t), \quad (3)$$

where $\lambda \in (0, 1]$ is the optimal Kalman gain,

$$\lambda = \frac{\delta - (1 - \rho^2) + \delta A}{2 + \delta - (1 - \rho^2) + \delta A}, \quad (4)$$

with $A = \sqrt{\frac{(1-\rho^2)^2}{\delta^2} + 1 + \frac{2}{\delta} + \frac{2\rho^2}{\delta}}$. Notice that the gain term λ depends on both the persistence parameter ρ and the signal-to-noise ratio, defined as $\delta \equiv \sigma_\varepsilon^2 / \sigma_\omega^2$. As shown in Appendix A.1, λ increases with δ , meaning that more precise signals (i.e., those with less noise) lead to a higher Kalman gain. Moreover, λ is increasing in ρ as well: the more persistent the underlying process, the more weight is assigned to the incoming signals.

2.1 Theories of overreaction

Individual rationality requires agents to make the best possible use of the available information, as described by eq. (3). We now examine some behavioral models that characterize deviations from rationality leading to an overreaction of expectations.

Overconfidence. Suppose agent i misspecifies the signal-to-noise ratio to be $\hat{\delta}$, where $\hat{\delta} \neq \delta$. Overconfidence corresponds to case where $\hat{\delta} > \delta$. As the weight on observed signals λ increases with perceived $\hat{\delta}$, overconfidence results in excessive reliance on signals relative to rational Bayesian updating.

Overpersistence. Suppose agent i misperceives the persistence of the random variable, so that $\hat{\rho} \neq \rho$. By “overpersistence,” we mean the case where the agent overestimates the degree of persistence in x_t , that is, $\hat{\rho} > \rho$. Consequently, the agent places too much weight on new information relative to a rational Bayesian updater, since the signal weight λ increases with perceived persistence. $\hat{\rho}$.

Diagnostic expectations. The theory of diagnostic expectations (DE), based on the “representativeness heuristic” of Kahneman and Tversky (1972), is formalized by Gennaioli and Shleifer (2010) and Bordalo et al. (2016). In this framework, an agent overweights the probability of states that become more likely upon receiving new information.⁴

In our experiment, the process for x_t follows an AR(1) with normally distributed shocks, $\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. Under diagnostic expectations, agents correctly perceive the variance σ_ε^2 but distort the conditional mean:

$$\mathbb{E}_t^\theta[x_{t+1}] = \mathbb{E}_t[x_{t+1}] + \theta \left(\mathbb{E}_t[x_{t+1}] - \mathbb{E}_{t-1}[x_{t+1}] \right),$$

where $\mathbb{E}_t^\theta[\cdot]$ denotes the diagnostic expectations operator and $\theta \geq 0$ is the diagnosticity parameter governing the degree of overreaction.⁵ The term $\left(\mathbb{E}_t[x_{t+1}] - \mathbb{E}_{t-1}[x_{t+1}] \right)$ represents the news (the revision in rational expectations based on information available at time t). If this revision is positive, signifying “good news,” diagnostic expectations become excessively optimistic. Conversely, if the revision is negative, indicating “bad news,” expectations become

⁴See Bordalo et al. (2018, Proposition 1, p. 209) for details; verify pagination for the version cited.

⁵We assume that the reference distribution is based on the $t-1$ information set. However, selective recall may involve more distant information sets, as formally modeled in Bianchi et al. (2024a).

excessively pessimistic.

Bordalo et al. (2020) extend this theory to a framework of learning from noisy signals as in our experimental setup. Let $F_t^{i,\theta} x_t$ denote diagnostic agent i 's forecast of x_t at time t . As shown in Appendix A.2 under DE, the dynamics of $F_t^{i,\theta} x_t$ are given by:

$$F_t^{i,\theta} x_t = F_{t-1}^i x_t + (1 + \theta) \lambda (s_t^i - F_{t-1}^i x_t), \quad (5)$$

where $\lambda \in (0, 1]$ is the optimal Kalman gain computed from the true variance and $\theta \geq 0$ captures the departure from rational updating due to representativeness. When $\theta = 0$, the diagnostic Kalman filter (5) reduces to the standard Kalman filter (3), i.e., rational Bayesian updating.

2.2 Hypotheses

We derive two hypotheses based on the models presented above. Let the forecast error (FE_t^i) and the forecast revision (FR_t^i) for agent i in period t be:

$$FE_t^i = x_t - F_t^i x_t, \quad (6)$$

$$FR_t^i = F_t^i x_t - F_{t-1}^i x_t. \quad (7)$$

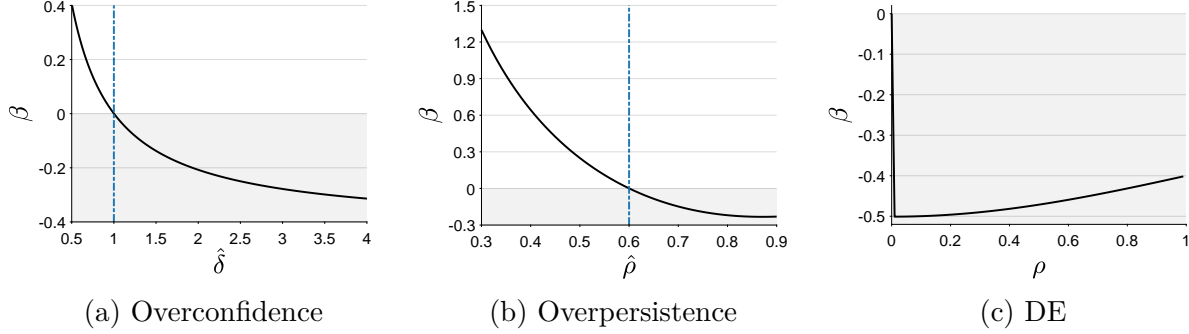
The first hypothesis is about the coefficient β in the the following regression that projects forecast errors onto forecast revisions at the individual level:

$$FE_t^i = \alpha + \beta \times FR_t^i + u_t^i. \quad (8)$$

Angeletos et al. (2020, Proposition 1, p.15) show that β , under overconfidence and overpersistence, is given by

$$\beta = -\kappa_1(\hat{\delta}^{-1} - \delta^{-1}) - \kappa_2(\hat{\rho} - \rho), \quad (9)$$

Figure 1: Forecast Revision Coefficients in Overconfidence, Overpersistence, and DE



Notes: In Panel (a) forecast revisions coefficients are given as a function of perceived signal precision, $\hat{\delta}$. The parameters δ and ρ are set at 1 and 0.6, respectively. In Panel (b) forecast revisions coefficients are given as a function of perceived persistence $\hat{\rho}$. There, the parameters δ and ρ are again set at 1 and 0.6, respectively. In Panel (c) forecast revision coefficients are given as a function of the persistence parameter ρ . We set both parameters δ and θ equal to 1.

where $\kappa_1 = \frac{1}{\mathcal{V}_{ind}}(\hat{\rho} - \hat{\Lambda})\frac{\hat{\Lambda}}{\delta(1 - \hat{\Lambda}^2)}$, $\kappa_2 = \frac{1}{\mathcal{V}_{ind}}(\hat{\rho} - \hat{\Lambda})\frac{(1 + \hat{\Lambda}^2)(1 - \rho^2) + (\hat{\Lambda} + \rho)(\rho - \hat{\rho})}{(1 - \hat{\Lambda}^2)(1 - \rho^2)(1 - \hat{\Lambda}\rho)}$, \mathcal{V}_{ind} denotes the variance of FR_t^i , and $\hat{\Lambda} = \frac{1}{2} \left(\hat{\rho} + \frac{1 + \hat{\delta}}{\hat{\rho}} + \sqrt{\left(\hat{\rho} + \frac{1 + \hat{\delta}}{\hat{\rho}} \right)^2 - 4} \right)$.⁶

