Discussion Paper No. 1293

ISSN (Print) 0473-453X
ISSN (Online) 2435-0982

OVERREACTION IN EXPECTATIONS UNDER SIGNAL EXTRACTION: EXPERIMENTAL EVIDENCE

John Duffy
Nobuyuki Hanaki
Donghoon Yoo

Revised July 2026
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‡

July 25, 2026

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 experiment, participants repeatedly forecast the contemporaneous and one-step-ahead values of a random variable. They are incentivized for accuracy and are fully informed about the DGP and its past realizations. We consider a signal-extraction treatment in which participants receive noisy private signals about the unobserved contemporaneous value, as well as a treatment without signals, and we vary the persistence of the AR(1) process. 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 forecasts exhibit underreaction when agents observe different private signals, consistent with an information-friction explanation. We then study the same behavioral mechanisms in a simple macroeconomic consumption model and find that misperceptions about the persistence of productivity generate sizable distortions in consumption dynamics and provide a better fit to U.S. consumption and productivity data than overconfidence, diagnostic expectations, or rational Bayesian learning.

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; Collegio Carlo Alberto; ISER; Sungkyunkwan University; UC Irvine; the Sogang Summer Research Workshop; the IAER Macroeconomics Workshop; the 1st Asian Conference on Expectations in Macroeconomics; the BSE Workshop on Theoretical and Experimental Macroeconomics; and Annual Meeting of the Society for Experimental Finance for helpful comments. Financial support: Korea University Grant (K2513021, K2615201), JSPS KAKENHI (20H05631, 21H04397, 23H00055, 24K00238), and the Joint Usage/Research Center at ISER, The University of Osaka. The experiment was approved by the IRB of ISER (No. 20231201). We thank Hyunzi Oh and Romain Tranchecoste for excellent research assistance; Yoo gratefully acknowledges the hospitality of ISER, Eesti Pank and Collegio Carlo Alberto. 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. (2025) 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 motivated various models to explain deviations from the FIRE hypothesis—particularly the tendency for beliefs to overreact to new information—under alternative behavioral assumptions, including (1) diagnostic expectations (Bordalo et al., 2018; Bordalo et al., 2019; Bianchi et al., 2024; L'Huillier et al., 2024; Bianchi et al., 2026; Maxted, 2024; Na and Yoo, 2025); (2) belief misspecification, in particular, misperceptions of the persistence of the underlying data-generating process (Angeletos et al., 2020; Candian and Leo, 2025; Falato and Xiao, 2024); and (3) overconfidence in private information (Adam et al., 2025).[1] While each approach can account for departures from FIRE, to the best of our knowledge, a rigorous empirical comparison across these three frameworks does not yet exist. We do not claim that diagnostic expectations, overpersistence, and overconfidence exhaust all possible departures from Bayesian updating. Rather, we focus on these three mechanisms because they are prominent in the macroeconomic expectations literature, map naturally into our signal-extraction environment, and generate distinct empirical predictions.

Thus, in this paper we pose the following research question: among three behavioral models of expectation formation—(1) diagnostic expectations, (2) overpersistence, defined as the mistaken overestimation of the persistence of the underlying data-generating process (DGP), and (3) overconfidence 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

[1] In contrast, level-k thinking (Farhi and Werning, 2019; Garcìa-Schmidt and Woodford, 2019) and cognitive discounting (Gabaix, 2020; Hajdini, 2026) inherently lead to belief underreaction.

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 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.

In our experimental design, we vary two key features of the environment. First, we consider whether participants receive a noisy private signal about the contemporaneous realization of the random variable (signal vs. no-signal environments).[2] Second, we vary the persistence of the data-generating process (DGP), comparing a persistent AR(1) process ($\rho = 0.6$) with a non-persistent process ($\rho = 0$).

[2]The signal extraction environment is widely implemented in studying expectation formation. See, for example, Coibion and Gorodnichenko (2015), Bordalo et al. (2020), Angeletos et al. (2020), Adam et al. (2025), among others. Bao and Duffy (2021) provide experimental evidence on the extent to which individuals can solve signal extraction problems.

This 2×2 design allows us to distinguish the three behavioral models we consider. As discussed in Section 2, the models generate distinct predictions depending on the persistence of the DGP and on whether subjects observe noisy private signals. Under diagnostic expectations (DE), beliefs overreact only when the DGP is persistent, $\rho > 0$. Overconfidence in private signals can generate overreaction in the signal treatments, including when $\rho = 0$. Overpersistence predicts overreaction when perceived persistence affects how subjects extrapolate current information; however, in the signal treatment with $\rho = 0$, this is a boundary case in which overpersistence predicts $\beta = 0$, while overconfidence predicts $\beta < 0$. The no-signal treatments are therefore central to our identification strategy, because overconfidence is absent by construction in those treatments, whereas misperceived persistence can still be detected from subjects' forecast-implied subjective persistence.

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 especially closely related to Afrouzi et al. (2023), who apply the Coibion and Gorodnichenko (2015) methodology to forecasts of an externally generated AR(1) process and document pervasive forecast overreaction across a wide range of treatments and parameter values. Our no-signal treatment is close in spirit to theirs, but differs in one important respect: subjects in our experiment are given full knowledge of the DGP, including its parameter values.[3] This reduces the scope for departures from FIRE to reflect uncertainty about the environment itself. We then build on this benchmark by introducing treatments with noisy private signals, which allow us to move beyond documenting overreaction and to distinguish among alternative mechanisms, including overpersistence, overconfidence, and diagnostic expectations.

Another 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

[3]In one treatment of Afrouzi et al. (2023), participants are told that the DGP is an AR(1) process, but the parameter values are not disclosed.

forces—diagnostic expectations, overpersistence, or overconfidence—that drive the systematic overreaction that we observe in our setting.

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 overestimate 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 overconfidence, computed from the estimated weight on the signal based on a signal extraction model and the estimated subjective persistence of the process, shows that overpersistence, and not overconfidence, is the model that best explains the observed belief overreaction of most of the participants. (Among those exhibiting belief overreaction, 80% are driven by overpersistence, either alone or in combination with overconfidence, whereas only 13% are driven by overconfidence alone.) The result from the no-signal treatment is consistent with the conclusion we drew from the signal treatment. Note that by construction, overconfidence 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 our experimental design and the procedures we followed in our treatments with and without signal extraction and section 4 presents our main results. Section 5 explores the implications of our findings of belief misspecification for a simple macroeconomic model of consumption, and shows that our finding in favor of overpersistence is also a prominent candidate for explaining consumption dynamics in U.S. data. Finally, section 6 concludes.

2 Theory

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.[4] 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*. The framework developed below is the natural benchmark for the experimental treatment involving signals and signal extraction. We also study an environment without signals. In that case, the overconfidence mechanism is not applicable, because subjects do not observe a noisy signal whose precision may be misperceived. However, overpersistence and, when the underlying process is persistent, diagnostic expectations still deliver predictions for forecast dynamics in the no-signal case. 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

[4]Throughout the paper, we use the terms *forecasters*, *individuals* and *agents* interchangeably.

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 realized but unobserved value of x_t:

$$s_t^i = x_t + \omega_t^i, \tag{2}$$

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 agent's own prior belief about x_t, $F_{t-1}^i x_t$. The analysis below therefore applies directly to the case in which agents receive noisy signals. For the no-signal environment, one can imagine agents as updating using only past realizations of $\{x_k\}_{k=1}^{t-1}$, so that hypotheses involving signal misperception are no longer relevant, whereas hypotheses involving perceived persistence remain relevant.

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

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

where $\lambda \in (0, 1]$ is the optimal Kalman gain. As shown in Appendix A.1,

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

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.2, λ is increasing in both δ and ρ. Thus, holding the innovation variance fixed, more precise signals imply a higher Kalman gain; similarly, a more persistent underlying process leads agents to place greater weight on incoming signals.[5]

2.1 Theories of overreaction

Individual rationality requires agents to make the best possible use of the available information, as described by equation (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 as $\hat{\delta}$ where $\hat{\delta} \neq \delta$. Overconfidence corresponds to the case in which $\hat{\delta} > \delta$. Since the weight on observed signals, λ, increases with the perceived value $\hat{\delta}$, 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, agents overweight the probability of states that become more likely upon receiving new information.[6]

[5] In our experimental design and in the empirical specification lagged realizations are observed each period, so the relevant one-step-ahead prior variance is simply σ_ε^2 and the gain $\lambda = \delta/(1+\delta)$ does not depend on ρ.

[6] See Bordalo et al. (2018, Proposition 1, p. 209) for details.

In our experiment, the process for x_t follows an AR(1) process 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:

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

where $\mathbb{E}_t^\theta[\cdot]$ denotes the diagnostic expectations operator and $\theta \geq 0$ is the diagnosticity parameter governing the degree of overreaction.[7] The term $\big(\mathbb{E}_t[x_{t+1}] - \mathbb{E}_{t-1}[x_{t+1}]\big)$ 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 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.3 under DE, the dynamics of $F_t^{i,\theta}x_t$ are given by

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

where $\lambda \in (0, 1]$ is the optimal Kalman gain computed using the true model parameters and $\theta > 0$ 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.

3 Experimental Design

Our experiment closely follows the theoretical setting presented in the previous section. The experiment consists of four treatments in a 2×2 design that varies signal availability (signal

[7]We 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. (2024).

vs. no signal) and persistence of the DGP ($\rho = 0$ vs. $\rho = 0.6$). 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 $F_t^i x_{t+1}$. In the no-signal treatment, subjects receive no signal about x_t prior to submitting their forecasts. By contrast, in the signal treatments, each participant receives a noisy signal s_t^i about x_t defined by $s_t^i = x_t + \omega_t^i$ with $\sigma_\omega = 3$. In both the no-signal and signal treatments, participants could observe the realizations of x over the previous 40 periods in both a time-series graph and a table displayed on their screens. In the signal treatments, they could also observe the corresponding past signals.

We consider two parameterizations in both the no-signal and signal environments: one with $\rho = 0$ and one with $\rho = 0.6$. In both cases, μ is chosen so that the long-run mean of 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.[8] 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 all four treatments, each participant faced the same sequence of realizations of x_t. In the signal treatments, each participant received a private signal s_t^i in every period and therefore observed a different signal realization from the other participants in the same treatment. Having multiple participants receive different private signals while facing the same DGP and forecasting task enables us to compare individual and consensus forecasts.

Participants were paid according to the accuracy of one randomly selected forecast of x_j, $F_j^i x_j$, and one randomly selected forecast of x_{k+1}, $F_k^i x_{k+1}$, where j and k denote randomly

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

selected periods. Specifically, participant i's payoff in experimental currency units (ECUs) is given by

$$\pi^i_{j,k} = \frac{100}{1 + |x_j - F^i_j x_j|} + \frac{100}{1 + |x_{k+1} - F^i_k x_{k+1}|}. \tag{6}$$

Participants were given on-screen instructions detailing the tasks they would perform in each of the 40 periods (see Appendix E 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 understood the tasks, they were required to answer control questions based on the instructions. These questions covered the data-generating process, including the means and variances of ε and ω, and required subjects to derive x_{t+1} given $x_t = 20$ and $\varepsilon_{t+1} = 0$. 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.[9] A total of 124 subjects were recruited using ORSEE (Greiner, 2015), with 31 assigned to each of the four treatments: no signal with $\rho = 0$, no signal with $\rho = 0.6$, signal with $\rho = 0$, and signal with $\rho = 0.6$. Participants were undergraduate and graduate students at Osaka University. Participants received an e-mail with the link to the experimental site, on which they were randomly allocated to the different treatments. They participated in the experiment using their own device (only computers and tablets were allowed) from a location of their choice, provided that they had a reliable internet connection and could concentrate on the experiment. Once they had completed all tasks, they were informed of their earnings 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, participants in the no-signal treatments earned 1,536 Japanese yen (JPY) while those in the signal treatments earned 1,710 JPY, including a show-up payment of 500 JPY. Participants were paid using Amazon Gift Cards (e-mail version) on the same day as the experiment.

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

3.1 Hypotheses

We derive four hypotheses based on the models and information structure presented above. Hypothesis 1 concerns whether individual forecast errors are systematically related to forecast revisions. We carefully derive predictions based on the candidate models with and without signals. An important qualification is that the no-signal treatments cannot evaluate the overconfidence model, since subjects do not observe noisy private signals. They remain informative about overpersistence and, when $\rho > 0$, diagnostic expectations, since both theories imply restrictions on forecast revisions. Hypothesis 2 concerns subjective persistence, measured by the forecast-implied estimate $\hat{\rho}$. Hypothesis 3 concerns forecast accuracy and compares treatments with and without signals. Finally, Hypothesis 4 concerns the consensus forecast dynamics in the treatments with private signals.

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

$$FE_t^i = \alpha + \beta \times FR_t^i + u_t^i, \tag{7}$$

where

$$FE_t^i = x_t - F_t^i x_t, \quad FR_t^i = F_t^i x_t - F_{t-1}^i x_t.$$

Under overconfidence, $\hat{\delta} > \delta$, and, as shown in Appendix B, we have

$$\beta = -\frac{\hat{\delta}(\hat{\delta} - \delta)}{\delta\rho^2 + \delta^2(\delta + 1 - \rho^2)}. \tag{8}$$

Thus, if $\hat{\delta} = \delta$, then $\beta = 0$, whereas if $\hat{\delta} > \delta$, the overconfidence model predicts $\beta < 0$.

Under overpersistence, $\hat{\rho} > \rho$, and, as shown in Appendix B, we have

$$\beta = \frac{-(\hat{\rho} - \rho)\left[\hat{\rho}(1 - \hat{\rho}\rho) - \delta(\hat{\rho} - \rho)\right]}{(\hat{\rho} + \delta\rho)^2(\hat{\rho} - \rho)^2 + (1 - \rho^2)\left[(\hat{\rho} - \delta(\hat{\rho} - \rho))^2 + \delta^2 + \delta(1 + \hat{\rho}^2)\right]}. \tag{9}$$

This implies that if $\hat{\rho} = \rho$, then $\beta = 0$. When $\hat{\rho} > \rho$, the denominator is positive; therefore, the sign of β depends on the signal-to-noise ratio, δ.

Under diagnostic expectations (DE), the diagnosticity parameter θ distorts Bayesian updating by shifting the conditional predictive distribution toward states that become more likely following new information. When $\theta = 0$, expectations coincide with the Bayesian benchmark. When $\theta > 0$, DE increases the probabilities assigned to 'representative" future states and decreases those assigned to 'unrepresentative" states. This mechanism is forward-looking: in an AR(1) environment, news about x_t from the signal s_t^i changes the predictive distribution of x_{t+1} only when $\rho > 0$. Thus, DE predicts forecast overreaction when the process is persistent.

The case $\rho = 0$ is a distinct boundary case and must be treated separately. Equation (5) is the diagnostic Kalman filter for the persistent case and is derived under the maintained assumption $\rho > 0$; mechanically evaluating it at $\rho = 0$ is therefore not part of the model in Bordalo et al. (2020). Following their treatment of the boundary case, when the process is i.i.d., current news does not alter the distribution of future states, the representativeness distortion is inoperative, and DE coincides with Bayesian updating.[10] In the context of our individual forecast-error specification (7), Appendix B therefore gives

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

Accordingly, when $\rho > 0$ and $\theta > 0$, DE predicts $\beta < 0$, whereas when $\rho = 0$, it predicts $\beta = 0$. This distinction is especially useful in the no-signal treatments, where overconfidence is absent by construction, so evidence on β helps discriminate primarily among Bayesian updating, overpersistence, and, when $\rho > 0$, diagnostic expectations.

[10]See Appendix A of Bordalo et al. (2020). Their diagnostic Kalman filter and the associated forecast-error–revision coefficients are derived under the explicit restriction $\rho > 0$. They consider $\rho = 0$ separately and obtain zero forecast revision and a zero Coibion–Gorodnichenko coefficient. Thus, the $\rho = 0$ prediction is a separate boundary result, not the value obtained by substituting $\rho = 0$ into the persistent-case filter.

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.

In Table 1, the first two rows of each panel summarize the model-specific predictions for β under $\rho = 0$ and $\rho > 0$, respectively. Panel (a) reports predictions for treatments with private signals, while Panel (b) reports predictions for treatments without private signals. As noted earlier, overconfidence is applicable only in the signal treatments, where subjects observe noisy private signals whose precision may be misperceived. The no-signal treatments nevertheless provide a useful complementary test of overpersistence, because they remove the overconfidence channel and therefore help isolate whether excessive subjective persistence alone can account for biased forecast dynamics.

The signal-treatment case with $\rho = 0$ is a useful boundary case. Since the true process is i.i.d., both diagnostic expectations and overpersistence imply $\beta = 0$, whereas overconfidence in the private signal implies $\beta < 0$. Thus, a negative β in this cell is informative about overconfidence, but does not by itself support overpersistence.

