BANCA D'ITALIA
EUROSISTEMA

Temi di discussione

(Working Papers)

Overconfident forecasters and the impact of inflation information: evidence from a randomized survey experiment

by Filippo Natoli and Sharath Sonti

April 2026

Number 1532

BANCA D'ITALIA
EUROSISTEMA

Temi di discussione

(Working Papers)

Overconfident forecasters and the impact of inflation information: evidence from a randomized survey experiment

by Filippo Natoli and Sharath Sonti

Number 1532 - April 2026

ISSN 2281-3950 (online)

Designed by the Printing and Publishing Division of Banca d'Italia

OVERCONFIDENT FORECASTERS AND THE IMPACT OF INFLATION INFORMATION: EVIDENCE FROM A RANDOMIZED SURVEY EXPERIMENT

by Filippo Natoli* and Sharath Sonti**

Abstract

We exploit a randomized information intervention in a quarterly survey of Italian firms to study how access to recent inflation data affects deviations from the full-information rational expectations (FIRE) benchmark in managers' inflation forecasts. Treated firms receive the latest inflation reading before submitting their forecasts and, relative to non-treated firms, display *less underreaction* at the consensus level and *less overreaction* at the individual level, moving forecasts closer to the FIRE benchmark. A model that combines noisy information with *overconfidence in private information* provides the best overall fit to the data, outperforming alternative frameworks featuring diagnostic expectations, internal cognitive constraints, or over-extrapolation. Intuitively, with noisy information, average forecasts react sluggishly to news, so an informative public signal speeds up aggregate updating and reduces consensus underreaction. With overconfidence, managers overweight private signals and the public signal shifts weight away from private information, attenuating individual overreaction.

JEL Classification: C53, D83, D84, E31, E37.

Keywords: overconfidence, information treatment, belief updating, firms, inflation expectations.

DOI: 10.32057/0.TD.2026.1532

* Bank of Italy.

** UC Berkeley.

1 Introduction[1]

Inflation expectations play an important role in shaping the economic decisions of households and firms. While most macroeconomic models assume full-information rational expectations (FIRE), this assumption has been challenged based on survey data on expectations of firms, households, and professional forecasters. In a seminal study, Coibion and Gorodnichenko (2015) find that on average, inflation forecasts of professional forecasters underreact to news – that is, they adjust their forecasts less than they should in response to new information –, and explain this by maintaining rational expectations but departing from the full information assumption (Sims, 2003; Woodford, 2001). More recent work has shown that individual inflation forecasts overreact to news (Bordalo et al., 2020), which cannot be explained by standard models of information frictions. To simultaneously explain sluggishness in average forecasts and excess sensitivity in individual forecasts, the literature has resorted to combining information rigidities with other frictions/biases, such as noisy memory (Sung, 2025), over-extrapolation (Angeletos et al., 2021), and diagnostic expectations (Bordalo et al., 2020).

Despite these advances, a clear consensus on the right model of expectation formation has yet to emerge. Concurrently, a growing body of literature leveraging randomized controlled trials has found that reminding firms of recent macroeconomic data – such as the latest inflation rate – at the time of the interview has a significant impact on the macroeconomic expectations they report (Coibion et al., 2022, 2023). Regarding inflation expectations, this effect is particularly strong in periods of low inflation (Weber et al., 2025), where providing such information also reduces disagreement about future inflation (Bartiloro et al., 2019; Coibion et al., 2019). However, much of this literature relies on newly established surveys, which makes these settings ill-suited to assessing whether sustained access to information improves the efficiency with which agents incorporate new data into their forecasts. These gaps motivate two central questions for this paper: Does enhanced access to information

[1]We are grateful to Marco Bottone, Luisa Carpinelli, Yuriy Gorodnichenko, Chen Lian, Giacomo Mangiante, Peter Maxted, Emi Nakamura, Tiziano Ropele, Sergio Santoro, Jon Steinsson, Alex Tagliabracci and seminar participants at UC Berkeley Macroeconomics seminars for their insightful comments. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Bank of Italy or ESCB. Address for correspondence: Bank of Italy, filippo.natoli@bancaditalia.it; UC Berkeley, sharath_sonti@berkeley.edu.

about recent macroeconomic dynamics bring forecasts closer to the rational expectations benchmark? And what does this reveal about the underlying model of expectation formation?

We answer these questions using a randomized information experiment in the Bank of Italy's Survey of Inflation and Growth Expectations (SIGE), a quarterly survey of managers of Italian firms. Until Q2 2012, questions about respondents' views on future consumer price inflation always included a preamble referencing the latest inflation release. Starting in Q3 2012, however, this preamble was omitted for roughly one-third of the interviewees. In this environment, our strategy involves testing the FIRE hypothesis by regressing firms' forecast errors on forecast revisions *separately* for treated and non-treated firms. We estimate these regressions both at the consensus level (as in Coibion and Gorodnichenko, 2015) and at the individual level (as in Bordalo et al., 2020).

We find that non-treated firms exhibit consensus underreaction and individual overreaction (as in previous studies based on professional forecasters' expectations). By contrast, the treatment drives the two estimated coefficients markedly toward zero. At the consensus level, treated firms display *less underreaction*, which is consistent with standard models of information frictions. At the individual level, we find that compared to untreated firms, treated firms display *less overreaction*, and we show that this allows us to distinguish between competing models that feature departures from individual rationality. Our individual-level finding is not driven by treated firms mechanically anchoring their forecasts to the provided inflation release, by the treatment merely attenuating measurement-error bias or by the information treatment differentially enabling managers to learn about structural features of the inflation process (e.g., persistence or the long-run mean). Overall, our findings suggest that providing timely information on recent inflation at the time of the survey enables firms to update their beliefs more efficiently, moving their forecasts closer to the full-information rational expectations benchmark.