One can clearly see from eq. (9) that when $\hat{\delta} = \delta$ and $\hat{\rho} = \rho$, then $\beta = 0$. Figure 1 illustrates the relationship between the forecast revision coefficient, β , and the degree of overconfidence, $\hat{\delta}$ (Panel a), as well as the degree of overpersistence, $\hat{\rho}$ (Panel b). Note that a smaller perceived variance for the noisy signal, $\hat{\sigma}_\omega^2$ relative to the true value σ_ω^2 reflects overconfidence in the private signal, implying $\hat{\delta} > \delta$. As shown in Panel (a), when $\hat{\delta} > \delta$ (i.e., overconfidence), the resulting forecast revision coefficient satisfies $\beta < 0$. Likewise, Panel (b) shows that when $\hat{\rho} > \rho$ (i.e., overpersistence), β also becomes negative.

Under diagnostic expectations (DE), the diagnosticity parameter θ distorts Bayesian updating by shifting the conditional predictive distribution toward states that become more likely given the signal (representativeness). When $\theta = 0$, the updating process conforms to rational expectations characterized by the Bayesian conditional probability distribution. Conversely for $\theta > 0$. DE inflates the probabilities of “representative” future states and deflates “unrepresentative” ones, relative to the Bayesian benchmark. Crucially, this mechanism is

⁶We note that $\kappa_1 > 0$ and $\kappa_2 > 0$ since $\hat{\rho} > \hat{\Lambda}$.

forward-looking and depends on how news changes the likelihood of future outcomes. In an AR(1) environment, news about x_t from the signal s_t^i shifts the predictive distribution of x_{t+1} only if $\rho > 0$. Hence, DE predicts forecast overreaction only when the process is persistent. If the process is non-persistent ($\rho = 0$), the predictive distribution of x_{t+1} does not vary with current news; all future states are equally likely given today's signal, so the DE shift is inoperative and the model collapses to Bayesian updating. In the context of our individual forecast-error specification (8), we have, under DE:

$$\beta = \begin{cases} -\frac{\theta(1+\theta)}{(1+\theta)^2 + \theta^2\rho^2} & \text{if } \rho > 0, \\ 0 & \text{if } \rho = 0. \end{cases}$$

Thus, if the random variable exhibits persistence and agents have diagnostic expectations so that $\theta \neq 0$, then we should find that $\beta < 0$ when $\theta > 0$ and $\beta > 0$ when $\theta < 0$, respectively. If the random variable does not exhibit any persistence ($\rho = 0$), then current news provides no information about future deviations from the mean.

The preceding discussion leads to our first hypothesis:

Hypothesis 1 (Forecast revision coefficient β).

- If agents are Bayesian, then $\beta = 0$.
- If agents have diagnostic expectations and the fundamental is i.i.d. ($\rho = 0$), then $\beta = 0$.
- In all other non-Bayesian settings, $\beta < 0$ indicates overreaction and $\beta > 0$ indicates underreaction.

Table 1 (first two rows) summarizes the different predictions of the four models considered above. When agents update their beliefs rationally (i.e., under Bayesian learning), $\beta = 0$. In contrast, theories of overconfidence and overpersistence predict $\beta < 0$, as agents give greater weight to new information than under Bayesian learning, leading to overreaction. A similar

prediction applies to diagnostic agents when the random variable is persistent. However, when the fundamental lacks persistence ($\rho = 0$), diagnostic agents do not overreact to new information, implying $\beta = 0$.

Table 1: Summary of Theoretical Predictions

	Bayesian	Overconfidence	Overpersistence	Diagnostic Expectations (DE)
β ($\rho = 0$)	0	< 0	< 0	0
β ($\rho > 0$)	0	< 0	< 0	< 0
$\hat{\rho}$	ρ	ρ	$\hat{\rho} > \rho$	ρ

Notes: With DE, the behavioral parameter θ is assumed to be strictly positive, $\theta > 0$.

Suppose that agent i makes two forecasts in period t : $F_t^i x_t$ and $F_t^i x_{t+1}$, as is the case in our experiment. Since no additional information is provided, it is reasonable to assume that the forecast $F_t^i x_{t+1}$ is related to $F_t^i x_t$ by:

$$F_t^i x_{t+1} = \hat{\mu} + \hat{\rho} F_t^i x_t, \quad (10)$$

where $\hat{\mu}$ and $\hat{\rho}$ represent the subjective mean and persistence, respectively. As noted above, overpersistence implies $\hat{\rho} > \rho$. By regressing $F_t^i x_{t+1}$ on $F_t^i x_t$, one can directly test this model of overpersistence. We summarize our hypothesis regarding subjective persistence in our second hypothesis:

Hypothesis 2 (Forecast-implied subjective persistence $\hat{\rho}$).

- If agents are Bayesian, overconfident or have diagnostic expectations, then $\hat{\rho} = \rho$.
- If agents exhibit overpersistence, then $\hat{\rho} > \rho$.

Table 1 (third row) summarizes these model-specific predictions for $\hat{\rho}$. Under Bayesian learning, agents correctly perceive the persistence of the fundamental, so $\hat{\rho} = \rho$. Overconfident agents, while misperceiving the precision of the signal, also have correct beliefs about persistence, implying $\hat{\rho} = \rho$. Similarly, under diagnostic expectations, the weight placed on signals is

biased, but there is no bias in the persistence parameter. Only the overpersistence model predicts a distorted belief about persistence, with $\hat{\rho} > \rho$, which can generate biased forecasts even without new information.

3 Experimental Design

Our experiment closely follows the theoretical setting presented in the previous section. Participants repeatedly forecast (over 40 periods), future values of a random variable that was known to follow the stochastic process $x_t = \mu + \rho x_{t-1} + \varepsilon_t$ with $\sigma_\varepsilon = 3$. In each period t , participants submit their forecasts for x_t and x_{t+1} , denoted by $F_t^i x_t$ and $F_t^i x_{t+1}$. Before submitting these forecasts, each participant also received a noisy signal s_t^i about x_t defined by $s_t^i = x_t + \omega_t^i$ with $\sigma_\omega = 3$. Furthermore, participants could observe realized x and s for the past 40 periods both in the form of a time series graph and a table on their screen.

We considered two parameterizations or treatments, $\rho = 0$ and $\rho = 0.6$, with μ being set so that in both treatments, the long-run mean value for x , $\mu/(1 - \rho) = 75$. That is, for the $\rho = 0$ treatment, we set $\mu = 75$ and for the $\rho = 0.6$ treatment, we set $\mu = 30$. Participants were provided with the exact formula for the data generating process for x_t (eq. (1)) and the noisy signal process s_t^i (eq. (2)) along with the values of the model's parameters: μ , ρ , σ_ω , and σ_ε .⁷ Our intent in providing this information was to minimize the effects of model misspecification, such as overconfidence or overpersistence. The formulas were carefully explained in the instructions and they were also displayed on subjects' screens during each period. In both treatments, each participant faced the same sequence of realizations for x_t , but received their own private signal s_t^i in every period, meaning they saw a different signal realization than others in the same treatment. Our use of multiple subjects receiving different signals but all facing the same DGP and forecasting task, enables us to compare individual forecasts with consensus (mean) forecasts.