Suppose that agent i makes two forecasts in period t, $F_t^i x_t$ and $F_t^i x_{t+1}$, as 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:

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

where $\hat{\mu}$ and $\hat{\rho}$ represent the subjective mean and persistence, respectively. As noted above, overpersistence implies $\hat{\rho} > \rho$. Thus, by regressing $F_t^i x_{t+1}$ on $F_t^i x_t$, one can directly test this model of overpersistence. This prediction is relevant both with and without signals, since

Table 1: Theoretical Predictions

(a) With Signals

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

(b) Without Signals

	Bayesian	Overconfidence	Overpersistence	Diagnostic Expectations
β $(\rho = 0)$	–	N/A	< 0	–
β $(\rho > 0)$	0	N/A	< 0	< 0
$\hat{\rho}$	ρ	N/A	$\hat{\rho} > \rho$	ρ

Notes: With DE, the behavioral parameter θ is assumed to be strictly positive, $\theta > 0$. In the no-signal treatment (without signals), the overconfidence model is not applicable because subjects do not observe noisy private signals. The overpersistence prediction for the $\rho > 0$ case with signals depends on our choice of $\delta = 1$. A dash indicates that the forecast revision regression is not identified or not informative in that case.

perceived persistence can be distorted even when agents observe only past realizations and no contemporaneous signal.

Our second hypothesis summarizes the prediction regarding subjective persistence:

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 $\hat{\rho} > \rho$.

Row 3 of each panel in Table 1 summarizes these model-specific predictions for $\hat{\rho}$.

The presence of signals may also affect the overall accuracy of forecasts. Under Bayesian learning, an additional informative signal expands the agent's information set and should weakly improve forecast accuracy relative to an otherwise identical environment without signals. More generally, even if agents do not update fully rationally, one may still expect forecasts to be more accurate on average when subjects observe informative signals than when they must rely solely on past realizations. This leads to our third hypothesis:

Hypothesis 3 (Signals and forecast accuracy)**.**

- Forecasts should be weakly more accurate in treatments with signals than in otherwise identical treatments without signals.

- This improvement should be stronger when the underlying process is persistent, because the contemporaneous signal is then more informative about future realizations.

To evaluate this prediction, we will compare the mean squared error of individual forecasts across treatments with and without signals.

We next turn to the behavior of consensus forecasts, defined as the cross-sectional median of individual forecasts in each period. This leads to the following hypothesis. In the overpersistence model, the consensus coefficient β^c reflects two opposing forces: information frictions push it in a positive direction, whereas overpersistence pushes it in a negative direction. Similarly, for agents with diagnostic expectations when $\rho > 0$, the sign of β^c depends on the interaction between information frictions and diagnosticity.

Hypothesis 4 (Consensus forecast revision coefficient β^c)**.**

- If agents are Bayesian or overconfident, then $\beta^c > 0$.

- If agents have diagnostic expectations or exhibit overpersistence, and the fundamental is i.i.d. ($\rho = 0$), then $\beta^c > 0$.

- In all other non-Bayesian settings, the sign of β^c depends on the relative magnitudes of belief distortions and information frictions.

Table 2 summarizes the predictions for consensus forecasts in the treatments with signals across the four models considered above with parameters used in our experiment: $\rho = \{0, 0.6\}, \delta = 1$.[11]

[11]Figure D.1 in Appendix D maps the signs of β and β^c over an even broader parameter region, while Tables 1 and 2 provide the predictions relevant to our experiment.

Table 2: Theoretical Predictions for Consensus Forecasts

	Bayesian	Overconfidence	Overpersistence	Diagnostic Expectations (DE)
β^c $(\rho = 0)$	> 0	> 0	> 0	> 0
β^c $(\rho > 0)$	> 0	> 0	> 0	$\gtreqless 0$

Notes: With DE, the behavioral parameter θ is assumed to be strictly positive, $\theta > 0$. The sign of β^c under diagnostic expectations depends on the exact value of θ. For the other models, the predictions do not depend on particular values of the underlying parameters.

4 Results

We begin by investigating Hypothesis 1, which concerns the relationship between forecast errors and forecast revisions. We next examine Hypothesis 2 on individuals' perceived parameter values for the DGP, in order to assess which model best accounts for the observed forecasting behavior. We then turn to Hypothesis 3 and compare forecast accuracy across treatments with and without signals, asking whether the availability of signals reduces the deviation of forecasts from realized outcomes. Finally, we examine Hypothesis 4 by analyzing the relationship between forecast errors and forecast revisions in consensus forecasts.

4.1 Overreaction to new information in individual forecasts

As noted in Section 3.1, we analyze forecast error predictability at the individual level by adapting the regression in equation (7). Specifically, we conduct participant-by-participant regressions to obtain a distribution of individual coefficients β^i, $i = 1, \ldots, I$, and we also estimate a common coefficient β^p by pooling all participants in a given treatment and including participant fixed effects.

We begin with the treatments in which subjects receive signals. Table 3 reports the estimates of the individual-level regressions for these treatments. In both cases, we find a negative relationship between forecast revisions and subsequent forecast errors, so that $\beta^p < 0$. The estimated pooled coefficient β^p is negative and statistically significant at the 1 percent level in both treatments; see Columns (1) and (3). Moreover, the medians of

Table 3: Predicting Forecast Errors with Forecast Revisions; Signal Treatment

	$\rho = 0$		$\rho = 0.6$	
	(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		0.175	
N	1029		1036	

Notes: The table reports coefficients from individual-level forecast error on forecast revision regression as in equation (7). 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 D.2 in Appendix D plots the distribution of the regression coefficient β^i.

the estimated β^i's from the participant-by-participant regressions are also negative in both treatment groups; see Columns (2) and (4). Specifically, β^i is negative for 26 out of 29 participants when $\rho = 0$ and for 29 out of 30 participants when $\rho = 0.6$. Figure D.2 in Appendix D shows the distributions of the individual-level regression coefficients β^i in each treatment and provides further support for the conclusion that $\beta^i < 0$. Overall, these findings indicate that expectations tend to overreact at the individual level, consistent with the overreaction observed in macroeconomic and financial forecasts by Bordalo et al. (2020).

We next turn to the no-signal treatments. Table 4 reports the corresponding estimates based on the same regression in equation (7). As in the signal treatments, we again find a negative relationship between forecast revisions and subsequent forecast errors for both $\rho = 0$ and $\rho = 0.6$. The estimated coefficients on forecast revisions are negative and statistically significant at the 1 percent level in both cases; see Columns (1) and (3). Likewise, the medians of the participant-level coefficients β^i are negative in both treatment groups; see Columns (2) and (4). Thus, expectations continue to overreact even in the absence of noisy signals. These results are robust to excluding outliers. See Table C.1 in Appendix C.

Table 4: Predicting Forecast Errors with Forecast Revisions: No-Signal Treatment

	$\rho = 0$		$\rho = 0.6$	
	(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 (7). 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).

This comparison is informative for distinguishing among the competing behavioral models. In the treatments with signals, the negative relationship between forecast revisions and subsequent forecast errors rejects the Bayesian learning model. It also rejects diagnostic expectations when $\rho = 0$, since that model predicts $\beta = 0$ in the i.i.d. case. The $\rho = 0$ signal treatment is a boundary case for overpersistence as well: as shown in Table 1, overpersistence also predicts $\beta = 0$ in this case, while overconfidence predicts $\beta < 0$. Thus, the negative coefficient in the signal treatment with $\rho = 0$ is inconsistent with overpersistence. The case for overpersistence comes instead from the forecast-implied subjective persistence estimates reported below. By contrast, in the signal treatment with $\rho = 0.6$, both overconfidence and overpersistence can generate $\beta < 0$. The no-signal treatments are therefore important because they rule out overconfidence as a general explanation for overreaction: the overconfidence mechanism requires noisy private signals and cannot operate when subjects do not observe such signals. Taken together with the forecast-implied subjective persistence estimates reported below, these findings point to overpersistence as the most robust explanation for the overall pattern of forecast dynamics, while allowing for some role for overconfidence in the signal treatments.

An interesting comparison is with Afrouzi et al. (2023), who find that forecasts exhibit more overreaction when the underlying process is more transitory. In our treatments with signals, by contrast, the estimated forecast error–revision coefficient β^p is very similar across the two values of ρ, at roughly -0.31, and the medians of the β^i's are likewise close, at around -0.29. However, in the no-signal treatments, forecast overreaction is substantially stronger when the underlying process is less persistent: the estimated coefficient is -0.681 when $\rho = 0$ and -0.384 when $\rho = 0.6$. Since our design differs from that of Afrouzi et al. (2023) in an important respect—subjects in our no-signal treatments are informed of the exact data-generating process—the results are not directly comparable. Still, the comparison of the signal and no-signal treatments suggests that the absence of a noisy signal may help explain why overreaction becomes more pronounced when the process is less persistent, as in Afrouzi et al. (2023).

4.2 Forecast-implied subjective persistence

We now evaluate Hypothesis 2 concerning forecast-implied subjective persistence and assess whether overpersistence may account for the negative β documented in the previous section. To this end, we estimate the empirical regression specified in equation (10).

We begin with the treatments in which subjects receive signals. Table 5 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

Table 5: Forecast-Implied Subjective Persistence: Signal Treatments

	$\rho = 0$ ($\mu = 75$)		$\rho = 0.6$ ($\mu = 30$)	
	(1)	(2)	(3)	(4)
$\hat{\rho}^p$	0.539		0.678	
	$(0.381, 0.697)$		$(0.587, 0.769)$	
SE	0.081		0.046	
t-stat	6.68		14.64	
$\hat{\rho}^i$ (Median)		0.455		0.645
$\hat{\mu}^p$	34.5		23.9	
	$(22.9, 46.1)$		$(17.4, 30.4)$	
SE	5.911		3.304	
t-stat	5.84		7.23	
$\hat{\mu}^i$ (Median)		40.5		26.3
$\hat{\mu}^p/(1-\hat{\rho}^p)$	74.87		74.17	
$\hat{\mu}^i/(1-\hat{\rho}^i)$		74.31		74.08
N	1063		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 ($\hat{\rho}^i$).

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.

We next examine subjective estimates of persistence and the process mean in the no-signal treatments, using the same approach. Table 6 reports the corresponding results. Consistent with the signal treatments, subjects again tend to overestimate the persistence of the process in both treatments. Columns (1) and (3) show that the point estimates of $\hat{\rho}$ exceed the true value of ρ in both cases. Columns (2) and (4) likewise show that the median participant-level estimates of $\hat{\rho}$ are greater than the actual persistence. In particular, forecast-implied persistence exceeds the actual ρ for 29 of 30 participants when $\rho = 0$, and for 20 of 29 participants when $\rho = 0.6$.

Table 6: Forecast-Implied Subjective Persistence: No-Signal Treatments

	$\rho = 0$ ($\mu = 75$)		$\rho = 0.6$ ($\mu = 30$)	
	(1)	(2)	(3)	(4)
$\hat{\rho}^p$	0.637		0.756	
	$(0.469, 0.805)$		$(0.624, 0.887)$	
SE	0.086		0.067	
t-stat	7.43		11.26	
$\hat{\rho}^i$ (Median)		0.733		0.707
$\hat{\mu}^p$	27.0		18.2	
	$(17.8, 36.2)$		$(8.9, 27.5)$	
SE	4.673		4.750	
t-stat	5.78		3.83	
$\hat{\mu}^i$ (Median)		20.0		21.8
$\hat{\mu}^p/(1-\hat{\rho}^p)$	74.43		74.49	
$\hat{\mu}^i/(1-\hat{\rho}^i)$		74.9		74.4
N	1049		1036	

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 ($\hat{\rho}^i$).

These findings reinforce the signal-treatment results and provide further support for overpersistence as a plausible explanation for belief overreaction. The no-signal treatments are especially informative because the overconfidence mechanism is absent by construction. Thus, subjects' tendency to overestimate persistence in both environments strengthens the case that overpersistence is the most robust explanation for the negative β documented above.

The estimated subjective parameters also suggest a behavioral interpretation of overpersistence. Subjects were incentivized to make accurate point forecasts, as specified in (6), rather than to report or estimate DGP parameters. The unconditional mean may therefore serve as a salient forecasting benchmark, while the persistence of deviations from that mean is a subtler dynamic feature. Consistent with this interpretation, subjects do not appear primarily to misperceive the unconditional mean. Instead, they overstate the persistence of deviations from that mean: when the process is above or below the long run mean of

75, their forecasts imply too slow a return toward 75 relative to the true DGP. In all four treatments, the estimated subjective persistence exceeds the true persistence: with signals, $\hat{\rho}^p = 0.539$ when $\rho = 0$ and $\hat{\rho}^p = 0.678$ when $\rho = 0.6$; without signals, $\hat{\rho}^p = 0.637$ when $\rho = 0$ and $\hat{\rho}^p = 0.756$ when $\rho = 0.6$. At the same time, the estimated intercepts move in the opposite direction. In the $\rho = 0$ treatments, where the true intercept is $\mu = 75$, the estimates are $\hat{\mu}^p = 34.5$ with signals and $\hat{\mu}^p = 27.0$ without signals. In the $\rho = 0.6$ treatments, where the true intercept is $\mu = 30$, the estimates are $\hat{\mu}^p = 23.9$ with signals and $\hat{\mu}^p = 18.2$ without signals. These two distortions largely offset each other, so that the implied long-run mean, $\hat{\mu}^p/(1 - \hat{\rho}^p)$, remains close to the true value of 75 in all treatments.

In this sense, overpersistence can be viewed as a form of *projection*: subjects use the correct long-run center of the process as a benchmark for point forecasting, but project deviations from that center forward more strongly than is warranted by the known DGP. This direction of projection is consistent with the behavioral heuristic of anchoring and insufficient adjustment. If the current realization provides the anchor, insufficient adjustment toward the long-run mean implies too much persistence of the current deviation, or $\hat{\rho} > \rho$ (Tversky and Kahneman, 1974; Epley and Gilovich, 2006). Because this projected process preserves approximately the correct unconditional mean, the distortion may be difficult for subjects to detect from observed realizations alone, especially in a finite sample. Realizations continue to fluctuate around 75, as both the true and subjective models imply; the error concerns the rate at which deviations from 75 should dissipate. This interpretation is also consistent with experimental evidence that subjects often underappreciate mean reversion in time-series forecasting tasks, even when the stochastic process is stable (Beshears et al., 2013; Afrouzi et al., 2023).

4.3 Estimating a signal extraction model

We now focus on the treatments with signal extraction in order to estimate the structural parameters governing individuals' belief formation. This exercise is specific to the signal treatments, since overconfidence is defined as excessive confidence in the precision of noisy private signals and is therefore not applicable in the no-signal environment. Our aim is to assess the role played by *overconfidence* ($\hat{\delta} > \delta = 1$), relative to overpersistence ($\hat{\rho} > \rho$), in explaining the overreaction observed in subjects' forecasts when signals are present. To do so, we follow the Bayesian model and assume 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

$$F_t^i x_t = (1 - \lambda) F_{t-1}^i x_t + \lambda s_t, \tag{11}$$

where λ represents the weight placed on the signal s_t.[12] 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:

$$F_t^i x_t = \text{const} + \gamma x_{t-1} + \lambda s_t + u_t^i, \tag{12}$$

where γ captures the persistence of the prior forecast and u_t^i denotes the error term.[13]

Given an estimate of $\hat{\lambda}$, we can recover the forecast-implied signal-to-noise ratio $\hat{\delta}$ using the parameter relationship, $\hat{\lambda} = \hat{\delta}/(1 + \hat{\delta})$. Appendix C.6 reports the distributions of $\hat{\lambda}$ and $\hat{\rho}$ pairs. Suppose that, for individual forecaster i, the revision coefficient β estimated in Section 4.1 is negative. Disregarding the role of the representativeness heuristic in the form of DE (this is especially true for $\rho = 0$), there are three possibile sources of belief overreaciton: (1) overpersistence ($\hat{\rho} > \rho$), (2) overconfidence ($\hat{\delta} > \delta$), and (3) both overconfidence and overpersistence.

Table 7 shows that overpersistence is the primary driver of belief overreaction as indicated by a negative β. When $\rho = 0$, 27% exhibit overpersistence alone and 65% exhibit both

[12] Note that equation (11) is simply another version of the rational Bayesian updating equation (3).

[13] See Appendix C.3 for the results based on an alternative specification: $F_t^i x_t = \text{const} + \gamma F_{t-1}^i x_t + \lambda s_t + u_t^i$. The results are very similar.

Table 7: Source of Overreaction

	$\rho = 0$	$\rho = 0.6$
Only overpersistence $(\hat{\rho} > \rho)$	7 (27%)	14 (48%)
Only overconfidence $(\hat{\delta} > 1)$	1 (4%)	6 (21%)
Both	17 (65%)	6 (21%)
Neither	1 (4%)	3 (10%)
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.

overpersistence and overconfidence; for $\rho = 0.6$, the corresponding shares are 48% and 21%, respectively. Instances of overconfidence alone are rare (4% and 21%), 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 in the signal treatments, whereas overconfidence alone plays only 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.

4.4 Informativeness of Signals

We now turn to Hypothesis 3, which concerns whether the availability of signals improves forecasting accuracy. Unlike the previous subsection, which focused only on the signal-extraction treatments in order to identify the role of overconfidence, this comparison necessarily uses data from both the signal and no-signal treatments. To evaluate this hypothesis, we compare the deviation of subjects' forecasts from actual realizations across the signal and no-signal treatments. Forecast accuracy is measured by the mean squared error (MSE) for

Table 8: Mean Square Error (MSE)

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

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 8 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 the treatments without signals than in the corresponding treatments 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 ($t = 1.5297$, $p = 0.1327$).