To distinguish among competing models of departures from individual rationality, we embed these mechanisms in a common noisy-information framework and conduct a formal model-selection exercise. Specifically, we start from a noisy-information environment in which each firm observes an idiosyncratic private signal about inflation, while treated firms additionally observe a common public signal. We then incorporate four leading departures from individual rationality in this setting: internal cognitive constraints/noisy memory (Sung, 2025), overconfidence in private information (Daniel et al., 1998), diagnostic expectations (Bordalo et al., 2020), and over-extrapolation (Angeletos et al., 2021). For each mechanism, we derive closed-form expressions for the consensus- and individual-level error-on-revision coefficients as functions of the model parameters. We then estimate each model's parameters in an overidentified GMM exercise that targets the treated and non-treated coefficients and the treatment-induced change in forecast dispersion (which disciplines the precision of the public inflation signal) and we assess overall fit using a GMM J-test. Because the number of targeted moments exceeds the number of free parameters, this approach yields a disciplined set of overidentifying restrictions that lets us formally compare competing models.

Our model-selection exercise indicates that a model with *overconfidence in private information* best explains the observed post-treatment reduction in individual-level overreaction. According to this model, excess sensitivity of individual forecasts is due to forecasters over-weighting their private signals relative to what rational agents would do. All else equal, the provision of a public signal leads them to place a smaller weight on their private information when forming their forecasts, thereby reducing overreaction.

The overconfidence model outperforms all competing frameworks that we evaluate. Among them, the model featuring diagnostic-expectations is excluded from consideration because in this model, as shown by Bordalo et al. (2020), the diagnostic distortion scales the Kalman gains on *all* signals (public and private) proportionally, implying that the individual-level error-on-revision coefficient is invariant to the provision of a public signal. As such, it cannot account for the post-treatment reduction in individual overreaction we document. Between the remaining alternatives – internal cognitive constraints and over-extrapolation – neither fits the data well. While the model with internal cognitive constraints matches the non-treated patterns reasonably well, it predicts too large a reduction in consensus underreaction among treated firms because noisy priors lead agents to place disproportionately high weight on the common public signal. It also generates too little attenuation of individual overreaction: since the memory friction limits how much information is carried across periods, excess sensitivity arises from weak priors and is not substantially remedied by better information in the current period. Finally, the over-extrapolation model is decisively rejected because matching the non-treated individual error-revision coefficient at the horizons we study would require a large gap between true and perceived persistence, which is infeasible when inflation is already highly persistent and perceived persistence is bounded above by one.

Our evidence that firms' inflation expectations reflect noisy information together with overconfidence in private signals points to an important role for central bank communication in shaping firms' inflation expectations. When information is noisy, consensus expectations adjust slowly to news; when managers are overconfident, individual forecasts respond too strongly to idiosyncratic signals. A timely public signal about inflation can counteract both forces by speeding up aggregate updating and tempering the weight placed on private information. As a result, communication that delivers clear, salient signals about inflation can improve the efficiency with which firms incorporate new information into their forecasts, with potential implications for inflation persistence through firms' pricing decisions.

Literature Review. This paper contributes to three strands of the literature. First, we add to the large and growing literature that identifies and explains deviations from full-information rational expectations (FIRE) by studying the predictability of forecast errors. This includes evidence from (i) professional forecasters' macroeconomic projections (Angeletos et al., 2021; Bordalo et al., 2020; Coibion & Gorodnichenko, 2012, 2015; Kohlhas & Walther, 2021), (ii) expectations about financial variables - such as stock prices, credit spreads and corporate earnings (Bordalo et al., 2018, 2024; Bouchaud et al., 2019; Greenwood & Shleifer, 2014) and (iii) randomized controlled trials (Afrouzi et al., 2023). We contribute to this literature by studying firm managers' inflation forecasts (rather than professional forecasters') and exploiting a randomized information intervention to document new stylized facts about how information provision affects forecast-error predictability.

Second, we contribute to a growing empirical literature that seeks to understand the expectations formation process and the subsequent decisions of households and firms through randomized information interventions. From a theoretical standpoint, sluggish forecasts are typically explained by noisy information or rational inattention (Coibion & Gorodnichenko, 2015), which can be heterogeneous (Link et al., 2023) and altered through the provision of information in an RCT (Mackowiak & Wiederholt, 2024). Empirically, news about inflation has been shown to significantly influence inflation expectations, especially in low-inflation environments (Coibion et al., 2022, 2023; Rosolia, 2024; Weber et al., 2025). The post-pandemic inflation surge temporarily heightened firms' attention to inflation;[2] as inflation subsided, however, firms' focus on aggregate inflation diminished once again (Mangiante & Tagliabracci, 2025). Providing salient information has proven crucial for belief updating. Reminding agents of recent inflation outcomes reduces disagreement at shorter horizons (Bartiloro et al., 2019; Coibion et al., 2019), whereas highlighting the central banks inflation target lowers disagreement over medium- to long-term horizons (Bottone et al., 2022). We contribute to this literature both empirically and theoretically. Empirically, we leverage the randomized information intervention to identify how providing recent inflation data changes firms' belief updating in response to new information. Theoretically, we embed the treatment in a unified noisy-information framework and incorporate leading mechanisms of departures from individual rationality. This structure delivers closed-form expressions for the error-on-revision coefficients and supports a disciplined comparison of model fit.

Finally, we contribute to the literature that emphasizes the role of overconfidence in private information in explaining forecast anomalies. A seminal contribution by Daniel et al. (1998) shows that investor overconfidence, reinforced by biased self-attribution, jointly generates excess volatility, short-run momentum, long-run reversals, and post-announcement earnings drift. More recently, Broer and Kohlhas (2024) show that forecasters overreact to some public signals but underreact to others – behavior at odds with standard noisy-information models. They explain this empirical fact by extending noisy rational expectations to allow both absolute overconfidence (overstating the precision of one's own signal) and relative overconfidence (understating the precision of others' signals). Most closely related to our paper is Adam et al. (2025), who model the forecasters in the Survey of Professional Forecasters as being overconfident in their private information. We contribute to this literature by documenting overconfidence in private information in firm managers' inflation forecasts, and by using the randomized information intervention to show how the provision of a public signal tempers the weight managers place on their private signals. We also show that in a formal model-selection exercise, the noisy-information-plus-overconfidence specification provides the best overall fit to the data, outperforming alternative leading mechanisms of deviations from individual rationality.

The remainder of the paper proceeds as follows. Section 2 provides an overview of the

[2] This is in line with the island illusion framework (Born et al., 2022), in which firms' expectations about their own output or prices overreact to firm-specific information but underreact to aggregate signals.