Participants were paid according to the accuracy of one randomly chosen forecast for

⁷Furthermore, we explained the distribution of ε and ω graphically in the instructions as well.

x_j ($F_j^i x_j$) as well as one randomly chosen forecast for x_{k+1} ($F_k^i x_{k+1}$) where j and k were randomly chosen periods. Specifically, participant i 's payoff in experimental currency units (ECUs) is given by

$$\pi_{j,k}^i = \frac{100}{1 + |x_j - F_j^i x_j|} + \frac{100}{1 + |x_{k+1} - F_k^i x_{k+1}|}. \quad (11)$$

Participants were given instructions on their screen detailing the tasks they would perform in each of the 40 periods (see Appendix C for an English translation). The instructions clarified the concept of a random variable drawn from a normal distribution, as used in the experiment. To ensure that subjects had a good understanding of the tasks they faced, subjects were required to answer control questions based on the instructions which included questions on the data generating process, namely, the mean and the variance of ϵ and ω , as well as correctly deriving x_{t+1} conditional on the given values of x_t ($=20$) and ϵ ($=0$). Only after all control questions were correctly answered could subjects begin the experiment.

The online experiments were conducted at the University of Osaka in June 2024.⁸ A total of 62 subjects (31 for each treatment) were recruited through ORSEE (Greiner, 2015). The subjects are undergraduate and graduate students of the University of Osaka. Participants received an e-mail with the link to the experimental site, on which they were randomly allocated to different treatments. They participated in the experiment using their own device (only computers and tablets were allowed) from the location of their choice, provided that it had a good internet connection and they could concentrate on the experiment. Once they had completed all tasks, they were informed of their payoff and could exit the study. Each subject's total earnings in ECU were converted into Japanese yen at a known, predetermined exchange rate of 1 ECU to 15 Japanese yen. On average, subjects earned 1710 Japanese yen (JPY) including a show-up payment of 500 JPY. They were paid using Amazon Gift Cards (e-mail version) on the same day of the experiment.

⁸The experiments were computerized and programmed in oTree (Chen et al., 2016).

4 Results

We first investigate Hypothesis 1 regarding the relationship between forecast errors and forecast revisions. We then proceed to investigate Hypothesis 2 regarding individuals' perceived parameter values for the DGP with the aim of understanding which model might best explain our data.

4.1 Overreaction to new information in individual forecasts

As noted in Section 2.2, we analyze forecast error predictability at the individual level by adapting the regression in eq. (8). Specifically, we conduct participant-by-participant regressions to obtain a distribution of individual coefficients β^i , $i = 1, \dots, I$ and we also estimate a common coefficient β^p by pooling all participants in a given treatment and including participant fixed effects.

Table 2 reports estimates of the individual level regressions in the two treatments. In both treatments, we find a negative relationship between forecast revisions and subsequent forecast errors: $\beta^p < 0$. The estimated β^p is negative and statistically significant at the 1 percent level—see Columns (1) and (3). Moreover, the median of the estimated β^i s in the participant-by-participant regressions are negative in all treatment groups (see Columns (2) and (4)). Specifically, the estimated β^i is negative for 26 out of 29 participants when $\rho = 0$ and for 28 out of 30 participants when $\rho = 0.6$. (Figure B.2 in Appendix B shows the distributions of the individual-level regression coefficient β^i in each treatment. The evidence clearly supports the notion that $\beta^i < 0$. This finding suggests that expectations tend to overreact at the individual level, consistent with the overreaction observed in macroeconomic and financial forecasts from Bordalo et al. (2020).

These results reject not only the Bayesian learning model but also the diagnostic expectations (DE) model, which predicts $\beta = 0$ when $\rho = 0$. Thus, of the four models considered in Section 2, only the overconfidence and overpersistence models align with our experimental

Table 2: Predicting Forecast Errors with Forecast Revisions

	$\rho = 0$		$\rho = 0.6$	
	(1)	(2)	(3)	(4)
β^p	−0.314		−0.319	
SE	0.071		0.045	
t -stat	−4.40		−7.04	
β^i (Median)		−0.297		−0.299
R^2	0.161		0.165	
N	1029	1029	1036	1036

Notes: The table reports coefficients from individual-level forecast error on forecast revision regression as in equation (8). Columns (1) and (3) show the estimation results from the individual-level fixed effect panel regression (β^p). Standard errors are clustered by both forecaster and time. Columns (2) and (4) show the median coefficients in forecaster-by-forecaster regressions (β^i). Figure B.2 in Appendix B plots the distribution of the regression coefficient β^i .

findings.

In a related experiment, Afrouzi et al. (2023) show that forecasts feature more over-reaction when the process is more transitory. However, in our experiment, the error revision coefficient β^p (or the median of β^i) is estimated to be very close to each other at around -0.31 (-0.29) for both $\rho = 0$ and $\rho = 0.6$. Due to two key differences between our experiment and that of Afrouzi et al. (2023)—namely, the presence of a noisy signal and the fact that subjects are informed of the exact data-generating process (DGP)—the results from the two studies are not directly comparable. We address this issue in Section 5, where we report results from an additional experiment in which participants do not receive a signal, as in Afrouzi et al. (2023), but are still informed of the exact DGP specification as in our experiment.

4.2 Forecast-implied subjective persistence

We now turn to the forecast-implied subjective persistence to assess whether overpersistence may account for the negative β documented in the previous section. To this end, we estimate the empirical regression specified in eq. (10).

Table 3: Forecast-Implied Subjective Persistence

	$\rho = 0$ ($\mu = 75$)		$\rho = 0.6$ ($\mu = 30$)	
	(1)	(2)	(3)	(4)
$\hat{\rho}^p$	0.539 (0.380, 0.697)		0.678 (0.587, 0.768)	
SE	0.080		0.046	
t-stat	6.68		14.64	
$\hat{\rho}^i$ (Median)		0.455		0.645
$\hat{\mu}^p$	34.5 (22.4, 46.6)		23.8 (17.1, 30.6)	
SE	5.911		3.304	
t-stat	5.84		7.23	
$\hat{\mu}^i$ (Median)		40.5		26.3
$\hat{\mu}^p / (1 - \hat{\rho}^p)$	74.84		73.91	
$\hat{\mu}^i / (1 - \hat{\rho}^i)$		74.31		74.08
N	1063	1063	1069	1069

Notes: Columns (1) and (3) show the estimation results from the individual-level fixed effect panel regression ($\hat{\rho}^p$), clustered S.E. The parentheses indicate a 95% confidence interval. Columns (2) and (4) show the median coefficients in forecaster-by-forecaster regressions ($\hat{\rho}^i$).

Table 3 reports the results for the two treatments: $\rho = 0$ and $\rho = 0.6$. Columns (1) and (3) estimate a common coefficient obtained by pooling participants. In both treatments, the point estimate of $\hat{\rho}$ exceeds the actual persistence ρ , suggesting that *overpersistence*, rather than overconfidence, is the more likely explanation for the negative β .

Columns (2) and (4) report the median of the participant-by-participant estimates of $\hat{\rho}$. Consistently, the median estimated $\hat{\rho}$ is greater than ρ in both treatments. Moreover, the subjective persistence implied by forecasts exceeds the actual persistence for 24 out of 29 participants when $\rho = 0$ and for 21 out of 30 participants when $\rho = 0.6$.

Finally, we note that the forecast-implied subjective long-run mean, $\hat{\mu}/(1 - \hat{\rho})$, is close to the actual long-run mean of 75 in both treatments. However, regardless of the value of the subjective long-run mean, as long as overpersistence prevails ($\hat{\rho} > \rho$), forecasts tend to

overreact to new information, as reflected in the negative β documented earlier.