4.5 Consensus forecasts and information frictions

Finally, we test Hypothesis 4 concerning the behavior of consensus forecasts, defined as the cross-sectional median of individual forecasts in each period. When participants observe different private signals, aggregation may generate information frictions even if those signals improve individual forecast accuracy. These frictions tend to produce underreaction of the consensus forecast relative to the full-information rational-expectations benchmark and therefore push the consensus forecast revision coefficient, β^c, in a positive direction. Behavioral distortions may reinforce or offset this effect. As stated in Hypothesis 4, β^c is predicted to be positive for Bayesian and overconfident agents and, when the fundamental is i.i.d. ($\rho = 0$), for agents with diagnostic expectations or overpersistence. In the remaining non-Bayesian cases, its sign depends on the relative magnitudes of information frictions and belief distortions. Following Coibion and Gorodnichenko (2015), we assess the predictability of consensus forecast errors using revisions in the consensus forecast:

$$\overline{\mathrm{FE}}_t = \alpha + \beta^c \overline{\mathrm{FR}}_t + u_t, \tag{13}$$

where

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

and u_t is the error term. Here, $\bar{F}_t x_t$ denotes the median forecast of x_t across all participants in a given treatment. A positive estimate of β^c indicates that consensus forecast revisions predict subsequent forecast errors, consistent with underreaction. Coibion and Gorodnichenko (2015) document such underreaction in the consensus forecasts of professional forecasters and attribute it to information frictions arising because forecasters observe different private signals in addition to publicly available information.

Our two information 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 9 presents estimates from the aggregate-level forecast revision regressions across treatment groups. Consistent with Coibion and Gorodnichenko (2015), we find that in the signal treatments, the consensus forecast exhibits underreaction: $\beta^c > 0$ for both

Table 9: 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
t-stat	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 eq. (13). 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.

$\rho \in \{0, 0.6\}$. By contrast, in the no-signal treatments, 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.

5 Consumption Dynamics with Behavioral Distortions

The previous section showed that individual forecasts systematically overreact to new information and that overpersistence provided the most robust explanation for this pattern. The comparison of consensus forecasts across treatments also highlighted the role of information frictions when agents observe different private signals. We now ask whether these experimental findings have implications beyond the laboratory. To address this question, we study a simple macroeconomic consumption model that maps naturally into the signal-extraction environment of our experiment. The model allows us to examine how model misspecification and diagnostic expectations affect aggregate consumption dynamics and the model's fit to U.S. business-cycle data relative to a rational Bayesian benchmark.

The terminology in this section is more general than that used in the experimental analysis. What we previously described as overpersistence corresponds here to agents using a model in which the persistence of the underlying process is too high, while overconfidence corresponds to agents using a model in which the signal is too precise. Because our empirical framework allows perceived persistence and signal precision to be either above or below their rational Bayesian values, it encompasses both over- and underpersistence and both over- and underconfidence. We therefore use the neutral terms "misspecified persistence" and "misspecified precision" throughout this section. Here, model misspecification refers to the model agents use to form expectations, rather than to misspecification of the econometric model. Diagnostic expectations constitute a separate departure from rational Bayesian updating. The purpose of this exercise is to connect our experimental evidence to issues of direct interest to macroeconomists: the formation of expectations, the interpretation of signals, and the extent to which departures from rational Bayesian updating help explain aggregate fluctuations.

The model environment is analytically tractable and provides a useful starting point for studying post-war U.S. data.[14] Productivity is driven by two shocks: a permanent shock with lasting effects and a transitory shock whose effects gradually dissipate. Consumers base their spending decisions on expectations of future outcomes, particularly long-run productivity. Although agents observe aggregate productivity, they cannot separately identify its permanent and transitory components. In addition to current and past productivity, they receive a noisy signal about the permanent component. We allow agents to depart from rational Bayesian learning in three ways: by misperceiving productivity persistence, which we call misspecified persistence; by misperceiving signal precision, which we call misspecified precision; or by forming expectations diagnostically. Given this information structure and possible model

[14] See L'Huillier (2012) and Blanchard et al. (2013) for detailed descriptions of the model.

misspecification, agents solve a signal-extraction problem and choose consumption accordingly.

Information structure. We consider a representative-agent environment in which log productivity is driven by exogenous permanent and transitory components. Specifically, observed log productivity, a_t, is given by

$$a_t = x_t + z_t, \tag{14}$$

where x_t denotes the permanent component and z_t the transitory component.

The permanent component evolves according to an AR(1) process in first differences,

$$\Delta x_t = \rho_x \Delta x_{t-1} + \sigma_\varepsilon \varepsilon_t, \tag{15}$$

and the transitory component follows a stationary AR(1) process,

$$z_t = \rho_z z_{t-1} + \sigma_\eta \eta_t, \tag{16}$$

where ρ_x and ρ_z are the persistence parameters of the shock processes and ε_t and η_t are standard Normal shocks. We assume that $\rho_x = \rho_z = \rho$ and $\rho\sigma_\varepsilon^2 = (1-\rho)^2\sigma_\eta^2$, under which the observed univariate process for log productivity is a random walk, i.e.,

$$\mathbb{E}[a_{t+1} \mid a_t, a_{t-1}, \ldots] = a_t. \tag{17}$$

Agents observe productivity a_t, but do not separately observe its components x_t and z_t. This informational assumption is important since agents choose their current spending using their expectations about future productivity. Thus, agents must infer the permanent and transitory components from the history of observed productivity. In this linear-Gaussian setting, that inference is naturally characterized by the Kalman filter.

In addition to observing aggregate productivity a_t, agents observe a noisy signal about the permanent component:

$$s_t = x_t + \sigma_\omega \omega_t, \tag{18}$$

where ω_t is a standard normal shock that affects beliefs but is independent of fundamentals. This signal affects agents' beliefs about long-run productivity and, through those beliefs, their consumption decisions. The specification therefore provides a tractable signal-extraction framework that maps directly into our experimental design and is well suited to structural estimation with macroeconomic data.

Consumption. Following Blanchard et al. (2013), consumers choose spending equal to their long-run expectations of productivity (income):

$$c_t = \lim_{j \to \infty} \tilde{\mathbb{E}}_t[a_{t+j}], \tag{19}$$

where $\tilde{\mathbb{E}}_t[\cdot]$ is a generic expectation operator conditional on information available at time t. Thus, consumption depends on how agents form expectations.[15] For example, under imperfect information, rationality implies (among other restrictions) that agents correctly perceive the persistence of the underlying productivity process (ρ) and the precision of the signal (σ_ω).

We relax this rationality assumption by allowing agents to misperceive the economy's fundamentals and the precision of the information they receive. We consider two forms of misspecification. First, agents may misperceive the persistence of the productivity process. When they believe that productivity is more persistent than it truly is, they exhibit *overpersistence* and over-extrapolate from observed realizations. When they believe it is less persistent, they exhibit *underpersistence* and under-extrapolate. Second, agents may misperceive the precision of their signal. When they believe that the signal, s_t, is more precise

[15] As shown in Blanchard et al. (2013) and Cao and L'Huillier (2018), this limiting result for consumption can be derived from a standard New Keynesian model with highly infrequent price adjustments, or from an SOE-RBC model where the domestic interest rate is highly insensitive to debt holdings.

than it actually is, they exhibit *over-confidence* and place excessive weight on it. When they believe it is less precise, they exhibit *under-confidence* and place insufficient weight on the signal.

Under misspecification, consumption is governed by agents' distorted beliefs about long-run productivity (BLR).

Definition 1 (Misspecified BLR)**.** *Given information at time t, the agents' perceived long-run estimate of productivity is*

$$\lim_{j \to \infty} \hat{\mathbb{E}}_t[a_{t+j}] = \frac{\hat{\mathbb{E}}_t[x_t - \hat{\rho} x_{t-1}]}{1 - \hat{\rho}} = \frac{\hat{x}_{t|t} - \hat{\rho} \hat{x}_{t-1|t}}{1 - \hat{\rho}}, \tag{20}$$

where $\hat{x}_{t|t} \equiv \hat{\mathbb{E}}_t[x_t]$ and $\hat{x}_{t-1|t} \equiv \hat{\mathbb{E}}_t[x_{t-1}]$ denote beliefs that may depend on the misspecified parameters, $\hat{\rho}$ and/or $\hat{\sigma}_\omega$. The expectation operator $\hat{\mathbb{E}}_t[\cdot]$ represents expectations formed under the misspecified model.

Solving the Model. Solving the model for consumption amounts to applying the Kalman filter to form beliefs about long-run productivity, $\lim_{j \to \infty} \hat{\mathbb{E}}_t[a_{t+j}]$. Using eq. (20), consumption is given by

$$c_t = \frac{1}{1 - \hat{\rho}} \left(\hat{x}_{t|t} - \hat{\rho} \hat{x}_{t-1|t} \right), \tag{21}$$

where $\hat{x}_{t|t}$ and $\hat{x}_{t-1|t}$ denote agents' beliefs about current and lagged permanent productivity. These beliefs are obtained from the consumer's Kalman filter (possibly under misspecification), in which the unobservable state vector is $\mathbf{x}_t = (x_t,\, x_{t-1},\, z_t)'$ and the observable vector is $\mathbf{s}_t = (a_t,\, s_t)'$.[16]

While the information structure of this model is closely related to the signal-extraction environment used in the experiment, there are several important differences. First, in the experiment, subjects observe past realizations of the fundamental, whereas in the macro model

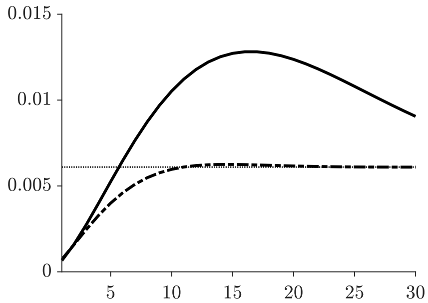
[16] Our empirical exercise also considers a consumption model with long-term beliefs guided by the diagnostic Kalman filter, introduced by Bordalo et al. (2020).

the permanent and transitory productivity components are never separately observed; agents observe only aggregate productivity and infer its decomposition. Second, in the experiment, subjects forecast a single random variable that follows an AR(1) process and, in the signal treatments, receive a noisy private signal about its current realization. In the macro model, by contrast, the signal is about the permanent component of productivity, so agents face a richer signal-extraction problem. Third, the macro model features a representative agent who observes a common signal, whereas the signal treatments in the experiment involve individual-specific private signals. Finally, in the macro model, expectations affect consumption choices, creating a channel from beliefs to actual spending that is directly disciplined by observed consumption data. Thus, while the model preserves the key signal-extraction logic of the experiment, the belief distortions estimated below should be interpreted as macroeconomic analogues of the experimental mechanisms rather than as direct one-to-one counterparts.

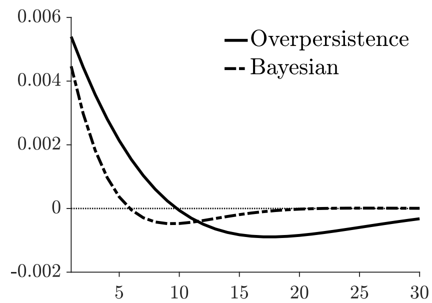
5.1 Consumption responses with model misspecification

In this section, we analyze how model misspecification affects consumption dynamics. Figures 1 and 2 plot, respectively, the consumption responses to permanent productivity and noise shocks under misspecified persistence, specifically overpersistence (with $\hat{\rho} > \rho$), and misspecified precision, specifically overprecision ($\hat{\sigma}_\omega < \sigma_\omega$). Here, time is measured in quarters. We use the estimated parameters reported in Table 10 (I). To analyze the implications of misspecified precision, we further assume that the perceived noise volatility is given by $\hat{\sigma}_\omega = 1/2\sigma_\omega$. All impulse responses are plotted alongside a Bayesian learning benchmark, for which $\hat{\rho} = \rho$ and $\hat{\sigma}_\omega = \sigma_\omega$.

We first examine responses under overpersistence as shown in Figure 1. In response to a positive permanent productivity shock ($\varepsilon_1 = 1$), consumption increases slowly because consumers do not immediately recognize the productivity boom, given the high volatility of the other shocks. Relative to Bayesian learning, overpersistence generates both a larger peak response and a slower time to converge to the new steady state. Under Bayesian learning,



(a) Permanent Tech. Shock



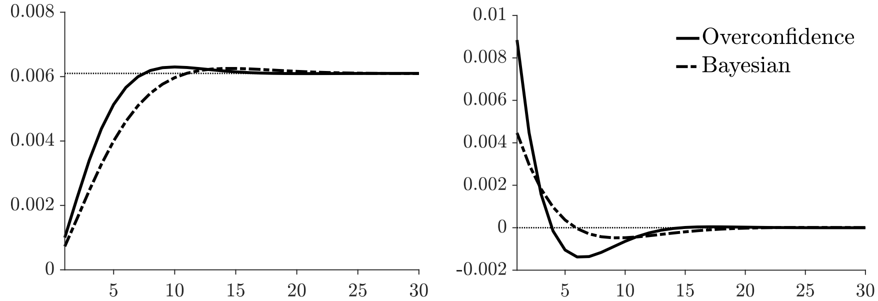
(b) Noise Shock

Figure 1: Impulse Responses of Consumption: Overpersistence

Notes: The panels show consumption impulse responses to a permanent productivity shock (left) and a noise signal (right). Solid lines denote overpersistence; dash-dotted lines denote Bayesian learning; black dotted lines denote the long-run response. Parameter values are from Table 10 (I): $\rho = 0.8814$, $\sigma_u = 0.0054$, $\sigma_\omega = 0.0082$. For the overpersistence case, the perceived persistence is $\hat{\rho} = 0.9729$.

consumption peaks at 0.0064 after 14 quarters and is essentially at its new long-run steady level after about 21 quarters. (The very small overreaction in the medium run is due to a mechanical effect induced by the persistence of beliefs.) By contrast, under overpersistence, consumption peaks at 0.0128 after 16 quarters and reaches its new steady-state level after 48 quarters (not shown in the figure). For a positive noise shock $\omega_1 = 1$, although productivity does not move, consumption initially increases and then gradually returns to its original level over time. Thus, overpersistence produces a larger initial response of 0.0054 (0.0045 under Bayesian learning) and a more protracted cycle.

Figure 2 plots consumption responses under overconfidence. Following a permanent productivity shock, consumption adjusts slowly toward the new long-run steady state, with overconfident consumers reaching the peak faster than under Bayesian learning. While consumers correctly perceive the persistence of the underlying fundamental, their misperception of signal quality generates biased consumption responses. Since $\hat{\sigma}_\omega < \sigma_\omega$, consumers place more weight on the signal they receive than Bayesian consumers do. This amplifies the initial consumption response and leads to stronger short-run reversals. However, generating sizable consumption distortions (relative to the Bayesian benchmark) requires a substantial degree



(a) Permanent Tech. Shock (b) Noise Shock

Figure 2: Impulse Responses of Consumption: Overconfidence

Notes: The panels show consumption impulse responses to a permanent productivity shock (left) and a noise signal (right). Solid lines denote overconfidence; dash-dotted lines denote Bayesian learning; black dotted lines denote the long-run response. Parameter values are from Table 10 (I): $\rho = 0.8814$, $\sigma_u = 0.0054$, $\sigma_\omega = 0.0082$. For the overconfidence case, the standard deviation of the perceived noise shock is $\hat{\sigma}_\omega = 0.0041$.

of misperception regarding information quality.[17] More interestingly, the estimation in the next section, as shown in Table 10 (II), provides evidence of *under*confidence rather than overconfidence when fitting the model to U.S. consumption and productivity data.

5.2 Model Estimation

In this section, we estimate the simple consumption model using Bayesian methods, incorporating the behavioral restrictions implied by model misspecification and diagnostic expectations. We construct the econometrician's Kalman filter by incorporating agents' expectations as state variables. The model can then be written in state-space form, with productivity growth, Δa_t, and consumption growth, Δc_t, as observables.

Priors. We set prior distributions consistent with the literature. The shock persistence parameters follow Beta distributions with mean 0.6 and standard deviation 0.2. The shock standard deviation parameters follow inverse-Gamma distributions with mean 1 and standard deviation 1, scaled by 100. For the diagnosticity parameter, θ, we assume a uniform prior

[17] Here, we illustrate overconfidence with $\hat{\sigma}_\omega = 1/2\sigma_\omega$. The effect is substantially lower for values closer to the benchmark, e.g., $\hat{\sigma}_\omega = 4/5\sigma_\omega$.

over $[-1, 3]$. (See Table D.1 in Appendix D for details.)

Estimation Results. The models are estimated using U.S. data for the consumption and productivity series. Using Bayesian methods, we evaluate the relative fit of belief misspecification. The consumption series is constructed by taking the log-difference of the ratio of NIPA consumption to population. The labor productivity series is constructed by taking the log-difference of the ratio of real gross domestic product (GDP) to employment. The data on real GDP (GDPC1), real personal consumption expenditure (PCECC96), employment (LNS12000000Q), and population (LNS10000000Q) are from 1976:II to 2019:III. We also remove secular drift in the consumption-to-productivity ratio from the consumption series. The simplicity of the model allows us to extract a significant amount of information using only these two series.

We consider four models: (I) misspecified persistence, (II) misspecified precision, (III) diagnostic expectations (DE), and (IV) Bayesian learning.[18] Specifically, in the misspecified persistence model, agents correctly perceive σ_ω but misperceive the persistence parameter as $\hat{\rho}$. In the misspecified precision model, agents correctly perceive ρ but misperceive the signal noise as $\hat{\sigma}_\omega$. In the DE model, agents correctly perceive the structural parameters but overreact to news, a tendency captured by θ. Finally, in the Bayesian learning model, agents update beliefs without any such systematic biases.