SIGE dataset and describes the design of the randomized information treatment. Section 3 details our methodology and reports the primary empirical findings. Section 4 presents targeted robustness checks to evaluate alternative explanations for our main results. Section 5 examines how our empirical evidence aligns with predictions from different theoretical models of expectation formation. Section 6 concludes.

2 Survey data

The dataset we use is the Survey of Inflation and Growth Expectations (SIGE) run by the Bank of Italy since 1999, which surveys a random sample of managers of Italian firms that have more than 50 employees. Initially targeted towards firms in non-financial private services and industry, it expanded in 2013 to include construction firms. Firms are selected to be part of the survey using a stratified design, with strata defined by (i) Sector: industry, non-financial private services, and construction; (ii) Geography: North-West, North-East, Centre, South and Islands; and (iii) Number of Employees: 50-199, 200-999, and 1,000 or more workers. The survey collects qualitative and quantitative data on firms' expectations about inflation, macroeconomic conditions, and their own business outlook - including anticipated prices, demand, investment and hiring plans, and credit availability. Information is gathered primarily through computer-assisted web interviews with the manager at each firm who is most knowledgeable about these topics. The survey also includes repeated questions to ensure consistency over time and occasionally introduces special modules on relevant economic themes. The SIGE is well-suited to our analysis for three reasons: (i) it spans a long time series; (ii) it elicits expectations at multiple horizons, allowing us to construct forecast revisions; and (iii) it embeds a randomized information treatment that provides the most recent inflation release to a subset of firms. We describe each in turn.

From 1999, the SIGE has elicited 1-year-ahead inflation expectations from firm managers. Over time, the questionnaire was expanded to elicit inflation expectations at different horizons: from 2009 Q2, the survey was expanded to elicit firms' year-on-year inflation expectations 2 years from now, from 2010 Q4, the survey was expanded to elicit firms' inflation expectations 6 months ahead[3], from 2014 Q1, the survey was expanded to elicit firms' year-

[3]Six-month-ahead year-on-year inflation refers to the YoY rate for the month or quarter occurring six

on-year inflation expectations 4 years from now. This feature is critical for our analysis, as it allows us to construct forecast revisions, defined as the difference between a firm's current forecast for a given horizon and the forecast it made in the past for the same target date. For example, a firm surveyed in period $t-h$ and again in period t provides two forecasts for inflation at time $t+h$, formed at different points in time. The difference between the firm's time t and time $t-h$ forecasts of inflation at $t+h$ captures how new information observed between $t-h$ and t is reflected in its updated forecast.

Information treatment. A unique feature of the SIGE survey is the inclusion of a randomized information treatment, introduced in 2012 Q3. Before this date, all surveyed firms were provided with information about the most recent realization of inflation in Italy before being asked to report their inflation expectations. Beginning in 2012 Q3, firms were randomly assigned to one of two groups. The control group comprising roughly 33% of the sample received no such information and were simply asked to report their inflation expectations across multiple horizons. In contrast, the treated group continued to receive information about the most recent 12-month inflation rate in Italy. Due to publication lags, this corresponds to the year-on-year inflation rate observed two months before the date of the survey. Treatment assignments are persistent: firms placed in the treatment group continue receiving information on recent inflation until the next reshuffling. These reshufflings occurred in 2012 Q4, 2017 Q2, and 2019 Q4.

To ensure data quality, we restrict the sample to firms who provided at least 20 forecasts in the sample and winsorize inflation expectations at 1-99 percentiles as in Born et al., 2022.[4] Figure 1 plots the resulting distribution of 6-month-ahead inflation forecasts across six different survey quarters, separately for treated and non-treated firms. The plot reveals that treated firms' forecast distributions tend to shift closer to the provided inflation rate (indicated by a vertical dashed line) while also exhibiting lower cross-sectional dispersion relative to the control group. This suggests that the information treatment may have an effect on expectations and reduce disagreement across firms (also shown in Table 2). Appendix Figure A.1 replicates the same structure as Figure 1 but focuses on 1-year ahead inflation

months after the interview. For example, if a firm is surveyed in March 2012, the six-month-ahead YoY rate is the rate of inflation from September 2011 to September 2012.

[4]As shown in Born et al., 2022, the resulting sample is representative of all surveyed firms in terms of geographic location, number of employees, export orientation, and sector.

expectations. As with the 6-month horizon, treated firms' expectations are more tightly clustered around the provided inflation figure. Relative to the one-year horizon, treated firms' six-month-ahead inflation forecasts cluster more tightly around the provided benchmark, consistent with firms viewing recent inflation as a better predictor of future inflation at shorter horizons.

Table 1 contrasts treated and non-treated firms across some of the variables reported in SIGE, such as inflation expectations, firm characteristics, and beliefs. Treated firms report higher inflation expectations, with the treated-control gap narrowing as the forecast horizon lengthens. Consistent with higher inflation expectations, they also expect to make larger price changes over the next 12 months. Otherwise, the two groups appear similar across a broad set of SIGE covariates, including assessments of current economic conditions and own business situation, export intensity, and the perceived sensitivity of business conditions to demand and to future credit conditions.

3 Empirical Analysis

3.1 Model

To empirically test the full-information rational expectations (FIRE) hypothesis in our setting, we rely on the error-on-revision regressions at the consensus level (Coibion & Gorodnichenko, 2015) and individual level (Bordalo et al., 2020). The consensus-level regression is as follows

$$x_{t+h} - x_{t+h|t} = \alpha_C + \beta_C(x_{t+h|t} - x_{t+h|(t-h)}) + \eta_{t+h|t} \tag{3.1}$$

where the dependent variable is the average h quarter-ahead forecast error (constructed as the difference between actual and expected inflation) and the main regressor is the h-quarter revision in the average inflation forecast for quarter $t + h$, i.e., the difference between the average forecast formed at quarter t and the one formed earlier at $t - h$. Similarly, the individual-level regression has the following form:

$$x_{t+h} - x^i_{t+h|t} = \alpha_i + \beta_I(x^i_{t+h|t} - x^i_{t+h|(t-h)}) + \eta^i_{t+h|t} \tag{3.2}$$