4.3 Estimating a signal extraction model

In this section, we estimate a signal extraction model to recover the structural parameters governing individuals' belief formation. Our aim is to understand the role played by *overconfidence* ($\hat{\delta} > \delta = 1$), relative to *overpersistence* ($\hat{\rho} > \rho$) in explaining the observed overreaction in subjects's forecasts. To do that, we follow the Bayesian model in assuming that individuals form forecasts as a weighted average of their prior forecast and the new signal. Specifically, agent i 's belief updating is given by

$$F_t^i x_t = (1 - \lambda)F_{t-1}^i x_t + \lambda s_t, \quad (12)$$

where λ represents the weight placed on the signal s_t .⁹ Since subjects observe the previous realization x_{t-1} , their prior forecast should correspond to $\hat{\rho}x_{t-1}$. Accordingly, we estimate the following empirical model:

$$F_t^i x_t = \text{const} + \gamma x_{t-1} + \lambda s_t + u_t^i, \quad (13)$$

where γ captures the persistence of prior forecast and u_t^i denotes the error term.

Given an estimate of $\hat{\lambda}$ and the implied persistence $\hat{\rho}$, we can recover the forecast-implied signal-to-noise ratio $\hat{\delta}$ using the parameter relationship in eq. (4). Suppose for individual forecaster i , the forecaster's revision coefficient β was estimated to be negative in Section 4.1. Disregarding the role of representativeness heuristic in the form of DE (this is especially true for $\rho = 0$), there are three possibilities to generate belief overreaction: (1) overpersistence ($\hat{\rho} > \rho$), (2) overconfidence ($\hat{\delta} > \delta$), and (3) both overconfidence and overpersistence.¹⁰

⁹Note that eq. (12) is just another version of the rational Bayesian updating eq. (3).

¹⁰Given an estimate of $\hat{\lambda}$, we infer the degree of diagnosticity θ using the relation $\hat{\lambda} = (1 + \theta)\lambda$, where λ denotes the optimal Kalman gain. For $\rho = 0.6$, we find that 38% exhibit diagnosticity ($\theta > 0$), with an average value of $\theta = 0.365$.

Table 4: Source of Overreaction

		$\rho = 0$	$\rho = 0.6$
Only overpersistence	$(\hat{\rho} > \rho)$	8 (31%)	15 (52%)
Only overconfidence	$(\hat{\delta} > 1)$	1 (4%)	5 (17%)
Both		16 (61%)	5 (17%)
Neither		1 (4%)	4 (14%)
Total	$(\hat{\beta} < 0)$	26	29

Notes: *Only overpersistence* refers to the case where $\hat{\rho} > \rho$ and $\hat{\delta} \leq 1$. *Only overconfidence* refers to the case where $\hat{\delta} > 1$ and $\hat{\rho} \leq \rho$. *Both overpersistence and overconfidence* describes the case where $\hat{\rho} > \rho$ and $\hat{\delta} > 1$. *Neither* corresponds to the case where neither $\hat{\rho} > \rho$ nor $\hat{\delta} > 1$ holds.

Table 4 shows that overpersistence is the primary driver of belief overreaction indicated by negative β . When $\rho = 0$, 31% exhibit overpersistence alone and 61% both overpersistence and overconfidence; for $\rho = 0.6$, the corresponding shares are 52% and 17%, respectively. Instances of overconfidence alone are rare (4% and 17%), and only a small fraction fall into the ‘neither’ category.

These findings indicate that overpersistence, whether alone or in combination with overconfidence, is the key driver of belief overreaction, whereas overconfidence alone plays a minor role. This pattern suggests that participants systematically overstate the persistence of the underlying process, which amplifies their responsiveness to new signals and results in belief overreaction, as evidenced by negative β .

5 Additional Experiment

The findings in the previous section indicate that individuals tend to overreact to new information. To investigate the underlying sources of this expectation bias, we evaluated which theoretical model of belief formation best fits our experimental data. We find that the predictions of both diagnostic expectations and overconfidence are inconsistent with the observed behavior, leaving overpersistence as the most plausible explanation for belief overreaction.

One key result—namely, that the estimated β is similar when $\rho = 0$ and when $\rho = 0.6$ —contrasts with the findings of Afrouzi et al. (2023), who report that β becomes more negative as the process becomes more transitory (i.e., when ρ is smaller). As discussed earlier, this discrepancy likely stems from differences between the two experimental designs. In our experiment, participants are fully informed of the data-generating process (DGP) and receive a noisy signal of the current realization of the random variable in addition to the history of past outcomes before making forecasts. In contrast, participants in Afrouzi et al. (2023) were neither informed about the DGP nor given a signal.

To isolate the effect of signal noise, we therefore conducted an additional experiment in which participants did not receive a noisy signal but were still fully informed about the DGP, so as to facilitate comparisons with our first experiment. All other features, including parameter values, were held constant across treatments. These additional experiments also serve as a further test of the overconfidence hypothesis: since agents do not face a signal extraction problem in this new experiment, if overreaction is observed, it cannot be attributed to overconfidence.

The online experiment was conducted in parallel with our main experiment on the same day with a *different* set of participants drawn from the same subject pool. A total of 62 subjects (31 each for $\rho = 0$ and $\rho = 0.6$) participated in this additional experiment. On average, participants earned 1536 JPY, including the 500 JPY show-up fee. The comparative static predictions of the rational Bayesian, model, the model with overpersistence, and diagnostic expectations regarding β and ρ_s are the same as those presented in Section 2.2—see Appendix A.3 for the details.

5.1 Results

5.1.1 Informativeness of Signals

We begin by assessing whether the presence of signals improved participants’ forecasting accuracy. To do so, we compare the deviation of subjects’ forecasts from actual realizations in

Table 5: Mean Square Error (MSE)

	$\rho = 0$	$\rho = 0.6$
With Signals	6.49 (0.380)	5.97 (0.286)
Without Signals	7.17 (0.227)	8.04 (0.423)

Notes: The mean squared error for each treatment, denoted MSE^i , is averaged across individuals within each treatment group. Standard errors are reported in parentheses.

the original experiment (with signals) to those in the additional experiment (without signals). Forecast accuracy is measured by the mean squared error (MSE) for each individual i , defined as

$$\text{MSE}^i = \frac{1}{40} \sum_{t=1}^{40} (x_t - F_t^i x_t)^2,$$

which captures individual i 's average squared forecast error over the 40 periods.

Table 5 reports the mean squared error, MSE^i , averaged across individuals for each treatment condition: $\rho = 0$ with and without signals, and $\rho = 0.6$ with and without signals. We find that average forecast errors are larger in treatments without signals compared to those with signals. However, the difference in MSE is relatively small when $\rho = 0$.

To formally test for differences in forecast accuracy, we conduct t -tests using individual-level MSEs as independent observations (with $n = 31$ per treatment). When the underlying process is persistent ($\rho = 0.6$), the average MSE is significantly higher without signals than with signals ($t = 4.0376$, $p = 0.0002$). In contrast, when the process is i.i.d. ($\rho = 0$), the difference is not statistically significant ($t = 1.5297$, $p = 0.1327$).