The parameter estimates are reported in Table 10. First, we find that the estimated DE parameter θ is negative and very close to zero, suggesting that the distortions generated by DE are negligible. Consequently, the estimated productivity and information parameters (ρ, σ_u, σ_ω) under DE are nearly identical to those obtained under Bayesian learning.

In contrast, the behavioral distortions induced by belief misspecification are more pronounced. Under misspecified persistence, the actual persistence of the productivity process is notably lower than its perceived counterpart ($\rho = 0.8814$ versus $\hat{\rho} = 0.9729$). Furthermore,

[18] For details on constructing a diagnostic Kalman filter, see Bordalo et al. (2020) and L'Huillier et al. (2021) (Section 4).

Table 10: Posterior Distribution

Parameter	Description	(I) Misspecified persistence		(II) Misspecified precision	
		Post. Mean	[05, 95]	Post. Mean	[05, 95]
ρ	persist.	0.8814	[0.8407, 0.9225]	0.9579	[0.9442, 0.9720]
$\hat{\rho}$	persist. (perc.)	0.9729	[0.9656, 0.9804]		
σ_u	tech. shock s.d.	0.0054	[0.0051, 0.0058]	0.0065	[0.0058, 0.0070]
σ_ω	noise s.d.	0.0082	[0.0048, 0.0114]	0.0073	[0.0042, 0.0102]
$\hat{\sigma}_\omega$	noise s.d. (perc.)			0.0092	[0.0057, 0.0126]
θ	diagnosticity				
log Data Density		1264.61		1260.66	

Parameter	Description	(III) DE		(IV) Bayesian learning	
		Post. Mean	[05, 95]	Mean	[05, 95]
ρ	persist.	0.9603	[0.9476, 0.9728]	0.9605	[0.9484, 0.9723]
$\hat{\rho}$	persist. (perc.)				
σ_u	tech. shock s.d.	0.0060	[0.0056, 0.0065]	0.0058	[0.0055, 0.0062]
σ_ω	noise s.d.	0.0108	[0.0059, 0.0157]	0.0115	[0.0062, 0.0165]
$\hat{\sigma}_\omega$	noise s.d. (perc.)				
θ	diagnosticity	-0.0774	[-0.1495, 0.0059]		
log Data Density		1255.48		1258.80	

Notes: With the random-walk assumption in (17), the variances of the two productivity shocks, σ_ε^2 and σ_η^2, are associated with the variance of Δa_t, denoted by σ_u^2: $\sigma_\varepsilon^2 = (1-\rho)^2\sigma_u^2$ and $\sigma_\eta^2 = \rho\sigma_u^2$. Thus, given ρ and σ_u, both σ_ε and σ_η are pinned down.

this estimated actual persistence is substantially lower than the values obtained in other models, where ρ is typically near 0.96. As illustrated in Figure 1, these estimates imply that overpersistence generates significant consumption distortions: consumption exhibits a gradual overreaction following a permanent technology shock, while a noise shock triggers an initial overreaction followed by a subsequent reversal.

Under misspecified precision, the estimated perceived standard deviation of the noise shock exceeds the true value ($\hat{\sigma}_\omega = 0.0092$ vs. $\sigma_\omega = 0.0073$). As a result, consumers view the signal as noisier—and therefore less precise—than it actually is. Relative to rational agents, they place too little weight on the signal, which implies underreaction to the signal. However the credible intervals overlap substantially, thus, we do not interpret this as strong evidence of underconfidence.

Bayesian learning is nested within the other three models: it corresponds to $\hat{\rho} = \rho$ under misspecified persistence, $\hat{\sigma}_\omega = \sigma_\omega$ under misspecified precision, and $\theta = 0$ under diagnostic expectations. We compare the empirical fit of the candidate models using Bayes factors. The log data density is 1264.61 for the misspecified-persistence model, 1260.66 for the misspecified-precision model, 1255.48 for the diagnostic-expectations model, and 1258.80 for the Bayesian-learning model. These differences provide strong evidence in favor of the misspecified-persistence model, whose estimates imply overpersistence, relative to the alternative models.

Variance Decompositions. Table 11 reports the implications of the estimated parameters for the variance decomposition, showing the contribution of the three shocks to the forecast error variance across different horizons. Noise shocks are the primary driver of short-run volatility across all specifications, explaining 83–92% of consumption fluctuations at a one-quarter horizon. However, as the horizon extends, permanent shocks play a much more significant role, particularly under misspecified persistence, where they account for 65.6% and 99.5% of the variance at the 4- and 8-quarter horizons, respectively. As evident from the impulse responses in Figure 1, permanent shocks generate much stronger consumption responses under this specification; this explains why permanent shocks contribute substantially to consumption fluctuations when agents misperceive shock persistence.

5.3 Discussion

This section considers several extensions and qualifications to our baseline empirical results. Specifically, we examine whether jointly misspecified persistence and precision can improve the model's empirical fit, and whether belief misspecification can be distinguished from diagnostic expectations using impulse responses.

Table 11: Variance Decomposition of Consumption

Horizon	(I) Misspecified persistence			(II) Misspecified precision		
	Permanent	Transitory	Noise	Permanent	Transitory	Noise
1	0.0135	0.0621	0.9243	0.0027	0.1631	0.8342
4	0.6562	0.0176	0.3262	0.3068	0.1844	0.5088
8	0.9952	0.0001	0.0047	0.9816	0.0168	0.0016
16	0.9951	0.0004	0.0046	0.9886	0.0002	0.0112
20	0.9951	0.0003	0.0046	0.9941	0.0004	0.0055
	(III) DE			(IV) Bayesian learning		
Horizon	Permanent	Transitory	Noise	Permanent	Transitory	Noise
1	0.0012	0.1302	0.8686	0.0011	0.1419	0.8570
4	0.1232	0.1690	0.7078	0.1104	0.1820	0.7075
8	0.9010	0.0495	0.0495	0.8611	0.0642	0.0747
16	0.9818	0.0000	0.0181	0.9834	0.0000	0.0166
20	0.9861	0.0004	0.0135	0.9855	0.0003	0.0142

5.3.1 Joint misspecification: allowing for both misspecified persistence and precision

We now allow both belief misspecification mechanisms to operate simultaneously, so that $\hat{\rho} \neq \rho$ and $\hat{\sigma}_\omega \neq \sigma_\omega$. This specification is estimated using the same U.S. dataset described in the previous section.

Table 12 reports the resulting parameter estimates. First, we observe that the model allowing for both misspecifications provides a better fit to the data than the model with *overpersistence* alone, which previously delivered the highest log marginal likelihood among the models considered in the last section. The log marginal likelihood is 1268.54 (versus 1264.61 in the model estimating both misspecifications). Second, the estimated persistence parameters, $\rho = 0.8727$ and $\hat{\rho} = 0.9723$, are very similar to those in the misspecified persistence model (0.8814 and 0.9729), while the estimated standard deviations of the perceived and actual noise in the signal, $\hat{\sigma}_\omega = 0.0087$ and $\sigma_\omega = 0.0066$, closely resemble those in the *misspecified precision* model (0.0092 and 0.0073). Although the posterior mean of the perceived noise slightly exceeds the actual noise, the 90% credible interval, $(0.0058, 0.0115)$, contains the mean of the

Table 12: Posterior Distribution: Jointly Incorporating Persistence and Precision Misspecification

		(V) Misspecified persistence and precision	
Parameter	Description	Mean	[05, 95]
ρ	persist.	0.8727	[0.8318, 0.9154]
$\hat{\rho}$	persist. (perceived)	0.9723	[0.9647, 0.9803]
σ_u	tech. shock s.d.	0.0060	[0.0055, 0.0066]
σ_ω	noise shock s.d.	0.0066	[0.0044, 0.0088]
$\hat{\sigma}_\omega$	noise shock s.d. (perceived)	0.0087	[0.0058, 0.0115]
log Data Density		1268.54	

Notes: With the random-walk assumption in (17) the variances of the two productivity shocks, σ_ε^2 and σ_η^2, are associated with the variance of Δa_t, denoted by σ_u^2: $\sigma_\varepsilon^2 = (1-\rho)^2\sigma_u^2$ and $\sigma_\eta^2 = \rho\sigma_u^2$. Thus, given ρ and σ_u, both σ_ε and σ_η are pinned down.

actual noise. Thus, the overall fit to data is best when both persistence and signal-precision misspecification are allowed, but the economically robust distortion is overpersistence, while signal-precision misspecification acts as a secondary adjustment margin in explaining U.S. consumption and productivity dynamics.

To further disentangle the individual contributions of each misspecification mechanism, we consider the following two scenarios. First, we fix the parameters obtained from the misspecified persistence model and estimate the remaining parameters; second, we fix the parameters from the misspecified precision model and estimate the remaining parameters. Table 13 (VI) reports the posterior distribution of $\hat{\sigma}_\omega$ when persistence parameters are held fixed, while Table 13 (VII) reports the posterior distribution of $\hat{\rho}$ when signal-precision parameters are held fixed. Importantly, Specification (VI) exhibits a substantial improvement in matching the observed data relative to Table 12, with a log marginal likelihood of 1279.15 versus 1268.54 in the model estimating both misspecifications. In contrast, the log marginal likelihood for Specification (VII) is similar to that of the model estimating both misspecifications. Taken together, the pattern of posterior estimates and data fit indicates that strong overpersistence ($\hat{\rho} > \rho$) constitutes the main mechanism explaining observed fluctuations in consumption.

Table 13: Posterior Distribution: Fixing Overpersistence/Overconfidence

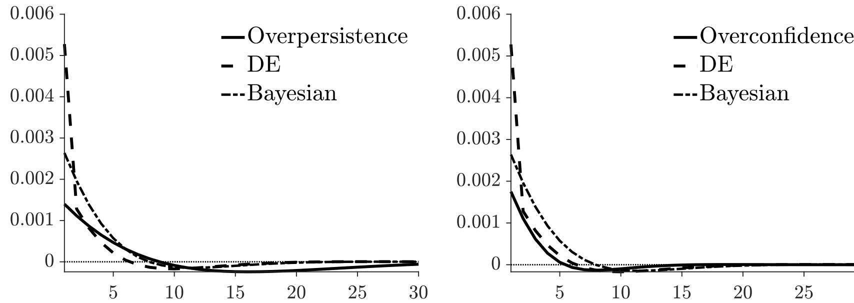
Parameter	Description	(VI) Fixing overpersistence		(VII) Fixing overconfidence	
		Post. Mean	[05, 95]	Mean	[05, 95]
ρ	persist.	0.8814		0.9579	
$\hat{\rho}$	persist. (perc.)	0.9729		0.9730	[0.9640, 0.9827]
σ_u	tech. shock s.d.	0.0061		0.0065	
σ_ω	noise s.d.	0.0082		0.0073	
$\hat{\sigma}_\omega$	noise s.d. (perc.)	0.0106	[0.0097, 0.0115]	0.0092	
log Data Density		1279.15		1267.60	

Notes: With the random-walk assumption in (17), the variances of the two productivity shocks, σ_ε^2 and σ_η^2, are associated with the variance of Δa_t, denoted by σ_u^2: $\sigma_\varepsilon^2 = (1-\rho)^2\sigma_u^2$ and $\sigma_\eta^2 = \rho\sigma_u^2$. Thus, given ρ and σ_u, both σ_ε and σ_η are pinned down.

5.3.2 Belief misspecification versus DE.

In this section, we ask how consumption behavior differs when agents misspecify economic fundamentals (through overpersistence or overconfidence) compared with when agents form expectations diagnostically. The consumption responses to permanent technology and noise shocks are qualitatively similar across these environments, as both mechanisms generate varying degrees of belief overreaction following such shocks in the short run. In contrast, consumption behavior differs markedly following a transitory technology shock.

Figure 3 plots the impulse responses of consumption to transitory technology shocks for overpersistence and diagnostic expectations (Panel (a)) and for overconfidence and diagnostic expectations (Panel (b)). While consumption responds more strongly under DE than under the Bayesian benchmark, the responses under overpersistence and overconfidence are substantially weaker in the short run. Under diagnostic expectations, agents overweight recent favorable news and project it too strongly into the future, so positive realizations generate an amplified consumption response. By contrast, under overpersistence, agents overestimate the persistence of transitory shocks; they interpret the disturbance as bad news, implying weaker future income dynamics, and thus dampening current consumption. Similarly, under overconfidence, if agents overestimate the likelihood of transitory shocks, they view observed improvements as temporary rather than persistent, which lowers expected



(a) Overpersistence and DE (b) Overconfidence and DE

Figure 3: Impulse Responses of Consumption: Transitory Technology Shocks

Notes: The panels show consumption impulse responses to a transitory technology shock. The left panel shows the responses under overpersistence (solid) and under DE (dashed) while the right panel shows the responses under overconfidence (solid) and DE (dashed). All models share the baseline parameters $\rho = 0.8814$, $\sigma_u = 0.0054$, and $\sigma_\omega = 0.0082$. Under overpersistence, we set $\hat{\rho} = 0.9729$ and under overconfidence, we set $\hat{\sigma}_\omega = 0.0041$; the diagnosticity parameter θ is set to 1 under DE.

permanent income and reduces consumption. Thus, this result delivers a sharp theoretical prediction that separates the mechanism of DE from belief misspecification arising from overpersistence and overconfidence.

Taken together, these results suggest that overpersistence is the main behavioral force in the data, while responses to transitory shocks provide the clearest margin along which belief misspecification can be distinguished from diagnostic expectations.

6 Conclusion

This paper studies three behavioral models of belief formation—diagnostic expectations, overconfidence in private signals, and overpersistence—using a laboratory forecasting experiment in a signal-extraction environment. Participants repeatedly forecast the current and next-period realizations of an AR(1) process, are fully informed of the DGP and its parameter values, and are paid for forecast accuracy, giving the FIRE benchmark its best chance to succeed. In the signal treatments, participants receive a noisy private signal of the current realization consistent with a signal extraction framework, while in the no-signal treatments they do not. The no-signal treatments permit a closer comparison with Afrouzi et al. (2023) and help distinguish overpersistence from overconfidence as a source of belief overreaction.

Across all treatments, we find robust evidence that individual beliefs overreact to new information, regardless of whether the DGP is i.i.d. or persistent and whether signals are available or not. Among the three models considered, overpersistence is most consistent with the experimental data, especially because overreaction persists even in the no-signal treatments, where overconfidence is absent by construction. Comparing treatments with and without signals also supports the information-friction hypothesis of Coibion and Gorodnichenko (2015): individual forecasts overreact, whereas consensus forecasts underreact when agents observe different private signals. We then incorporate these same mechanisms into a simple macroeconomic consumption model and find that overpersistence generates sizable distortions in consumption dynamics and fits U.S. consumption and productivity data better than overconfidence, diagnostic expectations, or rational Bayesian learning. Taken together, the experimental and macro evidence point to overpersistence as the most empirically relevant mechanism among the models considered here.

A limitation of both the experiment and the macroeconomic application is that beliefs do not feed back into fundamentals. By contrast, expectations and outcomes are jointly determined in many macroeconomic and financial environments, a feature emphasized in the "learning-to-forecast" literature (Marimon and Sunder, 1993; Hommes et al., 2005; Adam, 2007; Assenza et al., 2021; Anufriev et al., 2022; Evans et al., 2025; Hanaki and Takahasi, 2026, among others). Embedding the behavioral mechanisms studied here in such an environment is a promising direction for future research.

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 (2025). Overconfidence in private information explains biases in professional forecasts. *Journal of Monetary Economics 155*(Supplement), 103839.

Afrouzi, H., S. Kwon, A. Landier, Y. Ma, and D. Thesmar (2023). Overreaction in expectations: Evidence and theory. *Quarterly Journal of Economics 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*. 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.

Beshears, J., J. J. Choi, A. Fuster, D. Laibson, and B. C. Madrian (2013). What goes up must come down? experimental evidence on intuitive forecasting. *The American economic review 103*(3), 570–574.

Bianchi, F., C. Ilut, and H. Saijo (2024). Diagnostic business cycles. *Review of Economic Studies 91*(1), 129–162.

Bianchi, F., C. Ilut, and H. Saijo (2026). Smooth diagnostic expectations. *Review of Economic Studies (forthcoming)*.

Blanchard, O. J., J.-P. L'Huillier, and G. Lorenzoni (2013). News, noise, and fluctuations: An empirical exploration. *American Economic Review 103*(7), 3045–3070.

Bordalo, P., K. Coffman, N. Gennaioli, and A. Shleifer (2016). Stereotypes. *Quarterly Journal of Economics 131*(4), 1753–1794.

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), 2748–2782.

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 (2025). Imperfect exchange rate expectations. *Review of Economics and Statistics 107*(5), 1406–1423.

Cao, D. and J.-P. L'Huillier (2018). Technological revolutions and the three great slumps: A medium-run analysis. *Journal of Monetary Economics 96*, 93–108.

Chen, D. L., M. Schonger, and C. Wickens (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), 2644–2678.

Epley, N. and T. Gilovich (2006). The anchoring-and-adjustment heuristic: Why the adjustments are insufficient. *Psychological Science 17*(4), 311–318.

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.

Garcìa-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. (2026). Mis-specified forecasts and myopia in an estimated New Keynesian model. *American Economic Journal: Macroeconomics 18*(2), 188–226.