where the dependent variable is the h-quarter ahead individual forecast error and the main independent variable is the h-quarter revision in the individual forecast.[5] We estimate equation (3.2) with firm fixed effects to absorb time-invariant heterogeneity - for example, a firm manager's persistent optimism or pessimism. To maintain consistency with the literature that uses professional forecaster data, our baseline forecast horizon is $h = 2$ (six months).[6] Thus, in equation (3.1) the dependent variable is the average six-month-ahead forecast error, while the independent variable is the difference between the average six-month-ahead inflation forecast at quarter t and the average one-year ahead inflation forecast formed at quarter $t-2$. Relative to the large literature that studies deviations from FIRE by examining the predictability of forecast errors from forecast revisions, our specifications differ only in using a two-quarter (rather than a one-quarter) revision window. Since the independent variables in equations (3.1) and (3.2) are known at time t, FIRE implies that the coefficients β_C and β_I should be zero[7] . We estimate both specifications separately by treatment status to assess how the information intervention shapes systematic departures from rational expectations. As noted in Section 2, our sample begins in 2010 Q4-the first quarter in which six-month-ahead inflation forecasts were collected. Finally, to focus on typical business conditions and avoid the regime shift associated with the post-2021 inflation surge, we end the sample in 2020 Q4.[8]

Table 2 reports summary statistics on our main dependent and independent variables for treated and non-treated firms. Panel A reports statistics for the consensus forecast errors and forecast revisions, and Panel B reports statistics for individual forecast errors and forecast revisions. Panel A shows that while both treated and non-treated firms forecast higher inflation relative to realized inflation, non-treated firms make larger forecast errors. This pattern can also be seen in Figures 2 and 3, which plot both realized inflation and average inflation forecasts at the 6-month and 1-year horizons for treated and non-treated firms, respectively. Moreover, note that the consensus forecast errors and forecast revisions

[5]Here $x^i_{t+h|t}$ denotes the forecast of x_{t+h} made by firm i at time t and $x_{t+h|t} = \int x^i_{t+h|t}$ denotes the corresponding average forecast

[6]Coibion and Gorodnichenko (2015) and Bordalo et al. (2020) use a one-quarter revision window, $x_{t+h|t} - x_{t+h|t-1}$ and a baseline horizon of $h = 3$ quarters.

[7]To align with the empirics, the theoretical analysis in Section 5 indexes time in six-month periods; hence the two-quarter revision used here corresponds to a one-period revision in the model in Section 5.

[8]This choice is consistent with recent evidence that information treatments are most informative in low-inflation environments (Weber et al., 2025).

of treated firms are more dispersed than those of non-treated firms.

Panel B of Table 2 reports statistics on individual-level forecasts. First, we observe that there is more dispersion in the forecasts of non-treated firms[9], in line with what we show in Figure 1. We define the non-revision share as the fraction of firm-quarter observations in which the current six-month-ahead inflation forecast differs from the one-year-ahead forecast made two quarters earlier by less than 0.01 percentage points. This statistic is a measure of rigidity in forecasts with respect to new information, and used as a measure of forecaster attentiveness in the literature (Andrade & Le Bihan, 2013). By this metric, treated firms exhibit slightly lower but overall similar rigidity as compared to non-treated firms, suggesting that the two groups are not significantly different in terms of attentiveness to macroeconomic developments.[10] We also report average spell length by treatment status - the mean number of quarters a firm is observed as treated (and analogously as non-treated). Because the control group only stopped receiving the information treatment in 2012 Q3, treated spells are slightly longer on average. Finally, we observe that the individual forecast errors of non-treated firms are more dispersed than those of treated firms, while the standard deviation of individual forecast revisions is roughly the same for treated and non-treated firms.

3.2 Results

Table 3 reports our main results from estimating equations (3.1) and (3.2) for the treated and non-treated firms. First, we observe that non-treated firms (columns 2 and 4 of Table 3) exhibit underreaction at the consensus level and overreaction at the individual level, mirroring the evidence obtained from professional forecasters' expectations (Bordalo et al., 2020). Firms that receive the treatment show a lower consensus coefficient β_C, while the value of the individual coefficient β_I is less negative. This suggests that information treatments can help firms incorporate new information into their forecasts in a manner that is more consistent with the FIRE benchmark. Another noteworthy feature of Table 3 is the sharp decline in R-squared values when Equations (3.1) and (3.2) are estimated for treated firms.

[9] We define forecast dispersion as the average cross-sectional standard deviation of 6-month ahead inflation forecasts for treated and non-treated firms.

[10] It should be noted that both types of firms exhibit much higher rigidity than professional forecasters. For example, in Bordalo et al. (2020), the non-revision share of professional forecasters' inflation forecasts is close to 7%.

This implies that forecast errors are less predictable from forecast revisions among treated firms, consistent with the idea that the information treatment improves forecast efficiency and brings expectations closer to the rational benchmark.

Estimating equation (3.2) with firm fixed effects ensures that β_I is identified from within-forecaster variation in revisions, effectively controlling for time-invariant heterogeneity in forecasting behavior. However, our panel is relatively short, raising concerns about Nickell bias in fixed-effects estimators. This is particularly relevant for the non-treated group, whose time series begins only in 2012 Q3. To address this concern, Table 4 presents estimates of Equation (3.2) without firm fixed effects – but including fixed effects related to firms' size class, geographical area and sector – as a robustness check. Results confirm the negative and statistically significant coefficient for non-treated firms, and a less negative (here insignificant) coefficient for treated firms.

Overall, the information treatment moves both the consensus- and individual-level error-on-revision coefficients toward the FIRE benchmark. While the consensus evidence is consistent with standard models of information frictions (as formalized in Section 5), the treatment-induced attenuation of individual overreaction provides key evidence for distinguishing among competing models of deviations from individual rationality. Accordingly, Section 4 presents robustness checks targeted at alternative explanations for the individual-level evidence, and Section 5 embeds both the consensus and individual moments in a model-selection exercise.