5.1.2 Overreaction to new information

We next examine the relationship between forecast errors and forecast revisions in the experiment without signals. Table 6 presents estimates of the individual-level regression based on equation (8), using data from the no-signal treatment. These results can be compared

Table 6: Predicting Forecast Errors with Forecast Revisions without Signal Extraction

	$\rho = 0$		$\rho = 0.6$	
	(1)	(2)	(3)	(4)
β^p	-0.681		-0.384	
SE	0.108		0.136	
t -stat	-6.26		-2.82	
β^i (Median)		-0.731		-0.376
R^2	0.221		0.127	
N	1000		1003	

Notes: The table reports coefficients from individual-level forecast error on forecast revision regression as in equation (8). Columns (1) and (3) show the estimation results from the individual-level fixed effect panel regression (β^p). Standard errors are clustered by both forecaster and time. Columns (2) and (4) show the median coefficients in forecaster-by-forecaster regressions (β^i).

with those in Table 2, which report the corresponding estimates for the experiment *with* signals.

Consistent with the findings from the main experiment, we observe a negative relationship between forecast revisions and subsequent forecast errors for both $\rho = 0$ and $\rho = 0.6$. The estimated coefficients on forecast revisions, β , are negative and statistically significant at the 1% level (Columns (1) and (3)). This indicates that expectations tend to overreact at the individual level even in the absence of noisy signals.

Furthermore, the medians of the estimated β^i from participant-by-participant regressions are negative as shown in Columns (2) and (4). Figure B.2 in Appendix B illustrates the distributions of the individual-level regression coefficient β^i across treatment groups, further supporting the evidence that $\beta^i < 0$.

Note also that, consistent with Afrouzi et al. (2023), forecast overreaction is more pronounced when the underlying process is less persistent. The estimated forecast error–revision coefficient is more negative in the i.i.d. case ($\rho = 0$), at -0.681 , compared to -0.384 when $\rho = 0.6$. This pattern suggests that the *absence* of a noisy signal may be a key driver of the negative relationship between process persistence and overreaction observed in Afrouzi et al.

Table 7: Forecast-Implied Subjective Persistence

	$\rho = 0$ ($\mu = 75$)		$\rho = 0.6$ ($\mu = 30$)	
	(1)	(2)	(3)	(4)
$\hat{\rho}$	0.637 (0.469, 0.805)	0.733	0.755 (0.624, 0.887)	0.707
SE	0.085		0.067	
t-stat	7.43		11.26	
$\hat{\mu}$	27.0 (17.4, 36.5)	20.0	18.2 (8.48, 27.92)	21.8
SE	4.673		4.750	
t-stat	5.78		3.83	
$\hat{\mu}/(1 - \hat{\rho})$	74.3	74.9	74.2	74.4
N	1049	1049	1037	1037

Notes: Columns (1) and (3) show the estimation results from the individual-level fixed effect panel regression ($\hat{\rho}^p$), clustered S.E. The parentheses indicate a 95% confidence interval. Columns (2) and (4) show the median coefficients in forecaster-by-forecaster regressions ($\hat{\rho}^i$).

(2023).

5.1.3 Forecast-implied subjective persistence

Finally, we examine subjective estimates of persistence and the process mean using data from the experiment without signals, following the same approach as in Section 4.2. Columns (1) and (3) of Table 7 report the estimates of $\hat{\rho}$ and $\hat{\mu}$, respectively. Consistent with the main experiment with signals—see Table 3—subjects tend to overestimate the persistence of the process ($\hat{\rho} > \rho$) in both treatments.¹¹ These results reinforce our earlier findings and provide further support for the overpersistence model as a plausible explanation for belief overreaction. As discussed above, the overconfidence mechanism is not applicable in the absence of signals.

¹¹Columns (2) and (4) report the median participant-level estimates of $\hat{\rho}$. The median subjective persistence exceeds the true value in both treatments, and forecast-implied persistence exceeds the actual ρ for 29 of 30 participants when $\rho = 0$, and for 20 of 29 participants when $\rho = 0.6$.

5.2 Consensus forecasts and information frictions

Following Coibion and Gorodnichenko (2015), we also estimate the predictability of aggregate (“consensus”) rather than individual forecast errors, using information on consensus forecast revisions:

$$\overline{\text{FE}}_t = \alpha + \beta^c \overline{\text{FR}}_t + u_t, \quad (14)$$

where

$$\overline{\text{FE}}_t \equiv x_t - \bar{F}_t x_t, \quad \overline{\text{FR}}_t \equiv \bar{F}_t x_t - \bar{F}_{t-1} x_t,$$

and u_t is the error term. Here, $\bar{F}_t[\cdot]$ denotes the *median* expectation of the variable of interest from all participants in a given treatment. Coibion and Gorodnichenko (2015) present compelling evidence of *underreaction* in the consensus forecasts of professional forecasters. They attribute this pattern to information frictions, arising from the fact that each forecaster observes a different set of private signals in addition to publicly available information.

Our two experimental treatments allow us to test this conjecture. Specifically, if information frictions drive underreaction, we should also observe underreaction in our experiment *with* signal extraction—where each subject receives an individual, noisy private signal—but not in the experiment *without* signals, where all participants observe only common information.

Table 8 presents estimates from the aggregate-level forecast revision regressions across treatment groups. Consistent with Coibion and Gorodnichenko (2015), we find that in the signal extraction treatment, the consensus forecast exhibits underreaction: $\beta^c > 0$ for both $\rho \in \{0, 0.6\}$. By contrast, in the no-signal extraction treatment, we observe overreaction: $\beta^c < 0$ for both values of ρ . Taken together, these results support the interpretation that underreaction in consensus forecasts arises from information frictions among forecasters: when agents possess different private signals, the consensus aggregates them only partially, leading to muted forecast revisions; when only common information is available, this friction is absent and the consensus may even overshoot in response to news.

Table 8: Predicting Consensus Forecast Errors with Forecast Revisions

	With Signals				Without Signals			
	$\rho = 0$		$\rho = 0.6$		$\rho = 0$		$\rho = 0.6$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
β^c	0.653	0.512	0.761	0.573	-0.931	-0.969	-0.461	-0.363
SE	0.152	0.093	0.163	0.127	0.422	0.341	0.256	0.195
p -value	4.27	5.46	4.66	4.50	-2.21	-2.84	-1.80	-1.86
R^2	0.330	0.446	0.369	0.354	0.116	0.179	0.080	0.085
N	39	39	39	39	39	39	39	39

Notes: The table reports coefficients from consensus forecast error on consensus forecast revision regression as in equation (14). Columns (1), (3), (5), and (7) show the estimation results obtained using the median in the forecaster-by-forecaster regressions. Columns (2), (4), (6), (8) show the estimation results obtained using the mean in the forecaster-by-forecaster regressions.

6 Conclusion

This paper reports on a laboratory experiment designed to evaluate three behavioral models of belief formation: diagnostic expectations, overprecision (overconfidence in private signals), and overpersistence (misperception of the underlying process). The experimental framework is grounded in the standard signal-extraction environment widely used in theoretical analyses.

In each period t , participants forecast the realization of a random variable in periods t and $t + 1$. Forecast accuracy is incentivized through performance-based payments. Participants are fully informed of the AR(1) data-generating process (DGP), including all parameter values, and have access to the full history of past realizations (up to 40 periods). This design minimizes confounds such as model misspecification and gives the rational expectations (FIRE) benchmark the best chance to succeed. In the main experiment, participants also receive a noisy signal of the current realization, as in Coibion and Gorodnichenko (2015), Angeletos et al. (2020), and Bordalo et al. (2020). An additional treatment without signals allows for direct comparison with the experimental design of Afrouzi et al. (2023).