Hanaki, N. and Y. Takahasi (2026). An experiment on a multi-period beauty contest game. *Experimental Economics*, forthcoming. Published online 15 May 2026.

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. (2012). Did the US consumer overreact? A test of rational expectations. *Economics Letters 116*(2), 207–209.

L'Huillier, J.-P., S. R. Singh, and D. Yoo (2021). Diagnostic expectations and macroeconomic volatility. *Mimeo*.

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*, 103750.

Tversky, A. and D. Kahneman (1974). Judgment under uncertainty: Heuristics and biases. *Science 185*(4157), 1124–1131.

[FOR ONLINE PUBLICATION]

Appendix for

Overreaction in Expectations under Signal Extraction: Experimental Evidence

John Duffy, Nobuyuki Hanaki, Donghoon Yoo

July 25, 2026

A Derivations and further results

A.1 Derivation of the Kalman Gain

The state and observation equations are

$$x_t = \mu + \rho x_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2),$$

and

$$s_t^i = x_t + \omega_t^i, \qquad \omega_t^i \sim \mathcal{N}(0, \sigma_\omega^2).$$

Let P^- and P^+ denote the steady-state prior and posterior mean-squared forecast errors. The Kalman gain is

$$\lambda = \frac{P^-}{P^- + \sigma_\omega^2},$$

and the posterior variance is

$$P^+ = (1 - \lambda) P^-.$$

These satisfy the prediction-step recursion

$$P^- = \rho^2 P^+ + \sigma_\varepsilon^2. \tag{22}$$

Substituting P^+ and letting $\delta \equiv \sigma_\varepsilon^2/\sigma_\omega^2$ and $p \equiv P^-/\sigma_\varepsilon^2$, the steady-state condition reduces to

$$\delta p^2 + p\Big[1 - \rho^2 - \delta\Big] - 1 = 0. \tag{23}$$

Taking the positive root gives

$$p = \frac{\delta - (1-\rho^2) + \sqrt{\Big[\delta - (1-\rho^2)\Big]^2 + 4\delta}}{2\delta}.$$

Next, observe that

$$\Big[\delta - (1-\rho^2)\Big]^2 + 4\delta = \delta^2 A^2,$$

where

$$A \equiv \sqrt{\frac{(1-\rho^2)^2}{\delta^2} + 1 + \frac{2}{\delta} + \frac{2\rho^2}{\delta}}.$$

Thus,

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

Since

$$\lambda = \frac{\delta p}{\delta p + 1},$$

it follows that

$$\lambda = \frac{\delta - (1-\rho^2) + \delta A}{2 + \delta - (1-\rho^2) + \delta A}. \tag{24}$$

■

A.2 Comparative Statics of the Kalman Gain

Using the expression for the Kalman gain, (4) derived above, the comparative statics follow by differentiating λ with respect to δ and ρ. We first consider δ. Direct differentiation yields

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

where

$$\frac{\partial A}{\partial \delta} = -\frac{1}{A}\left[\frac{(1-\rho^2)^2}{\delta^3} + \frac{1+\rho^2}{\delta^2}\right].$$

Substituting this expression and using the definition of A gives

$$1 + A + \delta\frac{\partial A}{\partial \delta} = 1 + A - \frac{1}{A}\left[\frac{(1-\rho^2)^2}{\delta^2} + \frac{1+\rho^2}{\delta}\right] = 1 + \frac{1 + \frac{1+\rho^2}{\delta}}{A} > 0.$$

Therefore,

$$\frac{\partial \lambda}{\partial \delta} > 0. \tag{25}$$

Next, consider ρ. Direct differentiation gives

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

where

$$\frac{\partial A}{\partial \rho} = \frac{2\rho}{A}\left[\frac{1}{\delta} - \frac{1-\rho^2}{\delta^2}\right].$$

Thus,

$$2\rho + \delta\frac{\partial A}{\partial \rho} = 2\rho\left[1 + \frac{1 - \frac{1-\rho^2}{\delta}}{A}\right].$$

Finally, note that

$$A^2 - \left(\frac{1-\rho^2}{\delta} - 1\right)^2 = \frac{4}{\delta} > 0.$$

Thus,

$$A > \left|\frac{1-\rho^2}{\delta} - 1\right|.$$

This implies

$$A + \left(1 - \frac{1-\rho^2}{\delta}\right) > 0,$$

and therefore

$$1 + \frac{1 - \frac{1-\rho^2}{\delta}}{A} > 0.$$

Since $\rho \geq 0$, it follows that

$$\frac{\partial \lambda}{\partial \rho} \geq 0, \tag{26}$$

with strict inequality for $\rho > 0$.

Thus, the Kalman gain is increasing in the signal-to-noise ratio δ and weakly increasing in the persistence parameter ρ.

A.3 Diagnostic Kalman filter

A diagnostic agent i forms expectations about the unobserved fundamental x_t according to

$$F_t^{i,\theta} x_t = F_t^i x_t + \theta\Big(F_t^i x_t - F_{t-1}^i x_t\Big), \tag{27}$$

where $F_t^{i,\theta}[\cdot]$ denotes the diagnostic expectation operator, and θ is the diagnosticity parameter.

The rational expectations benchmark forecast, obtained from the signal-extraction problem, is given by

$$F_t^i x_t = F_{t-1}^i x_t + \lambda\Big(s_t^i - F_{t-1}^i x_t\Big). \tag{28}$$

Substituting this expression into eq. (27) gives

$$F_t^{i,\theta} x_t = F_{t-1}^i x_t + \lambda\Big(s_t^i - F_{t-1}^i x_t\Big) + \theta\Big(F_{t-1}^i x_t + \lambda\Big(s_t^i - F_{t-1}^i x_t\Big) - F_{t-1}^i x_t\Big). \tag{29}$$

We can further simplify it to obtain the diagnostic Kalman filter:

$$F_t^{i,\theta} x_t = F_{t-1}^i x_t + (1+\theta)\lambda\Big(s_t^i - F_{t-1}^i x_t\Big). \tag{30}$$

A.4 Predictions without signal extraction

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

$$x_t = \mu + \rho x_{t-1} + \varepsilon_t, \tag{31}$$

where ρ is the 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 $\{x_{t-1}, x_{t-2}, x_{t-3}, ..., x_{t-39}\}$. Over 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 experiment, σ_ε is set to 3, while μ is set to 75 when $\rho = 0$ and 30 when $\rho = 0.6$.)

A.4.1 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 (31) based on the information available at the time of her forecast, $\mathcal{I}_{t-1} = \{x_{t-1}, x_{t-2}, x_{t-3}, ..., x_{t-39}\}$.

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

$$F^i_{t-1}x_t = \rho x_{t-1},$$

$$F^i_{t-1}x_{t+1} = \rho^2 x_{t-1}.$$

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

$$F_{t-1}^i x_t = \hat{\rho} x_{t-1},$$
$$F_{t-1}^i x_{t+1} = \hat{\rho}^2 x_{t-1}.$$

Diagnostic expectations. Suppose that a diagnostic agent i forms forecasts of this random variable. The diagnostic forecasts $F_{t-1}^{i,\theta} x_t$ and $F_{t-1}^{i,\theta} x_{t+1}$ are given by

$$F_{t-1}^{i,\theta} x_t = F_{t-1}^i x_t + \theta \left(F_{t-1}^i x_t - F_{t-2}^i x_t \right) = \rho x_{t-1} + \theta \rho \epsilon_{t-1},$$
$$F_{t-1}^{i,\theta} x_{t+1} = F_{t-1}^i x_{t+1} + \theta \left(F_{t-1}^i x_{t+1} - F_{t-2}^i x_{t+1} \right) = \rho^2 x_{t-1} + \theta \rho^2 \epsilon_{t-1}.$$

B Derivations of Forecast Revision Coefficients

B.1 Bayesian learning

Individual forecasts with signals. The state and observations equations are

$$x_t = \mu + \rho x_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2),$$

and

$$s_t^i = x_t + \omega_t^i, \qquad \omega_t^i \sim \mathcal{N}(0, \sigma_\omega^2),$$

where $\mu = 0$ (for simplicity) and $0 \leq \rho < 1$. The agent's information set in period t is:

$$\mathcal{I}_t^i = \{s_t^i, x_{t-1}, x_{t-2}, ..., x_0\}.$$

Before observing the private noise signal s_t^i, the agent's prior is:

$$x_t | x_{t-1} \sim N(\rho x_{t-1}, \sigma_\varepsilon^2).$$

Define the signal-to-noise ratio, $\delta = \sigma_\varepsilon^2 / \sigma_\omega^2$. The agent solves a standard Gaussian signal-extraction problem:

$$F_t^i x_t = \rho x_{t-1} + \lambda (s_t^i - \rho x_{t-1}),$$

where $\lambda = \delta/(1+\delta)$. Thus, a higher signal-to-noise ratio δ puts more weight on the private signal. We then define the forecast error ($FE_t^i = x_t - F_t^i x_t$) and the forecast revision ($FR_t^i = F_t^i x_t - F_{t-1}^i x_t$):

$$FE_t^i = x_t - \Big(\rho x_{t-1} + \lambda (s_t^i - \rho x_{t-1})\Big) = (1-\lambda)\varepsilon_t - \lambda \omega_t^i,$$

and

$$FR_t^i = \rho x_{t-1} + \lambda(s_t^i - \rho x_{t-1}) - \rho \left(\rho x_{t-2} + \lambda(s_{t-1}^i - \rho x_{t-2})\right) = \rho(1-\lambda)\varepsilon_{t-1} - \rho\lambda\omega_{t-1}^i + \lambda\varepsilon_t + \lambda\omega_t^i.$$

Then, the forecast error regression coefficient β is given by

$$\beta = \frac{Cov(FE_t^i, FR_t^i)}{Var(FR_t^i)} = \frac{Cov((1-\lambda)\varepsilon_t - \lambda\omega_t^i, \lambda\varepsilon_t + \lambda\omega_t^i)}{Var\left(\rho(1-\lambda)\varepsilon_{t-1} - \rho\lambda\omega_{t-1}^i + \lambda\varepsilon_t + \lambda\omega_t^i\right)},$$

as only the ϵ_t and ω_t terms are shared in the covariance. Since $\lambda = \sigma_\varepsilon^2/(\sigma_\varepsilon^2 + \sigma_\omega^2)$, the covariance term is given by

$$\left(1 - \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_\omega^2}\right)\left(\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_\omega^2}\right)\sigma_\varepsilon^2 - \left(\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_\omega^2}\right)^2 \varepsilon_\omega^2 = 0,$$

whereas the variance term gives

$$\sigma_\varepsilon^2 \frac{\rho^2 + \delta}{1 + \delta}.$$

Therefore, under Bayesian learning, $\beta = 0$.

Consensus forecasts with signals. Consider the consensus prediction where the individual noise averages out. The consensus forecasts $\bar{F}_t x_t$ and $\bar{F}_{t-1} x_t$ are given by

$$\bar{F}_t x_t = \rho x_{t-1} + \lambda \varepsilon_t,$$

and

$$\bar{F}_{t-1} x_t = \rho^2 x_{t-2} + \rho\lambda\varepsilon_{t-1}.$$

We now compute the consensus forecast error ($\overline{FE}_t$) and forecast revision ($\overline{FR}_t$):

$$\overline{FE}_t = x_t - \bar{F}_t x_t = \rho x_{t-1} + \varepsilon_t - (\rho x_{t-1} + \lambda\varepsilon_t) = (1-\lambda)\varepsilon_t,$$

and

$$\overline{FR}_t = \bar{F}_t x_t - \bar{F}_{t-1} x_t = \rho^2 x_{t-2} + \rho\varepsilon_{t-1} + \lambda\varepsilon_t - (\rho^2 x_{t-2} + \rho\lambda\varepsilon_{t-1}) = \rho(1-\lambda)\varepsilon_{t-1} + \lambda\varepsilon_t.$$

Then, the consensus forecast error regression coefficient β^c is given by

$$\beta^c = \frac{Cov(\overline{FE}_t, \overline{FR}_t)}{Var(\overline{FR}_t)} = \frac{Cov((1-\lambda)\varepsilon_t, \rho(1-\lambda)\varepsilon_{t-1} + \lambda\varepsilon_t)}{Var(\rho(1-\lambda)\varepsilon_{t-1} + \lambda\varepsilon_t)} = \frac{\lambda(1-\lambda)}{\rho^2(1-\lambda)^2 + \lambda^2}.$$

Since $\lambda = \delta/(1+\delta)$,

$$\beta^c = \frac{\delta}{\rho^2 + \delta^2} > 0.$$

Forecasts without signals. Suppose agents do not receive a private noisy signal and their information set is given by $\mathcal{I}_t = \{x_{t-1}, x_{t-2}, ..., x_0\}$. The forecasts $F_{t-1}x_t$ and $F_{t-2}x_t$ are given by

$$F_{t-1}x_t = \rho x_{t-1},$$

and

$$F_{t-2}x_t = \rho^2 x_{t-2}.$$

Define the forecast error (FE_t) and the forecast revision (FR_t):

$$FE_t = x_t - F_{t-1}x_t = x_t - \rho x_{t-1} = \varepsilon_t,$$

and

$$FR_t = F_{t-1}x_t - F_{t-2}x_t = \rho x_{t-1} - \rho^2 x_{t-2} = \rho\varepsilon_{t-1}.$$

Therefore, the covariance between FE_t and FR_t is zero. Thus, for $\rho \neq 0$, the forecast error regression coefficient β is zero. For $\rho = 0$, the revision is degenerate, so the regression is not identified.

B.2 Overconfidence

Individual forecasts with signals. Suppose agent i misspecifies the signal-to-noise ratio as $\hat{\delta} \equiv \sigma_\varepsilon^2/\hat{\sigma}_\omega^2$, where $\hat{\sigma}_\omega^2$ denotes the perceived noise variance. Since $\hat{\lambda} = \hat{\delta}/(1+\hat{\delta})$, we have $\hat{\lambda} > \lambda$ whenever $\hat{\delta} > \delta$, and $\hat{\lambda} < \lambda$ whenever $\hat{\delta} < \delta$. Therefore, overconfidence corresponds to the case in which the agent underestimates the noise variance, or equivalently,

$$\hat{\delta} > \delta.$$

Then, the individual forecast $F_t^i x_t$ is:

$$F_t^i x_t = \rho x_{t-1} + \hat{\lambda}(s_t^i - \rho x_{t-1}) = \rho x_{t-1} + \hat{\lambda}(\varepsilon_t + \omega_t^i),$$

and the lagged forecast $F_{t-1}^i x_t$ is given by

$$F_{t-1}^i x_t = \rho \left(\rho x_{t-2} + \hat{\lambda}(\varepsilon_{t-1} + \omega_{t-1}^i)\right) = \rho^2 x_{t-2} + \rho\hat{\lambda}(\varepsilon_{t-1} + \omega_{t-1}^i).$$

Thus, the individual forecast error (FE_t^i) and the forecast revision (FR_t^i) are given by

$$FE_t^i = x_t - F_t^i x_t = x_t - \rho x_{t-1} - \hat{\lambda}(\varepsilon_t + \omega_t^i) = (1-\hat{\lambda})\varepsilon_t - \hat{\lambda}\omega_t^i,$$

and

$$FR_t^i = \rho x_{t-1} + \hat{\lambda}(\varepsilon_t + \omega_t^i) - \rho^2 x_{t-2} - \rho\hat{\lambda}(\varepsilon_{t-1} + \omega_{t-1}^i) = \rho(1-\hat{\lambda})\varepsilon_{t-1} - \rho\hat{\lambda}\omega_{t-1}^i + \hat{\lambda}\varepsilon_t + \hat{\lambda}\omega_t^i.$$

The covariance between FE_t^i and FR_t^i is $Cov((1-\hat{\lambda})\varepsilon_t - \hat{\lambda}\omega_t^i, \hat{\lambda}\varepsilon_t + \hat{\lambda}\omega_t^i)$:

$$Cov(FE_t^i, FR_t^i) = (1-\hat{\lambda})\hat{\lambda}\sigma_\varepsilon^2 - \hat{\lambda}^2\sigma_\omega^2.$$

Similarly, the variance of FR_t^i is:

$$Var(FR_t^i) = \rho^2(1-\hat{\lambda})^2\sigma_\varepsilon^2 + \rho^2\hat{\lambda}^2\sigma_\omega^2 + \hat{\lambda}^2\sigma_\varepsilon^2 + \hat{\lambda}^2\sigma_\omega^2.$$

Recall that the Kalman gain and its perceived counterpart depend on the signal-to-noise ratio via

$$\lambda = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_\omega^2} = \frac{\delta}{1+\delta}, \qquad \hat{\lambda} = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \hat{\sigma}_\omega^2} = \frac{\hat{\delta}}{1+\hat{\delta}},$$

where $\delta \equiv \sigma_\varepsilon^2/\sigma_\omega^2$ and $\hat{\delta} \equiv \sigma_\varepsilon^2/\hat{\sigma}_\omega^2$. Equivalently, $1-\lambda = 1/(1+\delta)$ and $1-\hat{\lambda} = 1/(1+\hat{\delta})$. Substituting $\hat{\lambda} = \hat{\delta}/(1+\hat{\delta})$ and $\sigma_\omega^2 = \sigma_\varepsilon^2/\delta$ into the expressions for $\text{Cov}(FE_t^i, FR_t^i)$ and $\text{Var}(FR_t^i)$ above and simplifying, the forecast error regression coefficient is

$$\beta = \frac{\text{Cov}(FE_t^i, FR_t^i)}{\text{Var}(FR_t^i)} = \frac{-\hat{\delta}(\hat{\delta}-\delta)}{\delta\rho^2 + \hat{\delta}^2(\delta + 1 + \rho^2)},$$

which is negative whenever $\hat{\delta} > \delta$.