4 Robustness Checks

Our setting and empirical strategy, which rely on information treatments, raise some unique concerns that challenge the interpretation of our results. First, if reported beliefs are noisy, estimates of β_I from equation (3.2) may reflect spurious negative values. If the treatment also reduces measurement error in reported forecasts, this would explain why β_I is less negative in our treated sample. Second, persistent treatment assignment may facilitate learning about features of the inflation process (e.g., persistence and the long-run mean), resulting in smaller anomalies in the treated sample. Finally, if treated firms anchor their responses to the provided inflation release, measured forecasts may be contaminated by experimenter-demand effects, rather than revealing true inflation expectations. We address these concerns and find that our results remain robust. Additional robustness checks are presented in Appendix B.

4.1 Measurement-Error Bias

As pointed out by Bordalo et al. (2020), if the firms' reported expectations contain measurement error, a negative β_I in (3.2) may reflect spurious correlation as opposed to over-reaction to news. If information treatments also reduce measurement error in reported inflation forecasts, this could lead to a less negative estimate of β_I for the treated firms. To address this concern, we estimate a regression of the 6-month ahead inflation forecast error on the 1-year forecast revision – constructed as the 1-year ahead inflation forecast in quarter t minus the 2-year ahead inflation forecast in quarter $t-4$. If firms update inflation expectations consistently across horizons, the estimates from this regression would still produce a negative coefficient while eliminating the mechanical bias arising from overlap between the dependent and independent variables (Bordalo et al., 2020).

Table 5 presents the results of this regression with and without firm-fixed effects. For non-treated firms, we continue to estimate a negative, statistically significant coefficient, whereas for the treated firms, we estimate a less negative (albeit statistically insignificant) coefficient. Appendix Table B.1 reports the results of regressing the one-year-ahead inflation forecast error on the six-month forecast revision - defined as the difference between the six-month-ahead inflation forecast in quarter t and the one-year-ahead forecast in quarter $t-2$ – and shows that the results are consistent with those in Table 5. This suggests that our main results in Tables 3 and 4 are not the result of information treatments systematically alleviating measurement error bias among treated firms.

4.2 Learning About the DGP

Farmer et al. (2024) show that the forecast anomalies documented in the literature can arise even under Bayesian updating, when agents do not know the underlying data-generating process and learn about parameters that characterize the long-run properties of the DGP. They also demonstrate that learning converges more quickly when agents face fewer unknowns.

In our context, treated firms consistently receive information about recent inflation over the sample period. As a result, treated firms may gradually develop a better understanding of structural features of the DGP - such as inflation persistence, or the long-run mean - which could lead to a reduction in forecast anomalies over time.

To test this hypothesis, we leverage another unique feature of the SIGE treatment design. As described in Section 3, prior to 2012 Q3, all surveyed firms received information about the most recent inflation release in Italy. Beginning in 2012 Q3, firms assigned to the control group were drawn from this previously informed pool. We therefore focus on the early years of the treatment rollout - when non-treated firms had only recently lost access to the information treatment after several quarters of receiving it. Specifically, we restrict the sample to the period from 2012 Q4 to 2015 Q4 and re-estimate equation (3.2) for both treated and non-treated firms. Table 6 presents the results from this exercise alongside our baseline estimates from Table 3. We find that, in this restricted window, non-treated firms exhibit the same level of individual overreaction, suggesting that our main results are not driven by treated firms learning structural features of the DGP (e.g., persistence or the long-run mean) more accurately due to the information treatment.

4.3 Experimenter Demand Effects

Treated firms may anchor their inflation forecasts to the most recent realization provided as part of the treatment, rather than reporting their underlying beliefs. This concern is especially relevant for 6-month-ahead forecasts, where firms may reasonably interpret recent inflation as a highly informative signal about near-term inflation dynamics. To gauge the severity of this issue, we first report statistics on the number of firms in the treated and non-treated samples that report the most recent inflation realization as their forecast at different horizons. At the 6-month horizon, approximately 15% of treated firm forecasts exactly match the provided recent inflation figure, compared to only 6% for non-treated firms. At the 12-month horizon, these proportions fall to 9% for treated firms and 4% for non-treated firms, which shows that mechanically anchoring forecasts to the provided inflation release is higher at shorter horizons. Even fewer firms (6% treated, 1.7% non-treated) align forecasts exactly at both horizons.

To address this issue, we pursue two approaches. First, we estimate our baseline regression at a longer forecast horizon. Specifically, we regress the 1-year-ahead forecast error on the forecast revision defined as the difference between the current 1-year-ahead inflation forecast and the 2-year-ahead forecast made four quarters earlier. Since inflation is less persistent at longer horizons, recent realizations are less directly informative, thereby reducing the likelihood that treated firms mechanically anchor their forecasts to the provided inflation figure (as shown above). Table 7 presents the results from these long-horizon individual-level regressions, alongside the baseline results from Table 3. The similarity of the coefficients across both horizons for treated and non-treated firms suggests that our main findings are not driven by experimenter demand effects.

Second, in our baseline regressions, we exclude firms in the treated sample that report both their 6-month and 1-year-ahead inflation forecasts to be exactly equal to the inflation figure provided to them as part of the treatment as this subset of firms is more likely to be mechanically anchoring their forecasts. Column (3) of Table B.3 presents the results from this specification. We find that the estimated β_I coefficient remains very similar to our baseline estimate for the treated sample, providing additional evidence that experimenter demand effects are not driving our main results.

5 Theoretical framework

In this section, we assess how well existing models of expectation formation can account for the new empirical facts documented in Section 3. The inflation process, x_t, follows an AR(1). Across all model variants, firms observe idiosyncratic private signals about x_t. We model the survey's information intervention as giving treated firms access to an additional public signal about recent inflation at the time forecasts are formed. The models differ in how they generate excess sensitivity of individual forecasts. Specifically, we evaluate the following leading mechanisms that feature departures from individual rationality - internal cognitive constraints (Sung, 2025), diagnostic expectations (Bordalo et al., 2020), overconfidence in private information (Daniel et al., 1998), and over-extrapolation (Angeletos et al., 2021). For each mechanism, we provide closed form expressions for the consensus and individual-level error-on-revision regression coefficients and evaluate model fit using a J-test of overidentifying restrictions. Detailed derivations are present in Appendix C.