We find robust evidence of belief overreaction to new information, regardless of whether the DGP is i.i.d. or persistent, and irrespective of signal availability. Among the three

models considered, overpersistence aligns most closely with the data. Moreover, comparing treatments with and without signals supports the information friction hypothesis of Coibion and Gorodnichenko (2015), wherein individual forecasts tend to overreact, even as consensus forecasts underreact.

This study contributes to the literature on belief formation by providing a clean experimental test of competing behavioral models under tightly controlled conditions. As theoretical models of expectations continue to evolve, experimental validation remains essential to distinguishing among them.

In our experiment, the variable that participants forecast is generated exogenously. By contrast, many macroeconomic and financial variables—such as GDP growth, inflation, and asset prices—are endogenously determined through the aggregated decisions of individuals, which depend in part on their expectations about future realizations of those variables. This endogenous belief–outcome interaction is explicitly modeled in a related experimental literature on expectation formation, known as the “learning-to-forecast” approach, which incorporates expectation–realization feedback loops (Marimon and Sunder, 1993; Hommes et al., 2005; Adam, 2007; Assenza et al., 2021; Anufriev et al., 2022; Evans et al., 2025, among others). While our motivation comes from survey evidence on belief updating by professional forecasters, firms, and households—where the feedback between expectations and outcomes is not explicitly modeled—a promising direction for future research is to embed the behavioral models studied here within a learning-to-forecast framework. Such an extension would incorporate belief–outcome interactions and thus enable richer tests of behavioral expectations in more realistic settings.

References

- Adam, K. (2007). Experimental evidence on the persistence of output and inflation. *Economic Journal* 117(603-636).
- Adam, K., P. Kuang, and S. Xie (2024). Overconfidence in private information explains biases in professional forecasts. *Mimeo*.
- Afrouzi, H., S. Kwon, A. Landier, Y. Ma, and D. Thesmar (2023). Overreaction in expectations: Evidence and theory. *Quarterly Journal of Economic* 138(3), 1713–1764.
- Angeletos, G.-M., Z. Huo, and K. A. Sastry (2020). Imperfect macroeconomic expectations: Evidence and theory. In *NBER Macroeconomics Annual 2020, volume 35*, NBER Chapters. National Bureau of Economic Research, Inc.
- Anufriev, M., J. Duffy, and V. Panchenko (2022). Learning in two-dimensional beauty contest games: Theory and experimental evidence. *Journal of Economic Theory* 201, 105417.
- Assenza, T., T. Bao, C. Hommes, and D. Massaro (2014). Experiments on expectations in macroeconomics and finance. In J. Duffy (Ed.), *Experiments in Macroeconomics*, Volume 17 of *Research in Experimental Economics*, pp. 11–69. Emerald Group Publishing Limited.
- Assenza, T., P. Heemeijer, C. Hommes, and D. Massaro (2021). Managing self-organization of expectations through monetary policy: A macro experiment. *Journal of Monetary Economics* 117, 170–186.
- Bao, T. and J. Duffy (2021). Signal extraction: Experimental evidence. *Theory and Decision* 90, 219–232.
- Bianchi, F., C. Ilut, and H. Saijo (2024a). Diagnostic business cycles. *Review of Economic Studies* 91(1), 129–162.
- Bianchi, F., C. Ilut, and H. Saijo (2024b). Smooth diagnostic expectations. *Mimeo*.

- Bordalo, P., K. Coffman, N. Gennaioli, and A. Shleifer (2016). Stereotypes. *Quarterly Journal of Economics* 131(4), 1753–94.
- Bordalo, P., N. Gennaioli, R. La Porta, and A. Shleifer (2019). Diagnostic expectations and stock returns. *Journal of Finance* 74(6), 2839–2874.
- Bordalo, P., N. Gennaioli, Y. Ma, and A. Shleifer (2020). Overreaction in macroeconomic expectations. *American Economic Review* 110(9), 2748–82.
- Bordalo, P., N. Gennaioli, and A. Shleifer (2018). Diagnostic expectations and credit cycles. *Journal of Finance* 73(1), 199–227.
- Candian, G. and P. D. Leo (2023). Imperfect exchange rate expectations. *Review of Economics and Statistics (Forthcoming)*.
- Chen, D. L., M. Schonger, and C. Wicken (2016). otree - an open-source platform for laboratory, online and field experiments. *Journal of Behavioral and Experimental Finance* 9, 88–97.
- Coibion, O. and Y. Gorodnichenko (2015). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review* 105(8), 2644–78.
- Evans, G. W., C. G. Gibbs, and B. McGough (2025). A unified model of learning to forecast. *American Economic Journal: Macroeconomics* 17(2), 101–133.
- Falato, A. and J. Xiao (2024). Expectations and credit slumps. *Mimeo*.
- Farhi, E. and I. Werning (2019). Monetary policy, bounded rationality, and incomplete markets. *American Economic Review* 109(11), 3887–3928.
- Gabaix, X. (2020). A behavioral New Keynesian model. *American Economic Review* 110(8), 2271–2327.

- García-Schmidt, M. and M. Woodford (2019). Are low interest rates deflationary? A paradox of perfect-foresight analysis. *American Economic Review* 109(1), 86–120.
- Gennaioli, N. and A. Shleifer (2010). What comes to mind. *Quarterly Journal of Economics* 125(4), 1399–1433.
- Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association* 1(1), 114–125.
- Hajdini, I. (2023). Mis-specified forecasts and myopia in an estimated new keynesian model. *American Economic Journal: Macroeconomics* (forthcoming).
- He, Z. and S. Kučinskas (2023). Expectations with correlated variables. *The Economic Journal* 134(660), 1517–1543.
- Hommes, C., J. Sonnemans, J. Tuinstra, and H. van de Velden (2005). Coordination of expectations in asset pricing experiments. *Review of Financial Studies* 18(3), 955–980.
- Kahneman, D. and A. Tversky (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology* 3(3), 430–454.
- Kohlhas, A. N. and A. Walther (2021). Asymmetric attention. *American Economic Review* 111(9), 2879–2925.
- L’Huillier, J.-P., S. R. Singh, and D. Yoo (2024). Incorporating diagnostic expectations into the new keynesian framework. *Review of Economic Studies* 91(5), 3013–3046.
- Marimon, R. and S. Sunder (1993). Indeterminacy of equilibria in a hyperinflationary world: Experimental evidence. *Econometrica* 61(5), 1073–1107.
- Maxted, P. (2024). A macro-finance model with sentiment. *Review of Economic Studies* 91(1), 438–475.

Na, S. and D. Yoo (2025). Overreaction and macroeconomic fluctuation of the external balance. *Journal of Monetary Economics* 151, 1–14.

Appendix for

Overreaction in Expectations under Signal Extraction:

Experimental Evidence

John Duffy, Nobuyuki Hanaki, Donghoon Yoo

October 4, 2025

A Further Derivation and Results

A.1 Kalman gain, signal-to-noise ratio, and persistence in the DGP

The optimal Kalman gain is given by

$$\lambda = \frac{\delta - (1 - \rho^2) + \delta A}{2 + \delta - (1 - \rho^2) + \delta A},$$

where $A = \sqrt{(1 - \rho^2)^2/\delta^2 + 1 + 2/\delta + 2\rho^2/\delta}$.