Consensus forecasts with signals. We now consider the consensus prediction where the individual noise averages out. The consensus forecasts $\bar{F}_t x_t$ and $\bar{F}_{t-1}x_t$ are given by

$$\bar{F}_t x_t = \rho x_{t-1} + \hat{\lambda}\varepsilon_t,$$

and

$$\bar{F}_{t-1}x_t = \rho^2 x_{t-2} + \rho\hat{\lambda}\varepsilon_{t-1}.$$

The consensus forecast error ($\overline{FE}_t$) and the forecast revision ($\overline{FR}_t$) are given by

$$\overline{FE}_t = x_t - \bar{F}_t x_t = (1-\hat{\lambda})\varepsilon_t,$$

and

$$\overline{FR}_t = \bar{F}_t x_t - \bar{F}_{t-1}x_t = \rho\varepsilon_{t-1} + \hat{\lambda}\varepsilon_t - \rho\hat{\lambda}\varepsilon_{t-1} = \hat{\lambda}\varepsilon_t + \rho(1-\hat{\lambda})\varepsilon_{t-1}.$$

The consensus forecast error regression coefficient β^c is:

$$\beta^c = \frac{Cov(\overline{FE}_t, \overline{FR}_t)}{Var(\overline{FR}_t)} = \frac{(1-\hat{\lambda})\hat{\lambda}}{\hat{\lambda}^2 + \rho^2(1-\hat{\lambda})^2} > 0.$$

Therefore, overconfidence predicts a positive consensus forecast error coefficient $\beta^c > 0$.

B.3 Overpersistence.

Individual forecasts with signals. Suppose agent i misperceives the persistence of the DGP: $\hat{\rho} \neq \rho$. With overpersistence, we consider the case where the agent overestimates the degree of persistence in x_t:

$$\hat{\rho} > \rho.$$

With signals, the individual forecast $F_t^i x_t$ is:

$$F_t^i x_t = \hat{\rho} x_{t-1} + \lambda(s_t^i - \hat{\rho} x_{t-1}) = ((1-\lambda)\hat{\rho} + \lambda\rho)x_{t-1} + \lambda\varepsilon_t + \lambda\omega_t^i,$$

and the lagged forecast $F_{t-1}^i x_t$ is given by

$$F_{t-1}^i x_t = \hat{\rho}((1-\lambda)\hat{\rho} + \lambda\rho)x_{t-2} + \hat{\rho}\lambda\varepsilon_{t-1} + \hat{\rho}\lambda\omega_{t-1}^i.$$

Since the individual forecast error is $FE_t^i = x_t - F_t^i x_t$, we have

$$FE_t^i = x_t - (1-\lambda)\hat{\rho} x_{t-1} - \lambda x_t - \lambda\omega_t^i = (1-\lambda)(x_t - \hat{\rho} x_{t-1}) - \lambda\omega_t^i.$$

Similarly, the individual forecast revision is $FR_t^i = F_t^i x_t - F_{t-1}^i x_t$, we have

$$FR_t^i = (1-\lambda)\hat{\rho} x_{t-1} + \lambda x_t + \lambda\omega_t^i - \left((1-\lambda)\hat{\rho}^2 x_{t-2} + \hat{\rho}\lambda x_{t-1} + \hat{\rho}\lambda\omega_{t-1}^i\right).$$

Rearranging it, we get

$$FR_t^i = (1-\lambda)\hat{\rho}(x_{t-1} - \hat{\rho}x_{t-2}) + \lambda(x_t - \hat{\rho}x_{t-1}) + \lambda\omega_t^i - \hat{\rho}\lambda\omega_{t-1}^i.$$

Define $u_t \equiv x_t - \hat{\rho}x_{t-1}$. Then, the covariance between FE_t^i and FR_t^i is given by

$$(1-\lambda)^2\hat{\rho}Cov(u_t, u_{t-1}) + (1-\lambda)\lambda Var(u_t) - \lambda^2\sigma_\omega^2.$$

Since $Cov(u_t, u_{t-1}) = -(\hat{\rho}-\rho)V_x(1-\hat{\rho}\rho)$ and $Var(u_t) = \sigma_\varepsilon^2 + (\hat{\rho}-\rho)^2V_x$ where $V_x \equiv Var(x_t)$, we have

$$Cov(FE_t^i, FR_t^i) = -(1-\lambda)(\hat{\rho}-\rho)V_x\left((1-\lambda)\hat{\rho}(1-\hat{\rho}\rho) - \lambda(\hat{\rho}-\rho)\right).$$

Furthermore, given that $\lambda = \delta/(1+\delta)$, we get

$$Cov(FE_t^i, FR_t^i) = -\frac{1}{(1+\delta)^2}(\hat{\rho}-\rho)V_x\left(\hat{\rho}(1-\hat{\rho}\rho) - \delta(\hat{\rho}-\rho)\right).$$

To compute $Var(FR_t^i)$, we first show that:

$$FR_t^i = -(\hat{\rho}-\rho)\left((1-\lambda)\hat{\rho} + \lambda\rho\right)x_{t-2} + \left((1-\lambda)\hat{\rho} - (\hat{\rho}-\rho)\lambda\right)\varepsilon_{t-1} + \lambda\varepsilon_t + \lambda\omega_t^i - \hat{\rho}\lambda\omega_{t-1}^i.$$

Since $x_{t-2}, \varepsilon_t, \varepsilon_{t-1}, \omega_t^i, \omega_{t-1}^i$ are mutually independent,

$$Var(FR_t^i) = (\hat{\rho}-\rho)^2((1-\lambda)\hat{\rho} + \lambda\rho)^2V_x + ((1-\lambda)\hat{\rho} - (\hat{\rho}-\rho)\lambda)^2\sigma_\varepsilon^2 + \lambda^2\sigma_\varepsilon^2 + \lambda^2(1+\hat{\rho}^2)\sigma_\omega^2,$$

such that

$$Var(FR_t^i) = \frac{V_x}{(1+\delta)^2}\left\{(\hat{\rho}+\delta\rho)^2(\hat{\rho}-\rho)^2 + (1-\rho^2)\left((\hat{\rho}-\delta(\hat{\rho}-\rho))^2 + \delta^2 + \delta(1+\hat{\rho}^2)\right)\right\}.$$

Therefore, the regression coefficient β is:

$$\beta = \frac{Cov(FE_t^i, FR_t^i)}{Var(FR_t^i)} = \frac{-(\hat{\rho}-\rho)\left(\hat{\rho}(1-\hat{\rho}\rho)-\delta(\hat{\rho}-\rho)\right)}{(\hat{\rho}+\delta\rho)^2(\hat{\rho}-\rho)^2+(1-\rho^2)\left[(\hat{\rho}-\delta(\hat{\rho}-\rho))^2+\delta^2+\delta(1+\hat{\rho}^2)\right]}.$$

The sign of β is determined by the sign of $\hat{\rho}-\rho$ and by the sign of $\hat{\rho}(1-\hat{\rho}\rho)-\delta(\hat{\rho}-\rho)$. Define $\hat{\rho}_+$ as the positive root of $\hat{\rho}(1-\hat{\rho}\rho)-\delta(\hat{\rho}-\rho)=0$. Equivalently, for $0<\rho<1$,

$$\hat{\rho}_+ = \frac{1-\delta+\sqrt{(1-\delta)^2+4\delta\rho^2}}{2\rho}.$$

Since $\hat{\rho}(1-\hat{\rho}\rho)-\delta(\hat{\rho}-\rho)$ is positive for $\hat{\rho}\in(\rho,\hat{\rho}_+)$ and negative for $\hat{\rho}>\hat{\rho}_+$, the sign of β can be summarized as follows:

$$\hat{\rho}<\rho \Longrightarrow \beta>0,$$
$$\rho<\hat{\rho}<\hat{\rho}_+ \Longrightarrow \beta<0,$$
$$\hat{\rho}>\hat{\rho}_+ \Longrightarrow \beta>0.$$

Thus, moderate overestimation of persistence generates a negative β, whereas sufficiently strong overestimation of persistence reverses the sign and generates a positive β. If $0<\rho<\hat{\rho}<1$ and $\delta\leq 1$, then $\hat{\rho}_+\geq 1$, so overpersistence gives $\beta<0$. If $\delta>1$, then $\hat{\rho}_+<1$, so the sign can reverse.

When $\rho=0$, the second term simplifies to $\hat{\rho}(1-\delta)$. Therefore, for $\hat{\rho}>0$, the sign of β is determined by the sign of $-(1-\delta)$: $\beta>0$ if $\delta>1$, $\beta<0$ if $\delta<1$, and $\beta=0$ if $\delta=1$.[19]

Consensus forecasts with signals. We now consider deriving predictions for the consensus forecasts under overpersistence. Let $u_t \equiv x_t - \hat{\rho}x_{t-1}$. The consensus forecast error ($\overline{FE}_t$) and

[19] For $\delta=1$, the sign of β is fully determined by the sign of $-(\hat{\rho}-\rho)$ for $0<\rho,\hat{\rho}<1$. Thus, overpersistence predicts a negative forecast error regression coefficient. When, $\rho=0$ with $\delta=1$, $\beta=0$ for any $\hat{\rho}$.

forecast revision ($\overline{FR}_t$) are given by

$$\overline{FE}_t = x_t - \bar{F}_t x_t = (1-\lambda)u_t,$$

and

$$\overline{FR}_t = \bar{F}_t x_t - \bar{F}_{t-1} x_t = (1-\lambda)\hat{\rho}u_{t-1} + \lambda u_t.$$

Then, the covariance between $\overline{FE}_t$ and $\overline{FR}_t$ is given by

$$(1-\lambda)^2\hat{\rho}Cov(u_t, u_{t-1}) + (1-\lambda)\lambda Var(u_t).$$

Since $Cov(u_t, u_{t-1}) = -(\hat{\rho}-\rho)V_x(1-\hat{\rho}\rho)$ and $Var(u_t) = \sigma_\varepsilon^2 + (\hat{\rho}-\rho)^2 V_x$ where $V_x \equiv Var(x_t)$, we have

$$Cov(\overline{FE}_t, \overline{FR}_t) = -(1-\lambda)(\hat{\rho}-\rho)V_x\left((1-\lambda)\hat{\rho}(1-\hat{\rho}\rho) - \lambda(\hat{\rho}-\rho)\right) + (1-\lambda)\lambda\sigma_\varepsilon^2.$$

Furthermore, given that $\lambda = \delta/(1+\delta)$, we get

$$Cov(\overline{FE}_t, \overline{FR}_t) = \frac{V_x}{(1+\delta)^2}\left[\delta(1-\rho^2) - (\hat{\rho}-\rho)\left(\hat{\rho}(1-\hat{\rho}\rho) - \delta(\hat{\rho}-\rho)\right)\right].$$

Similarly, we have

$$\overline{FR}_t = (1-\lambda)\hat{\rho}(\varepsilon_{t-1} - (\hat{\rho}-\rho)x_{t-2}) + \lambda(\varepsilon_t - (\hat{\rho}-\rho)x_{t-1}).$$

Given that $x_{t-1} = \rho x_{t-2} + \varepsilon_{t-1}$, we have

$$\overline{FR}_t = (1-\lambda)\hat{\rho}\varepsilon_{t-1} - (1-\lambda)\hat{\rho}(\hat{\rho}-\rho)x_{t-2} + \lambda\varepsilon_t - \lambda(\hat{\rho}-\rho)\rho x_{t-2} - \lambda(\hat{\rho}-\rho)\varepsilon_{t-1}.$$

Thus,

$$\overline{FR}_t = -(\hat{\rho}-\rho)((1-\lambda)\hat{\rho} + \lambda\rho)x_{t-2} + ((1-\lambda)\hat{\rho} - \lambda(\hat{\rho}-\rho))\varepsilon_{t-1} + \lambda\varepsilon_t.$$

Since $x_{t-2}, \varepsilon_t, \varepsilon_{t-1}$ are mutually independent, we have

$$Var(\overline{FR}_t) = V_x((1-\lambda)\hat{\rho} + \lambda\rho)^2(\hat{\rho}-\rho)^2 + \sigma_\varepsilon^2((1-\lambda)\hat{\rho} - \lambda(\hat{\rho}-\rho))^2 + \lambda^2\sigma_\varepsilon^2$$

Further simplification with $(1-\lambda) = 1/(1+\delta)$, $\lambda = \delta/(1+\delta)$, $\sigma_\varepsilon^2 = (1-\rho^2)V_x$ leads to

$$Var(\overline{FR}_t) = \frac{V_x}{(1+\delta)^2}\left[(\hat{\rho}+\delta\rho)^2(\hat{\rho}-\rho)^2 + (1-\rho^2)\left((\hat{\rho}-\delta(\hat{\rho}-\rho))^2 + \delta^2\right)\right].$$

Therefore, the consensus forecast regression coefficient β^c is given by

$$\beta^c = \frac{Cov(\overline{FE}_t, \overline{FR}_t)}{Var(\overline{FR}_t)} = \frac{\delta(1-\rho^2) - (\hat{\rho}-\rho)\left(\hat{\rho}(1-\hat{\rho}\rho) - \delta(\hat{\rho}-\rho)\right)}{(\hat{\rho}+\delta\rho)^2(\hat{\rho}-\rho)^2 + (1-\rho^2)\left((\hat{\rho}-\delta(\hat{\rho}-\rho))^2+\delta^2\right)}.$$

The sign of β^c cannot solely be determined by the sign of $\hat{\rho}-\rho$. Specifically, with overpersistence ($\hat{\rho} > \rho$),

$$\beta^c > 0 \Longleftrightarrow \delta > \frac{(\hat{\rho}-\rho)\hat{\rho}(1-\rho\hat{\rho})}{(1-\rho^2)+(\hat{\rho}-\rho)^2}.$$

Since the consensus forecast removes idiosyncratic signal noise, the remaining consensus revision is informative about the true innovation ε_t, which creates a positive error-revision covariance. Overpersistence creates a negative force, but it has to be strong enough, or δ has to be small enough, to dominate the positive information-friction force.[20]

Forecasts without signals. Consider now the case without signals such that agents do not receive a private noisy signal. The forecasts $F_{t-1}x_t$ and $F_{t-2}x_t$ are given by

$$F_{t-1}x_t = \hat{\rho}x_{t-1},$$

and

$$F_{t-2}x_t = \hat{\rho}^2 x_{t-2}.$$

[20]For $\delta = 1$, the sign of β^c is determined by the sign of $1-\rho^2-(\hat{\rho}-\rho)\rho(1-\hat{\rho}^2)$: $\beta^c > 0$ for $0 \leq \rho < \hat{\rho} < 1$.

Define the forecast error (FE_t) and the forecast revision (FR_t):

$$FE_t = x_t - \hat{\rho} x_{t-1} = u_t,$$

and

$$FR_t = \hat{\rho} x_{t-1} - \hat{\rho}^2 x_{t-2} = \hat{\rho}(x_{t-1} - \hat{\rho} x_{t-2}) = \hat{\rho} u_{t-1}.$$

Thus, the forecast error regression coefficient β is:

$$\beta = \frac{Cov(FE_t, FR_t)}{Var(FR_t)} = \frac{Cov(u_t, \hat{\rho} u_{t-1})}{Var(\hat{\rho} u_{t-1})},$$

where $u_t \equiv x_t - \hat{\rho} x_{t-1}$. Since $Cov(u_t, \hat{\rho} u_{t-1}) = -\hat{\rho}(\hat{\rho} - \rho) V_x (1 - \hat{\rho}\rho)$ and $Var(\hat{\rho} u_{t-1}) = V_x \hat{\rho}^2 \left((1 - \rho^2) + (\hat{\rho} - \rho)^2\right)$, we have

$$\beta = \frac{-(\hat{\rho} - \rho)(1 - \hat{\rho}\rho)}{\hat{\rho}\left((1 - \rho^2) + (\hat{\rho} - \rho)^2\right)} < 0,$$

with $\hat{\rho} > \rho$. Therefore, overpersistence predicts $\beta < 0$.