Throughout this section, we maintain the following assumptions about the underlying process x_t and the signal structure. A continuum of forecasters $i \in [0,1]$ forecast x_{t+h} for $h \in \{0, 1, ..H\}$. Its current value x_t follows an AR(1) process:

$$x_t = \rho x_{t-1} + u_t \ ; \quad u_t \sim N(0, \sigma_u^2) \tag{5.1}$$

where $u_t \sim N(0, \sigma_u^2)$. We assume that the agents are aware of the long-run mean, which we normalize to 0, as well as the parameters ρ and σ_u^2. The current value of x_t is not perfectly observed, and each forecaster i instead observes a noisy signal:

$$s_t^i = x_t + \epsilon_{it}; \quad \epsilon_{it} \sim N(0, \sigma_\epsilon^2) \tag{5.2}$$

Treated firms, in addition, also have access to an informative public signal:

$$s_t = x_t + \nu_t; \quad \nu_t \sim N(0, \sigma_\nu^2) \tag{5.3}$$

For non-treated firms, we assume that the public signal is shut down; i.e., $\sigma_\nu^2 \to \infty$. At each time t, forecasters minimize the expected quadratic loss from inaccurate projections across H horizons:

$$\sum_{h=0}^{H} \left(x_{t+h} - F_{i,t}(x_{t+h})\right)^2 \tag{5.4}$$

This implies that the forecast of x_{t+h} for agent i is $\rho^h F_{i,t}(x_t)$, where $F_{i,t}(x_t)$ is agent i's best guess of the current value of the hidden state variable, which we denote as $x_{t|t}^i$. Then, $F_{i,t}\ (x_{t+h}) = x_{t+h|t}^i = \rho^h x_{t|t}^i$. Also, we denote $x_{t+h|t} = \int_0^1 x_{t+h|t}^i$ We now derive the regression coefficients that are the theoretical counterparts of the empirical specifications in equations (3.1) and (3.2).

5.1 Internal Cognitive Constraints

This section adapts the framework in Section 2 of Sung (2025), but instead of modeling endogenous information acquisition, we assume that agents receive exogenous noisy signals about the underlying state, as specified in equations (5.2) and (5.3). While agents are assumed to know the process parameters, they face cognitive costs in retrieving information from previous periods, which increases their sensitivity to new information and generates overreaction at the individual level.

Let m_{it} denote the mental representation of existing knowledge at the start of period t, while $M_{it} = (m_{it-1}, s_{it-1}, s_{t-1})$ denotes the prior knowledge that is available to the agent before she observes the noisy signals at time t. Following Sung (2025), we assume that the mental representation of existing knowledge takes the following form:

$$m_{it} = \Lambda_t \; M_{it} + \tilde{\omega_{it}}; \quad \tilde{\omega_{it}} \sim N(0, \Sigma_{\omega t}) \tag{5.5}$$

The agent faces a constraint on the mutual information between m_{it} and M_{it}:

$$\mathcal{I}(m_{it}, M_{it}) \leq -\frac{\log \phi_m}{2} \tag{5.6}$$

In this case, Sung (2025) shows that the optimal memory representation takes the following form:

$$m_{it} = \lambda_t \, \mathbb{E}[x_t | M_{it}] + \omega_{it}; \quad \omega_{it} \sim N(0, \sigma_{\omega_t}) \tag{5.7}$$

where $\lambda_t \in [0,1]$ and $\mathbb{E}[x_t|M_{it}] = x^i_{t|t-1}$. In Appendix C.1, we show that under these assumptions, $\lambda_t = \lambda = 1 - \phi_m$. In general, there are multiple values of $\sigma^2_{\omega t}$ and λ that satisfy the constraint (5.6). Therefore, we impose the following normalization that allows us to determine $\sigma^2_{\omega t}$ purely as a function of λ:

$$\text{Cov}[\mathbb{E}[x_t|M_{it} - m_{it} \; , \; m_{it}]] = 0 \tag{5.8}$$

This allows us to derive the following expressions for the asymptotic limit of β_C and β_I (a detailed exposition is in Appendix C.1):

Proposition 1. *When forecasters form expectations subject to the Noisy Memory constraint and receive both idiosyncratic private signals and a public signal, the asymptotic error-on-revision coefficients are*

$$\beta_I \; = \; -\frac{(1-K)(1-\lambda)}{\rho^{-2} + 1 - 2\big[(1-K)\lambda + K\big]} \tag{5.9}$$

and

$$\beta_C \;=\; \frac{1-K}{K} \;-\; \frac{K_2^2\,\sigma_\nu^2}{K\,V_{agg}} \;+\; \frac{(1-K)(1-\lambda)}{K}\,\frac{C_{x_{t|t-1},agg}}{V_{agg}}\,. \tag{5.10}$$

where K_1 and K_2 denote the steady state Kalman gains on the private and public signals, respectively, with $K = K_1 + K_2$, V_{agg} denotes the variance of the aggregate forecast revision and $C_{x_{t|t-1},agg}$ denotes the covariance between $x_{t|t-1}$ and the aggregate forecast revision □

As long as $\lambda < 1$, it follows that $\beta_I < 0$. Intuitively, limits on recalling past information lead forecasters to place more weight on newly observed signals, which in turn generates individual-level overreaction. In the absence of a public signal, we also have $\beta_C > 0$. However, as shown by Coibion and Gorodnichenko (2015), introducing a common signal can bias β_C downward (given by the third term in 5.10) and may even result in negative values of β_C. Details on how we calculate $C_{x_{t|t-1},agg}$ and V_{agg} are in Appendix C.1.