First, take the derivative of λ with respect to δ :

$$\frac{\partial \lambda}{\partial \delta} = \frac{2(1 + A + \delta \frac{\partial A}{\partial \delta})}{(2 + \delta - (1 - \rho^2) + \delta A)^2},$$

where $A = \sqrt{(1 - \rho^2)^2/\delta^2 + 1 + 2/\delta + 2\rho^2/\delta}$. We need to show whether $1 + A + \delta \frac{\partial A}{\partial \delta} > 0$. Since $1 + A + 2\rho \frac{1 - \frac{1-\rho}{\delta}}{A} > 0$, $\frac{\partial \lambda}{\partial \delta} > 0$. Therefore, the more precise the signal is, the larger the value of λ , indicating the more weight is attached to the observed signal when it is more precise.

Now, we take the derivative of λ with respect to ρ :

$$\frac{\partial \lambda}{\partial \rho} = \frac{2(2\rho + \delta \frac{\partial A}{\partial \rho})}{(2 + \delta - (1 - \rho^2) + \delta A)^2},$$

where $\frac{\partial A}{\partial \delta} = \frac{2\rho(1 - \frac{1-\rho}{\delta})}{\delta \sqrt{(1-\rho^2)^2/\delta^2 + 1 + 2/\delta + 2\rho^2/\delta}}$.

Since $\frac{(1-\rho^2)^2}{\delta^2} + 1 + \frac{2}{\delta} + \frac{2\rho^2}{\delta} > 1 - \frac{2(1-\rho)}{\delta} + \frac{(1-\rho)^2}{\delta^2}$, $\frac{\partial \lambda}{\partial \rho} > 0$. Thus, the more persistent the actual process, the more weight is attached to the observed signal.

A.2 Diagnostic Kalman filter

Diagnostic agent i forms expectations about the unobserved fundamental x_t by

$$F_t^{i,\theta} x_t = F_t^i x_t + \theta(F_t^i x_t - F_{t-1}^i x_t), \quad (15)$$

where θ is the diagnosticity parameter. If we substitute the signal extraction forecast

$$F_t^i x_t = F_{t-1}^i x_t + \lambda(s_t^i - F_{t-1}^i x_t), \quad (16)$$

in equation (15), we have

$$F_t^{i,\theta} x_t = F_{t-1}^i x_t + \lambda(s_t^i - F_{t-1}^i x_t) + \theta(F_{t-1}^i x_t + \lambda(s_t^i - F_{t-1}^i x_t) - F_{t-1}^i x_t). \quad (17)$$

This further simplifies to the so-called diagnostic Kalman filter given by (5):

$$F_t^{i,\theta} x_t = F_{t-1}^i x_t + (1 + \theta)\lambda(s_t^i - F_{t-1}^i x_t). \quad (18)$$

A.3 Predictions without signal extraction

Suppose that the underlying fundamental process x_t is an AR(1) process with mean μ :

$$x_t = \mu + \rho x_{t-1} + \varepsilon_t, \quad (19)$$

where ρ is a persistence parameter and ε_t is a mean-zero i.i.d. Normal shock with variance σ_ε^2 . However, subjects no longer receive a noisy signal about x_t , limiting their information set to past realizations of fundamental x_t . Specifically, the information at the time t is given by $\{x_{t-1}, x_{t-2}, x_{t-3}, \dots, x_{t-39}\}$. During 40 periods, subjects predict the realization of x_t and x_{t+1} . We consider the two cases: the underlying fundamental process is i.i.d. ($\rho = 0$) or it is persistent ($\rho = 0.6$). (In the experiments, σ_ε is set to 3, while μ is set to 75 when $\rho = 0$ and 30 when $\rho = 0.6$.)

A.3.1 Theory

Suppose a group of agents, indexed by $i \in [0, 1]$, aims to forecast a random variable x_t (and x_{t+1}) which is given by equation (19) based on the information available at the time of her forecast, $\mathcal{I}_{t-1} = \{x_{t-1}, x_{t-2}, x_{t-3}, \dots, x_{t-39}\}$.

Rationality. When an agent i rationally updates her belief, the forecasts $F_{t-1}^i x_t$ and $F_{t-1}^i x_{t+1}$ are given by:

$$\begin{aligned} F_{t-1}^i x_t &= \rho x_{t-1}, \\ F_{t-1}^i x_{t+1} &= \rho^2 x_{t-1}. \end{aligned}$$

Overpersistence. Suppose an agent i misperceives the persistence of the actual process, such that $\hat{\rho} > \rho$. The resulting forecasts, denoted by $\check{F}_t^i x_t$ and $\check{F}_t^i x_{t+1}$ are given by:

$$\begin{aligned}\check{F}_{t-1}^i x_t &= \hat{\rho} x_{t-1}, \\ \check{F}_{t-1}^i x_{t+1} &= \hat{\rho}^2 x_{t-1}.\end{aligned}$$

Diagnostic expectations. Suppose a diagnostic agent i makes a forecast about the random variable. The diagnostic forecasts, $F_{t-1}^{i,\theta} x_t$ and $F_{t-1}^{i,\theta} x_{t+1}$ are given by:

$$\begin{aligned}F_{t-1}^{i,\theta} x_t &= F_{t-1}^i x_t + \theta (F_{t-1}^i x_t - F_{t-2}^i x_t) = \rho x_{t-1} + \theta \rho \epsilon_{t-1}, \\ F_{t-1}^{i,\theta} x_{t+1} &= F_{t-1}^i x_{t+1} + \theta (F_{t-1}^i x_{t+1} - F_{t-2}^i x_{t+1}) = \rho^2 x_{t-1} + \theta \rho^2 \epsilon_{t-1}.\end{aligned}$$

A.3.2 Predictions

We define the forecast error (FE_t^i) and the forecast revision (FR_t^i) by:

$$FE_t^i = x_t - F_{t-1}^i x_t, \tag{20}$$

$$FR_t^i = F_{t-1}^i x_t - F_{t-2}^i x_t. \tag{21}$$

We run the following regression to project forecast errors onto forecast revisions at the individual level:

$$FE_t^i = \alpha + \beta FR_t^i + u_{i,t}.$$

Hypothesis 3 (Forecast revision coefficient β without signal extraction). As Hypothesis 1 predicts, when agents update their beliefs rationally, the estimated β should be close to zero. In contrast, over-persistence predicts a negative coefficient. Similarly, when the fundamental is persistent, DE predict a negative coefficient. For the i.i.d. process, however, the estimated coefficient should be close to zero under DE. Table A.1 (first and second row)

summarizes these predictions.

After making two forecasts $F_{t-1}^i x_t$ and $F_{t-1}^i x_{t+1}$ using the same information set, agent i 's forecasts $F_{t-1}^i x_{t+1}$ and $F_{t-1}^i x_t$ are related by

$$F_{t-1}^i x_{t+1} = \hat{\mu} + \hat{\rho} F_{t-1}^i x_t.$$

where $\hat{\mu}$ and $\hat{\rho}$ are forecast-implied subjective mean and persistence, respectively.

Hypothesis 4 (Forecast-Implied Persistence). Under Bayesian learning, the forecast-implied persistence should be the same as true persistence, ρ . The same holds for diagnostic agents. However, under overpersistence, agents misperceive the persistence of the process, leading to a forecast-implied persistence $\hat{\rho}$ that exceeds ρ . As a result, even in the absence of new information, biased forecasts may arise due to distorted beliefs about persistence. (Table A.1 (third row) summarizes predictions for the forecast long-run mean without signal extraction.)