B.4 Diagnostic expectations

Individual forecasts with signals. Under diagnostic expectations with signals, the agent distorts the conditional mean:

$$F_t^{i,\theta} x_t = \rho x_{t-1} + (1 + \theta)\lambda(s_t^i - \rho x_{t-1}) = \rho x_{t-1} + \tilde{\theta}(\varepsilon_t + \omega_t^i)$$

where $\tilde{\theta} \equiv (1 + \theta)\lambda > \lambda$. We also assume that $\rho > 0$. The lagged expectation is then given by

$$F_{t-1}^{i,\theta} x_t = \rho \left(\rho x_{t-2} + (1 + \theta)\lambda(s_{t-1}^i - \rho x_{t-2})\right) = \rho^2 x_{t-2} + \rho\tilde{\theta}(\varepsilon_{t-1} + \omega_{t-1}^i).$$

The (diagnostic) forecast error ($FE_t^{i,\theta}$) and the (diagnostic) forecast revision ($FR_t^{i,\theta}$) are given by

$$FE_t^{i,\theta} = x_t - F_t^{i,\theta} x_t = (1-\hat{\theta})\varepsilon_t - \tilde{\theta}\omega_t^i,$$

and

$$FR_t^{i,\theta} = F_t^{i,\theta} x_t - F_{t-1}^{i,\theta} x_t = \rho(1-\tilde{\theta})\varepsilon_{t-1} + \tilde{\theta}(\varepsilon_t + \omega_t^i) - \rho\tilde{\theta}\omega_{t-1}^i.$$

The covariance between $FE_t^{i,\theta}$ and $FR_t^{i,\theta}$ is $Cov((1-\tilde{\theta})\varepsilon_t - \tilde{\theta}\omega_t^i, \tilde{\theta}\varepsilon_t + \tilde{\theta}\omega_t^i)$:

$$Cov(FE_t^{i,\theta}, FR_t^{i,\theta}) = (1-\tilde{\theta})\tilde{\theta}\sigma_\varepsilon^2 - \tilde{\theta}^2\sigma_\omega^2.$$

Similarly, the variance of $FR_t^{i,\theta}$ is:

$$Var(FR_t^{i,\theta}) = \rho^2(1-\tilde{\theta})^2\sigma_\varepsilon^2 + \tilde{\theta}^2\sigma_\varepsilon^2 + \tilde{\theta}^2\sigma_\omega^2 + \rho^2\tilde{\theta}^2\sigma_\omega^2.$$

Since $Cov(FE_t^{i,\theta}, FR_t^{i,\theta})$ is simplified to

$$(1-\tilde{\theta})\tilde{\theta}\sigma_\varepsilon^2 - \tilde{\theta}^2\frac{1-\lambda}{\lambda}\sigma_\varepsilon^2 = \tilde{\theta}\sigma_\varepsilon^2\left(1-\tilde{\theta}-\frac{\tilde{\theta}}{\lambda}+\tilde{\theta}\right) = \tilde{\theta}\sigma_\varepsilon^2\left(1-\frac{\tilde{\theta}}{\lambda}\right) < 0,$$

and $\sigma_\omega^2 = \sigma_\varepsilon^2/\delta$, the forecast error regression coefficient is given by

$$\beta = \frac{-\theta(1+\theta)\lambda}{\rho^2(1-(1+\theta)\lambda)^2 + (1+\theta)^2\lambda^2(1+(1+\rho^2)/\delta)} < 0, \quad \text{if } \rho > 0.$$

Consensus forecasts with signals. The consensus forecasts $\bar{F}_t^\theta x_t$ and $\bar{F}_{t-1}^\theta x_t$ are given by

$$\bar{F}_t^\theta x_t = \rho x_{t-1} + \tilde{\theta}\varepsilon_t,$$

and

$$\bar{F}_{t-1}^\theta x_t = \rho^2 x_{t-2} + \rho\tilde{\theta}\varepsilon_{t-1}.$$

The consensus forecast error ($\overline{FE}_t^\theta$) and the forecast revision ($\overline{FR}_t^\theta$) are given by

$$\overline{FE}_t^\theta = x_t - \bar{F}_t^\theta x_t = (1-\tilde{\theta})\varepsilon_t,$$

and

$$\overline{FR}_t^\theta = \bar{F}_t^\theta x_t - \bar{F}_{t-1}^\theta x_t = \rho x_{t-1} + \tilde{\theta}\varepsilon_t - \rho^2 x_{t-2} - \rho\tilde{\theta}\varepsilon_{t-1} = \tilde{\theta}\varepsilon_t + \rho(1-\tilde{\theta})\varepsilon_{t-1}.$$

The consensus forecast error regression coefficient β^c is:

$$\beta^c = \frac{Cov(\bar{FE}_t^\theta, \bar{FR}_t^\theta)}{Var(\bar{FR}_t^\theta)} = \frac{(1-\tilde{\theta})\tilde{\theta}}{\tilde{\theta}^2 + \rho^2(1-\tilde{\theta})^2}.$$

Thus, $\beta^c > 0$ if and only if $\tilde{\theta} < 1$; $\beta^c = 0$ if and only if $\tilde{\theta} = 1$; $\beta^c < 0$ if and only if $\tilde{\theta} > 1$. Since $\tilde{\theta} = (1+\theta)\lambda$, the threshold in terms of diagnosticity θ is $(1+\theta)\lambda = 1$ such that the threshold value of θ is:

$$\theta = \frac{1}{\lambda} - 1.$$

Since $\lambda = \delta/(1+\delta)$, we have $\beta^c > 0$ if and only if $\theta < 1/\delta$; $\beta^c = 0$ if and only if $\theta = 1/\delta$; $\beta^c < 0$ if and only if $\theta > 1/\delta$. Diagnosticity dominates the information friction if $\theta > 1/\delta$. If $\theta < 1/\delta$, the information-friction force dominates and $\beta^c > 0$.

Forecasts without signals. Consider now the case without signals such that diagnostic agents do not receive a private noisy signal: the forecasts $F_{t-1}^\theta x_t$ and $F_{t-2}^\theta x_t$ are given by

$$F_{t-1}^\theta x_t = F_{t-1}x_t + \theta(F_{t-1}x_t - F_{t-2}x_t) = \rho x_{t-1} + \theta(\rho x_{t-1} - \rho^2 x_{t-2}) = \rho x_{t-1} + \theta\rho\varepsilon_{t-1},$$

and

$$F_{t-2}^\theta x_t = F_{t-2}x_t + \theta(F_{t-2}x_t - F_{t-3}x_t) = \rho^2 x_{t-2} + \theta(\rho^2 x_{t-2} - \rho^3 x_{t-3}) = \rho^2 x_{t-2} + \theta\rho^2\varepsilon_{t-2}.$$

The diagnostic forecast error (FE_t^θ) and the diagnostic forecast revision (FR_t^θ) are given by

$$FE_t^\theta = x_t - F_{t-1}^\theta x_t = x_t - \rho x_{t-1} - \theta\rho\varepsilon_{t-1} = \varepsilon_t - \theta\rho\varepsilon_{t-1},$$

and

$$FR_t^\theta = F_{t-1}^\theta x_t - F_{t-2}^\theta x_t = \rho x_{t-1} + \theta\rho\varepsilon_{t-1} - \rho^2 x_{t-2} - \theta\rho^2\varepsilon_{t-2} = \rho(1+\theta)\varepsilon_{t-1} - \theta\rho^2\varepsilon_{t-2}.$$

Then, the forecast error regression coefficient under diagnostic expectations is:

$$\beta^\theta = \frac{Cov(FE_t^\theta, FR_t^\theta)}{Var(FR_t^\theta)} = \frac{Cov(-\theta\rho\varepsilon_{t-1}, \rho(1+\theta)\varepsilon_{t-1})}{Var(\rho(1+\theta)\varepsilon_{t-1} - \theta\rho^2\varepsilon_{t-2})} = \frac{-\theta(1+\theta)}{(1+\theta)^2 + \theta^2\rho^2}.$$

Therefore, diagnostic expectations ($\theta > 0$) predict $\beta < 0$.

C Additional analyses

C.1 Estimating β excluding outliers

Table C.1 reports the results of regressions based on eq. (7) in the main text. The All columns show the results reported in Tables 3 and 4, while Median, Trim 1%, Trim 5% are to check whether the results are affected by outliers by conducting median regression and regressions by excluding extreme 1% and 5% observations. The results are consistent across specifications.

Table C.1: Predicting Forecast Errors with Forecast Revisions Excluding Outliers

(a) Signal Treatment

	$\rho = 0$				$\rho = 0.6$			
	All	Median	Trim 1%	Trim 5%	All	Median	Trim 1%	Trim 5%
β^p	−0.314	−0.338	−0.331	−0.296	−0.319	−0.346	−0.285	−0.255
SE	0.071	0.029	0.055	0.056	0.045	0.026	0.050	0.062
t-stat	−4.40	−11.70	−5.99	−5.30	−7.04	−13.25	−5.66	−4.02
R^2	0.161	0.099	0.160	0.123	0.175	0.095	0.141	0.096
N	1029	1029	994	855	1036	1036	1004	866

(b) No-Signal Treatment

	$\rho = 0$				$\rho = 0.6$			
	All	Median	Trim 1%	Trim 5%	All	Median	Trim 1%	Trim 5%
β^p	−0.681	−0.635	−0.693	−0.504	−0.384	−0.319	−0.346	−0.247
SE	0.108	0.112	0.049	0.103	0.136	0.042	0.127	0.128
t-stat	−6.26	−12.98	−6.17	−4.91	−2.82	−7.59	−2.73	−1.93
R^2	0.221	0.094	0.204	0.094	0.127	0.055	0.091	0.040
N	1000	1000	969	826	1003	1003	969	866

Notes: The table reports coefficients from individual-level forecast error on forecast revision regression as in equation (7). Columns "All" show the estimation results from the individual-level fixed effect panel regression (β^p). Standard errors are clustered by both forecaster and time. Columns "Median" show the results of median regression. Columns "Trim 1%" ("Trim 5%") show the results where we drop top 1% and bottom 1% (top 5% and bottom 5%) observations of the forecast errors and forecast revisions.

C.2 Estimating ρ and μ excluding outliers

Table C.2 reports the results of individual fixed-effect panel regressions for estimating ρ and μ. The columns "All" reproduce the results reported in Tables 5 and 6, columns "Trim 1%" and "Trim 5%" report the results based on dropping outliers, defined as those extreme 1% and 5% (both top and bottom) observations, respectively. In $\rho = 0$ treatments, the estimated values of ρ becomes smaller as we drop more outliers, but they are all significantly different from the true value, i.e., $\rho = 0$. For $\rho = 0.6$ treatments, while the estimated values are not significantly different from the true value in the signal treatment, they are significantly different for No-Signal treatment.

Table C.2: Forecast-Implied Subjective Persistence Excluding Outliers

Signal Treatments

	$\rho = 0$ ($\mu = 75$)			$\rho = 0.6$ ($\mu = 30$)		
	All	Trim 1%	Trim 5%	All	Trim 1%	Trim 5%
$\hat{\rho}^p$	0.539	0.484	0.384	0.678	0.668	0.590
	(0.381, 0.697)	(0.335, 0.632)	(0.253, 0.515)	(0.587, 0.769)	(0.595, 0.741)	(0.510, 0.670)
SE	0.081	0.076	0.067	0.046	0.037	0.041
t-stat	6.68	6.38	5.75	14.64	17.93	14.47
$\hat{\mu}^p$	34.5	38.6	46.0	23.9	24.6	30.3
	(22.9, 46.1)	(27.4,49.9)	(35.8,56.2)	(17.4, 30.4)	(19.5,29.7)	(24.5,36.2)
SE	5.911	5.739	5.203	3.304	2.610	2.982
t-stat	5.84	6.73	8.84	7.23	9.43	10.17
$\hat{\mu}^p/(1-\hat{\rho}^p)$	74.87	74.79	74.72	74.17	74.08	73.90
N	1063	1039	910	1069	1046	931

No-Signal Treatments

	$\rho = 0$ ($\mu = 75$)			$\rho = 0.6$ ($\mu = 30$)		
	All	Trim 1%	Trim 5%	All	Trim 1%	Trim 5%
$\hat{\rho}^p$	0.637	0.577	0.441	0.756	0.759	0.704
	(0.469, 0.805)	(0.423, 0.731)	(0.283, 0.599)	(0.624, 0.887)	(0.664, 0.853)	(0.644, 0.765)
SE	0.086	0.079	0.081	0.067	0.048	0.031
t-stat	7.43	7.34	5.47	11.26	15.70	22.79
$\hat{\mu}^p$	27.0	31.5	41.6	18.2	18.0	22.0
	(17.8, 36.2)	(22.0,41.0)	(31.1, 52.2)	(8.9, 27.5)	(10.8,25.1)	(17.5,26.5)
SE	4.673	4.860	5.404	4.750	3.636	2.289
t-stat	5.78	6.48	7.71	3.83	4.94	9.60
$\hat{\mu}^p/(1-\hat{\rho}^p)$	74.43	74.42	74.53	74.49	74.51	74.26
N	1049	1026	924	1036	1007	925

Notes: Columns "All" show the estimation results from the individual-level fixed effect panel regression. Standard errors are clustered by both forecaster and time. Columns "Trim 1%" ("Trim 5%") show the results where we drop top 1% and bottom 1% (top 5% and bottom 5%) observations of the forecast errors and forecast revisions.

C.3 Source of overreaction: Results based on an alternative specification to estimate $\hat{\lambda}$

Here we estimate $\hat{\lambda}$ based on

$$F_t^i x_t = \text{const} + \gamma F_{t-1}^i x_t + \lambda s_t + u_t^i.$$

instead of eq. (12), and using it together with the implied persistence $\hat{\rho}$ to recover the forecast-implied signal-to-noise ratio $\hat{\delta}$. Table C.3 shows that the results do not change much by this alternative specification.

Table C.3: Source of Overreaction

		$\rho = 0$	$\rho = 0.6$
Only overpersistence	$(\hat{\rho} > \rho)$	10 (38%)	14 (48%)
Only overconfidence	$(\hat{\delta} > 1)$	1 (4%)	6 (21%)
Both		14 (54%)	6 (21%)
Neither		1 (4%)	3 (10%)
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* describe 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.

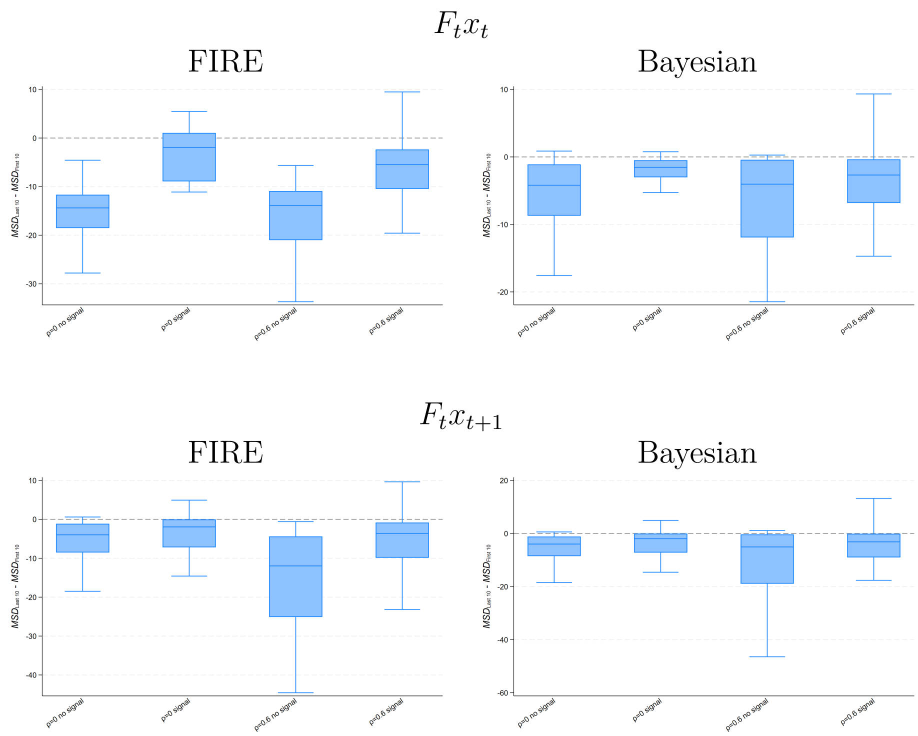
C.4 Participants' learning

This subsection analyze participants' learning by comparing the mean squared deviation of their forecasts from FIRE and Bayesian benchmark between the first and the last 10 periods.

Figure C.1 shows the distributions of changes in the mean squared deviation (MSD) from FIRE (left) and Bayesian (right) benchmark between the first and the last 10 periods of the experiment. The box shows inter-quartile range with the line inside representing the median. The lower and the upper whiskers shows the lowest and the highest non-outlier values.[21] The

[21] One outlier is dropped in each case.

Figure C.1: Changes in MSD from the benchmark between the first and the last 10 periods



Notes: The box shows inter-quartile range with the line inside representing the median. The lower and the upper whiskers show the lowest and the highest non-outlier values.

MSDs are significantly lower in the last 10 periods compared to the first 10 periods in all the treatments at the 5% significance level, except for $F_t x_t$ in Signal $\rho = 0$ with FIRE.[22] Thus, we observe significant evidence of subjects learning to forecast rationally as they gain experience with clear feedback in a stable environment.

C.5 Distribution of realized value of ε and ω

Figure C.2 shows the histogram of the realized values of ε during 40 periods of the experiment. The dotted line represents the $N(0, 3)$ benchmark. Because we have held these realizations to be the same across participants and treatments, all the subjects faced the same outcomes. The mean and the standard deviation are -1.121 and 2.461. The realized distribution is

[22] Based on t-test using the individual difference in MSD between the first and the last 10 periods as independent observation. Outliers are dropped.

significantly different from $N(0,3)$ according to the two-sided Kolmogorov-Smirnov Test ($p = 0.032$).