5.2 Diagnostic Expectations

Bordalo et al. (2020) develop a model of expectation formation in which agents receive idiosyncratic noisy signals and exhibit belief distortions driven by Kahneman and Tversky's representativeness heuristic that accounts for both consensus underreaction and individual-level overreaction. The representativeness of a state x at time t is defined by the following likelihood ratio:

$$R_t(x) = \frac{f(x \mid S_t^i)}{f(x \mid S_{t-1}^i \cup \{x_{t|t-1}^i\})}, \tag{5.11}$$

where S_t^i represents the complete set of signals observed by agent i up to time t (for treated firms, S_t^i includes both public and private signals), and $f(x \mid S_t^i)$ denotes the rational posterior belief about the state x after observing these signals. The forecaster then overweights representative states by using the following distorted posterior:

$$f^\theta\left(x_t|S_t^i\right) = f\left(x_t|S_t^i\right) R_t(x_t)^\theta \frac{1}{Z_t} \tag{5.12}$$

where Z_t is a normalizing factor that ensures that $f^\theta(x_t|S_t^i)$ integrates to 1. θ governs the extent to which forecasters overweight the absolute probability of states that have become relatively more likely after observing today's signals. Under these assumptions, we have the following expressions for the asymptotic limit of β_C and β_I:

Proposition 2. *Under Diagnostic Expectations, when forecasters receive both idiosyncratic private signals and a public signal, the asymptotic error-on-revision coefficients are*

$$\beta_I \;=\; -\frac{\theta(1+\theta)}{(1+\theta)^2+\theta^2\rho^2} \tag{5.13}$$

and

$$\beta_C = \frac{\left(\frac{1-K}{K}-\theta\right)\left([1+\theta-\theta\rho^2(1-K)]V^R_{agg}+\theta\rho^2K_2^2\sigma_\nu^2\right)-\frac{(1+\theta)K_2^2\sigma_\nu^2}{K}}{V_{agg}} \tag{5.14}$$

where K_1 and K_2 denote the steady state Kalman gains on the private and public signals, respectively, with $K = K_1 + K_2$, V_{agg} is the variance of the aggregate forecast revision under diagnostic expectations and V^R_{agg} is the variance of the forecast revision under rational expectations □

A detailed exposition of this proposition is provided in Section A of the Online Appendix of Bordalo et al. (2020) and is therefore omitted here. As shown in equation (5.13), the individual-level error-on-revision regression coefficient in the diagnostic expectations model depends only on the diagnostic distortion θ and the process persistence ρ. Thus, while the provision of a public signal affects β_C, it leaves β_I unchanged. Consequently, the diagnostic expectations framework of Bordalo et al. (2020) cannot account for the post-treatment increase in β_I documented in Section 3 [11]. Intuitively, the diagnostic distortion θ scales the Kalman gains on *all* signals proportionally, which implies that forecasters overreact to both public and private signals.

5.3 Overconfidence

We now consider a model in which forecasters are overconfident about the precision of their private signals (Broer & Kohlhas, 2024; Daniel et al., 1998). While the true variance of the idiosyncratic signal is σ^2_ϵ, agents believe it to be $\hat{\sigma}^2_\epsilon < \sigma^2_\epsilon$; that is, they perceive their private

[11] A natural extension is the Smooth Diagnostic Expectations model of Bianchi et al. (2024), in which the severity of the diagnostic distortion varies with conditional uncertainty. If the information treatment lowers conditional uncertainty, smooth DE would predict the rise in β_I that we observe empirically. We leave a full analysis to future work

signals as more informative than they actually are. We assume that forecasters correctly assess the informativeness of the public signal, as defined in equation (5.3). Under these assumptions, we have the following expressions for the coefficients β_C and β_I.

Proposition 3. *When forecasters observe idiosyncratic private signals as well as a public signal and are overconfident about the precision of their private signal, the asymptotic error-on-revision coefficients are*

$$\beta_I = \frac{1-K}{K} - \frac{K_1^2 \sigma_\epsilon^2}{K\, V_{ind}} - \frac{K_2^2 \sigma_\nu^2}{K\, V_{ind}}. \tag{5.15}$$

and

$$\beta_C = \frac{1-K}{K} - \frac{K_2^2 \sigma_\nu^2}{K\, V_{agg}}. \tag{5.16}$$

where K_1 and K_2 denote the steady state (distorted) Kalman gains on the private and public signals, respectively, with $K = K_1 + K_2$, V_{ind} denotes the variance of individual forecast revisions and V_{agg} denotes the variance of the aggregate forecast revisions. □

Because agents overestimate the precision of their private signals, they update more aggressively than a rational agent would in response to any given realization of that signal. This excessive updating generates systematic overreaction, and yields a negative value of β_I. Expressions for the Kalman gains and the variance of the forecast revisions are provided in Appendix C.2.

5.4 Over-extrapolation

In this model, we assume that while the true data-generating process (DGP) continues to follow equation (5.1), agents believe that the process is more persistent than it actually is, i.e., they believe the underlying state evolves according to:

$$x_t = \hat{\rho} x_{t-1} + u_t; \quad u_t \sim N(0, \sigma_u^2) \tag{5.17}$$

The asymptotic error-on-revision regression coefficients are then given by the following expressions:

Proposition 4. *When forecasters have extrapolative expectations and observe idiosyncratic private signals as well as a public signal, the asymptotic error-on-revision coefficients are*

$$\beta_I^h \;=\; \frac{1-K}{K} \;-\; \frac{K_1^2\sigma_\epsilon^2}{K\,V_{ind}} \;-\; \frac{K_2^2\sigma_\nu^2}{K\,V_{ind}} \;+\; \frac{\left(\rho^h-\hat{\rho}^h\right)}{\hat{\rho}^h}\frac{C_{x,Ind}}{V_{ind}}. \tag{5.18}$$

and

$$\beta_C^h \;=\; \frac{1-K}{K} \;-\; \frac{K_2^2\,\sigma_\nu^2}{K\,V_{agg}} \;+\; \frac{\left(\rho^h-\hat{\rho}^h\right)}{\hat{\rho}^h}\frac{C_{x,agg}}{V_{agg}}. \tag{5.19}$$

where K_1 and K_2 denote the steady state (distorted) Kalman gains on the private and public signals, respectively, with $K = K_1 + K_2$. Here, V_{ind} and V_{agg} denote the variances of the individual and aggregate forecast revisions, respectively, and $C_{x,Ind}$ and $C_{x,agg}$ denote the covariances of x_t with the individual and aggregate forecast revisions, respectively. □

Since agents overestimate the persistence of the underlying process, they expect shocks to x_t to have more prolonged effects than they truly do, leading to negative ex-post forecast errors and a negative β_I. Among the models considered in this section, the over-extrapolation model is unique in that the coefficients β_C^h and β_I^h depend on the forecast horizon h.[12] This is evident from equations (5.18) and (5.19). The reason is that in this model, agents make forecast errors because $\hat{\rho} \neq \rho$, and the magnitude of these errors varies with the forecast horizon h. Expressions for the Kalman gains, variances of the individual and aggregate forecast revisions and covariance of x_t with the individual and aggregate forecast revisions are given in Appendix C.3.