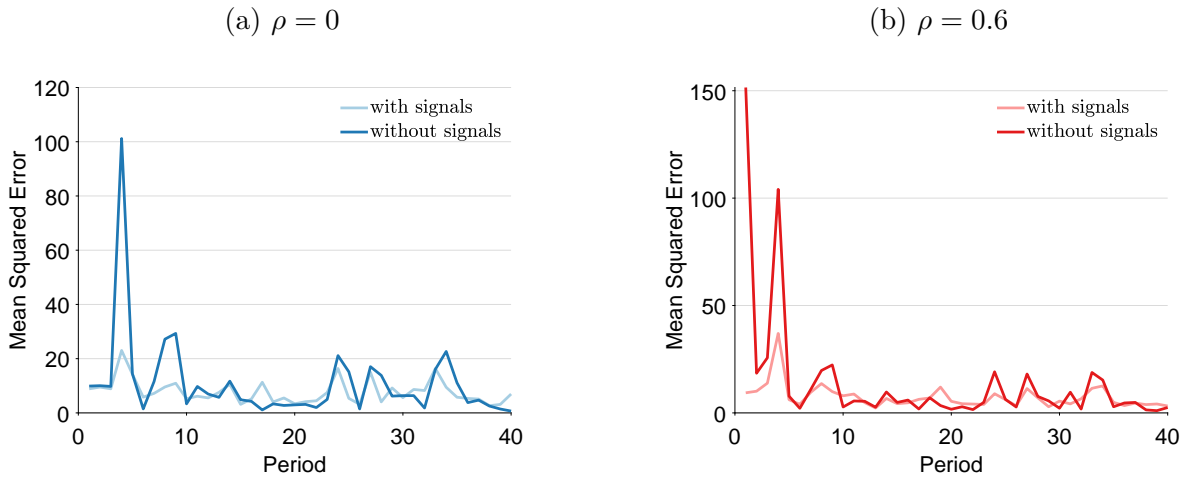
Table A.1: Summary of Theoretical Predictions without Signals

	<i>Rationality</i>	Overpersistence	<i>DE</i>
β ($\rho = 0$)	0	less than 0	0
β ($\rho > 0$)	0	less than 0	less than 0
$\hat{\rho}$	ρ	greater than ρ	ρ

Notes: With DE, the behavioral parameter θ is assumed to be strictly positive, $\theta > 0$.

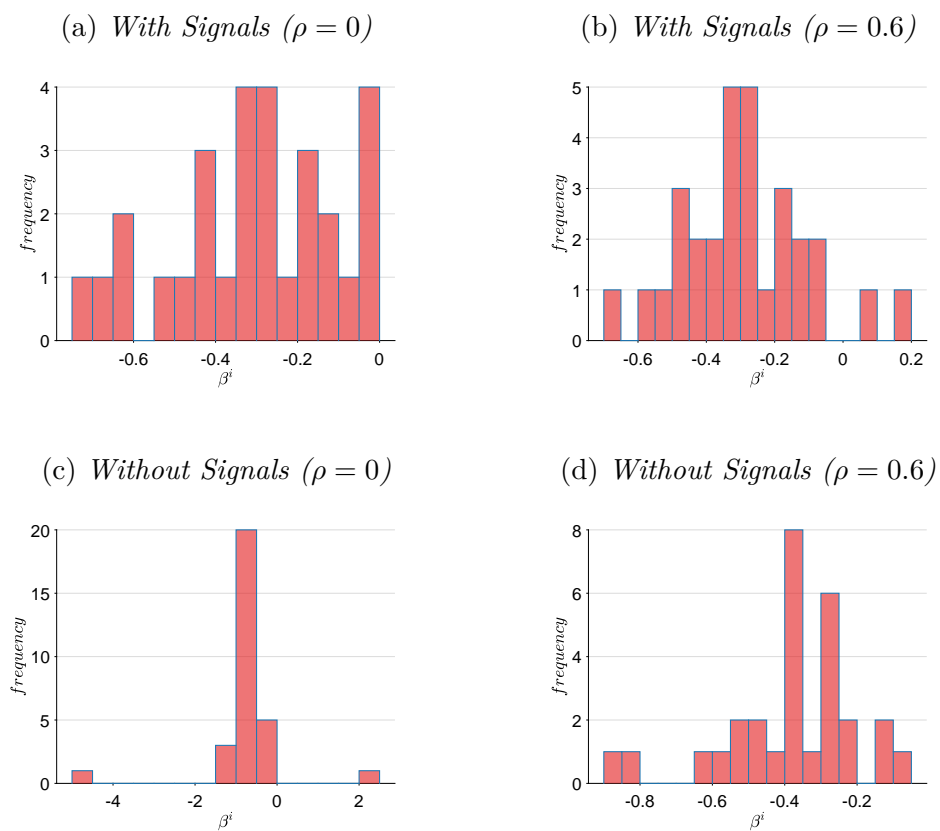
B Supplementary Figures and Tables

Figure B.1: The Period-by-period Average Forecast Errors



Notes: Panels (a) and (b) plot the period-by-period average mean squared errors for *with/without signals* ($\rho = 0$) and *with/without signals* ($\rho = 0.6$).

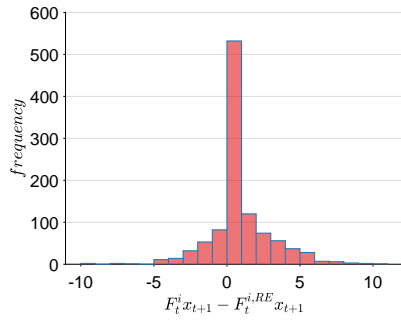
Figure B.2: The Distribution of Forecast Error on Forecast Revision Regression Coefficients



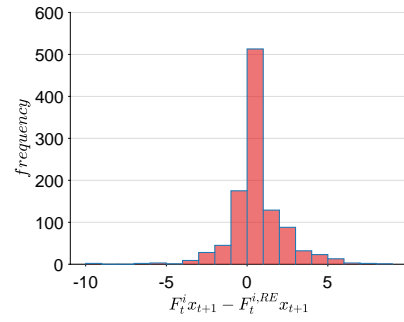
Notes: The panels present the distributions of the individual-level forecast error on forecast revision regression coefficients.

Figure B.3: The Distribution of the Measured Bias in Expectations

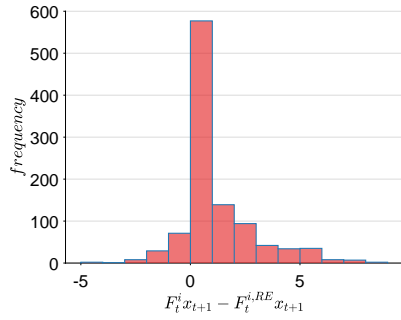
(a) *With Signals* ($\rho = 0$)



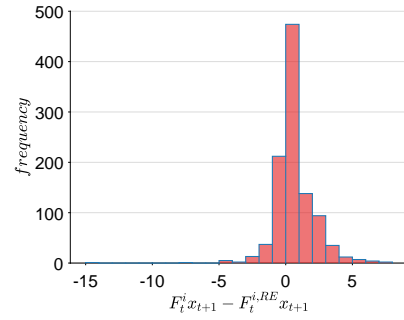
(b) *With Signals* ($\rho = 0.6$)



(c) *Without Signals* ($\rho = 0$)



(d) *Without Signals* ($\rho = 0.6$)



Notes: The panels present the distributions of the individual-level bias in expectations. This measure corresponds to the difference between the realized one-period-ahead forecast $F_t^i x_{t+1}$ and the one-step-ahead forecast a subject would have made under RE, $F_t^{i,RE} x_{t+1}$.

C Experimental Instructions

English translations of the instruction slides for the four treatments, as well as the quiz and screenshots of the decision screens, can be found at https://osf.io/xeg95/?view_only=2957785667844aa2846a06a42833259d.