Figure C.2: Distribution of realized values of ε

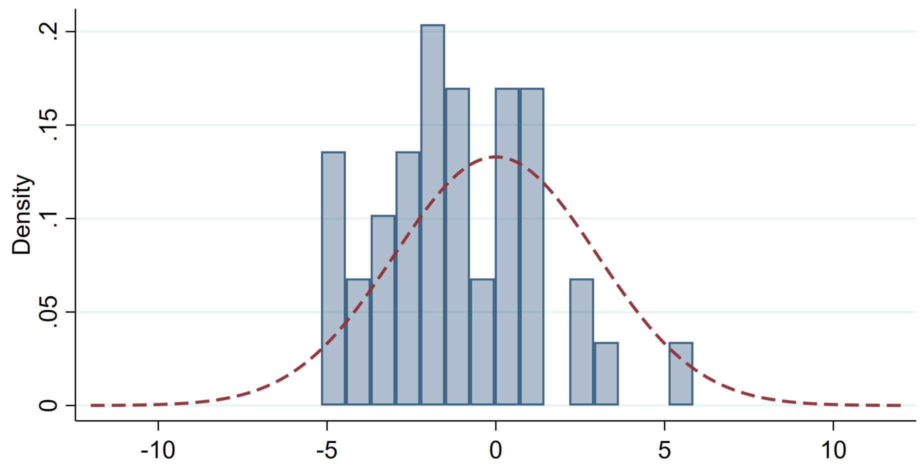
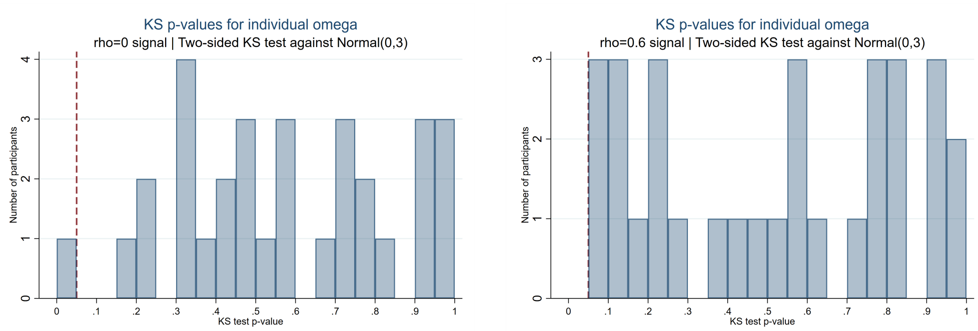


Figure C.3 shows the distribution of p-values testing, at the subject level, whether realized distribution of ω follows $N(0,3)$. There was one participant whose realized values of ω differed significantly from $N(0,3)$ at the 5% significance level in $\rho = 0$, but for all the other participants they were not. The mean p-values for $\rho = 0$ and $\rho = 0.6$ were 0.58 and 0.59, respectively.

Figure C.3: Distribution of p-values testing realized distribution of ω

(1) $\rho = 0$ (1) $\rho = 0.6$



This finite-sample feature does not affect our treatment comparisons, because all subjects faced the same realized aggregate history. More importantly, our behavioral benchmarks are computed conditional on the realized history and, where relevant, the realized signal. Thus

Table C.4: Predicting Forecast Errors with Forecast Revisions With Period-Fixed Effect

	Signal Treatment		No-Signal Treatment	
	$\rho = 0$	$\rho = 0.6$	$\rho = 0$	$\rho = 0.6$
β^p	-0.616	-0.520	-0.553	-0.412
SE	0.033	0.032	0.044	0.056
t-stat	-18.42	-16.08	-12.64	-7.35
R^2	0.746	0.688	0.908	0.857
N	1029	1036	1000	1003

Notes: The table reports coefficients from individual-level forecast error on forecast revision regression as in equation (7) with period-fixed effect.

the Bayesian forecast, the signal-news term, and the alternative behavioral models are not evaluated against the asymptotic distribution alone, but against the information actually available to subjects in the experiment.

Nevertheless, the realized shock path could matter for individual-level estimates over a short horizon. To account for this, we have re-estimated β^p, $r\hat{h}o^p$, and $\hat{\mu}^p$ with individual- as well as with period-fixed effects, by adding period dummies for each period.

Table C.4 reports the results of forecast-error forecast-revision regressions with period-fixed effects. The estimated values of β remain negative and significant in all the treatments.

Table C.5 shows the results for forecast-implied subjective persistence. While the estimated values of $\hat{\rho}^p$ remain significantly different from the true value for treatments with $\rho = 0$, they are no longer significantly different from its true values for $\rho = 0.6$.

C.6 Distribution of estimated values of $\hat{\rho}$ and $\hat{\lambda}$

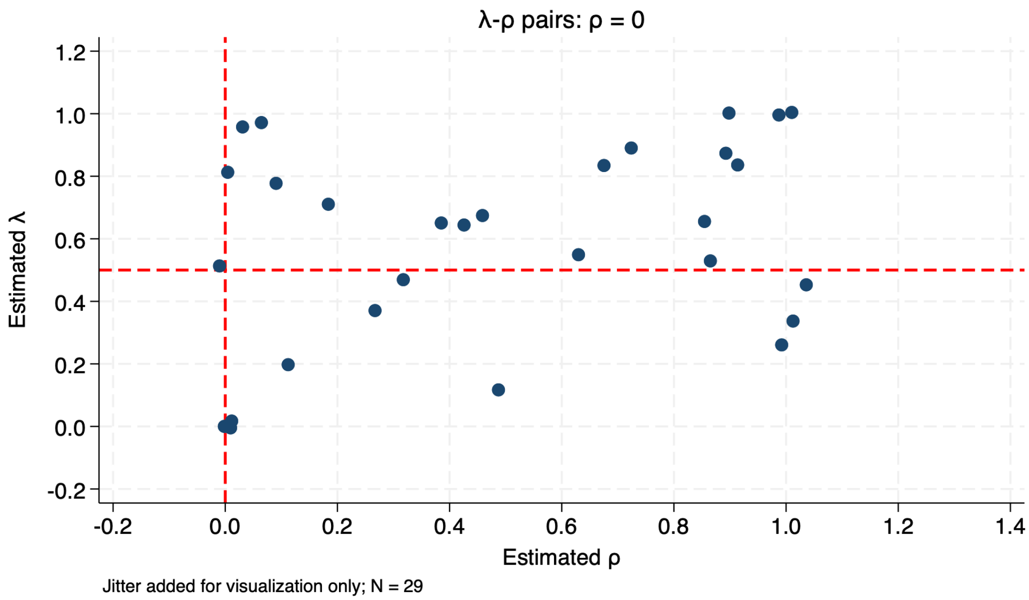
This subsection reports the distribution of the estimated weight on signal, $\hat{\lambda}$, and the forecast implied subjective persistence $\hat{\rho}$. Figure C.4 shows the distribution of estimated values of $\hat{\rho}$ and $\hat{\lambda}$ pairs. Each point represents a participant. Red dashed lines in the figure show the Bayesian benchmark. Note that overconfidence implies that $(\hat{\lambda})$ exceeds the Bayesian benchmark. Thus, an estimated value of $(\hat{\lambda})$ above the Bayesian benchmark is consistent

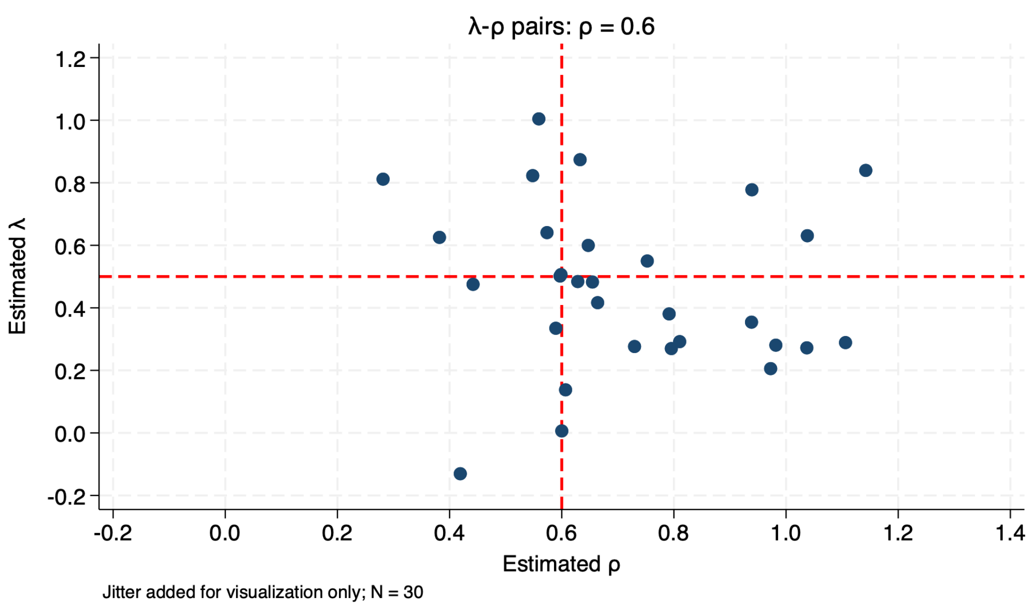
Table C.5: Forecast-Implied Subjective Persistence with Period-Fixed Effect

	Signal Treatment		No-Signal Treatment	
	$\rho = 0$	$\rho = 0.6$	$\rho = 0$	$\rho = 0.6$
$\hat{\rho}^p$	0.504	0.642	0.550	0.635
	(0.357,0.651)	(0.545,0.739)	(0.387,0.712)	(0.393,0.877)
SE	0.075	0.050	0.083	0.124
t-stat	6.721	12.960	6.635	5.138
$\hat{\mu}^p$	37.71	27.02	33.64	28.37
	(26.68,48.73)	(19.67,34.37)	(23.34,43.94)	(9.18,47.56)
SE	5.624	3.749	5.254	9.792
t-stat	6.70	7.21	6.40	2.90
$\hat{\mu}^p/(1-\hat{\rho}^p)$	76.02	75.47	74.70	77.63
N	1063	1069	1049	1036

with overconfidence.

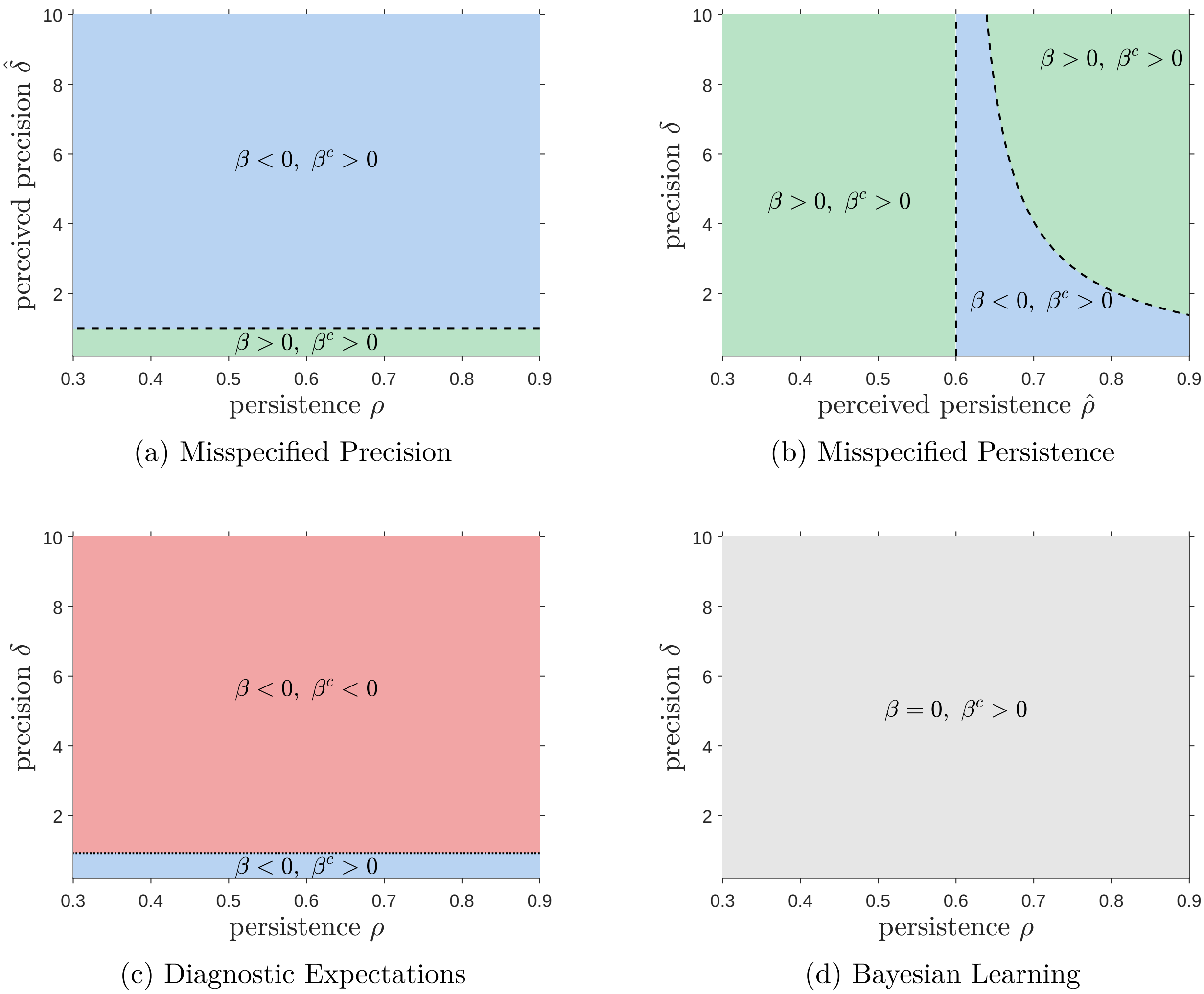
Figure C.4: Distribution of estimated values of $\{\hat{\rho},\, \hat{\lambda}\}$





Notes: A point represent a participant. Red dashed lines correspond to the Bayesian benchmark.

D Supplementary Figures and Tables



(a) Misspecified Precision

(b) Misspecified Persistence

(c) Diagnostic Expectations

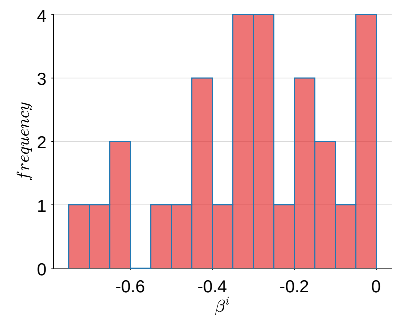
(d) Bayesian Learning

Figure D.1: Forecast Revision Coefficients in the Theory

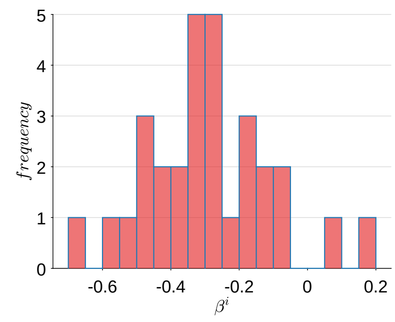
Notes: The figure reports sign predictions for the individual forecast-revision coefficient, β, and the corresponding consensus forecast-revision coefficient, β^c. The four regions correspond to the possible sign combinations of (β, β^c), as indicated within each panel. Panel (a) reports the prediction under misspecified precision as a function of actual persistence, ρ, and perceived signal precision, $\hat{\delta}$, with the true signal precision fixed at $\delta = 1$; Panel (b) reports the prediction under misspecified persistence as a function of perceived persistence, $\hat{\rho}$, and signal precision, δ, with true persistence fixed at $\rho = 0.6$; Panel (c) reports the prediction for diagnostic expectations as a function of persistence, ρ, and signal precision, δ, with the diagnosticity parameter fixed at $\theta = 1.1$; Panel (d) reports the prediction from Bayesian-learning as a function of ρ and δ; under Bayesian learning, $\beta = 0$ and $\beta^c > 0$ throughout the plotted region. The dashed lines indicate the sign-switching boundary for the individual coefficient, $\beta = 0$, whereas the dotted lines indicate the corresponding boundary for the consensus coefficient, $\beta^c = 0$.

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

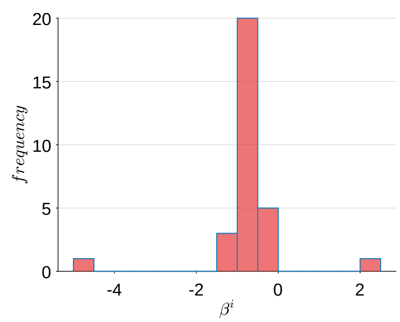
(a) *With Signals* ($\rho = 0$)



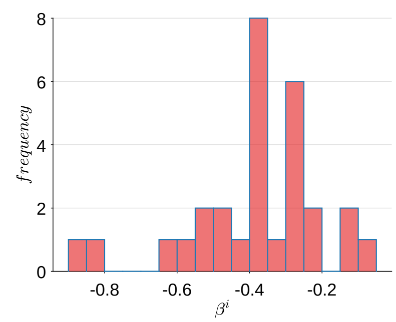
(b) *With Signals* ($\rho = 0.6$)



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



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



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

Table D.1: Prior Distribution

Parameter	Description	Distribution	Mean	Std. dev
ρ	persist.	Beta	0.6	0.2
$\hat{\rho}$	persist. (perc.)	Beta	0.6	0.2
σ_u	tech. shock s.d.	Inv. Gamma	0.01	0.01
σ_ω	noise shock s.d.	Inv. Gamma	0.01	0.01
$\hat{\sigma}_\omega$	noise shock s.d. (perc.)	Inv. Gamma	0.01	0.01
θ	diagnosticity	Uniform	1.00	1.15

Notes: The table reports the prior distribution of structural parameters in the estimation procedure.

E 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.