5.5 Analysis

We now assess the ability of the three models – internal cognitive constraints, overconfidence in private information, and over-extrapolation – to quantitatively match the empirical findings documented in Section 3.[13] To this end, we implement a model selection exercise by testing overidentifying restrictions using a J-test. We begin by estimating the parameters of the data-generating process, ρ and σ_u^2, using historical data on consumer price inflation

[12] An exception is the Noisy Memory model in which agents use past data to also learn about the long-run mean of the process; in that case β_C and β_I also vary with the horizon.

[13] Recall that we excluded the diagnostic expectations model since the expression for β_I in that model is independent of σ_ν^2

in Italy. We then specify the free parameters associated with each model and describe the moments used in our test of overidentifying restrictions. Finally, we present the results of the test, showing that the overconfidence model provides the best fit, and conclude with an intuition for why this model outperforms the others.

First, we fit an AR(1) process to the de-meaned quarterly series of year-on-year changes in the Harmonized Index of Consumer Prices (HICP) for Italy[14]. This allows us to estimate the persistence of the process, denoted ρ_q, and the variance of its innovations, denoted $\sigma^2_{u,q}$. To ensure consistency with the two-quarter revision window used in our empirical analysis, we transform these estimates using the relationships $\rho = \rho_q^2$ and $\sigma_u^2 = (1 + \rho_q^2)\sigma^2_{u,q}$. These parameters are treated as fixed throughout the analysis, although we examine the robustness of our results to several alternative specifications of the data-generating process in Appendix D.

Each model features three free parameters: the variance of the private signal, σ^2_ϵ; a model-specific behavioral parameter $\hat{\sigma}^2_\epsilon$ in the overconfidence model, λ in the internal cognitive constraints model, and $\hat{\rho}$ in the over-extrapolation model; and the variance of the public signal for treated firms, σ^2_ν. Model-implied non-treated moments are obtained by evaluating the same closed-form expressions at $\sigma^2_\nu \to \infty$ (i.e., zero public signal precision). Our test of overidentifying restrictions targets five empirical moments. The first four are the consensus and individual error-on-revision regression coefficients for non-treated firms, $\beta_C\,(NT)$ and $\beta_I\,(NT)$, and for treated firms, $\beta_C\,(T)$ and $\beta_I\,(T)$. The fifth moment is the ratio of the average cross-sectional standard deviation of six-month-ahead inflation forecasts for treated firms relative to non-treated firms, which disciplines the value of σ^2_ν. Intuitively, the availability of a common public signal causes all firms to update based on the same signal realization, thereby reducing forecast dispersion in theory and in the data (see Table 2). Analytical expressions for the cross-sectional variance of forecasts under different model specifications are given in Appendix C.

5.6 Overidentification test

We estimate the parameters of the three models using an overidentified Generalized Method of Moments (GMM) procedure and assess model fit using a J-test. Let $\hat{m}$ denote the empirical

[14]This is the inflation rate that firms in our survey are asked to forecast.

moments and $m(\theta)$ denote the model-implied moments for a given parameter vector θ. For each model, we choose θ to solve the following minimization problem:

$$\hat{\theta} = \arg\min_{\theta} \left(m(\theta) - \hat{m}\right)' W \left(m(\theta) - \hat{m}\right), \tag{5.20}$$

where W, the optimal weighting matrix, is the inverse of the variance-covariance matrix of the five moments that we consider. To estimate the optimal weighting matrix, we implement a fixed-length time-block bootstrap. In each bootstrap draw, we sample, with replacement, blocks of twenty consecutive quarters to assemble a bootstrap time path that is the same length as our original panel, and we retain all firms observed in the selected quarters. We recompute the five moments on this resampled panel and use the sample variance-covariance matrix across bootstrap draws as our estimate of the optimal variance matrix. Because we have five moments and three parameters, the minimized value of the GMM objective function in equation (5.20) is distributed as a chi-squared random variable with two degrees of freedom. The parameter estimates from this procedure are given in Table 8, and the estimated model fit and J-statistic from our GMM exercise are reported in Table 9.

First, note that the forecast dispersion ratio is the most precisely estimated moment in the bootstrap, so it receives the greatest weight in the optimal matrix and effectively disciplines σ_ν^2 across specifications. As a result, all three models match this moment fairly well. Among the alternatives, the Overconfidence (OC) model provides the best fit. It reproduces the non-treated coefficients β_C and β_I quite well, and for the implied value of σ_ν^2 it generates changes in the treated coefficients that are of comparable magnitude to those observed in the data. The resulting J-statistic is small, with an associated p-value of 0.91, indicating that the overidentifying restrictions are comfortably satisfied. Taken together, these results point to a strong overall fit of the OC model relative to the other specifications.

In contrast, the Noisy Memory model reproduces the non-treated coefficients reasonably well, but—given σ_ν^2 that is disciplined by the dispersion ratio—it implies a much larger drop in β_C for treated firms (0.03 vs. 0.43 in the data) and only a modest increase in β_I (-0.25 vs. -0.15 in the data). The resulting J-statistic is higher and the p-value is significantly lower ($p = 0.15$): we do not reject at conventional levels, but the fit is notably weaker than the OC model, with this model failing to match both treated coefficients for reasonable values of σ_ν^2.

Finally, the over-extrapolation model delivers the weakest fit out of the alternatives we consider. Since coefficients in this framework are horizon-dependent, we set $h = 1$ (corresponding to a six-month horizon, consistent with our empirical analysis in Section 3). To generate $\beta_I(NT) = -0.28$ at this short horizon, the model requires $\hat{\rho}$ to be substantially larger than the true persistence parameter ρ. However, with the restriction $\hat{\rho} < 1$, the model implied $\beta_I(NT)$ is much closer to zero (-0.15 in our calibration) relative to what we observe in the data. It also generates a counterfactually large value of β_C for the non-treated firms. As a result, the J-statistic is large and the associated p-value falls below 1% ($p < 0.01$), leading to a strong rejection of the model's overidentifying restrictions.

To build intuition, consider the following exercise: choose σ_ϵ^2 and the relevant model-specific behavioral parameter so that each model exactly matches β_C and β_I for non-treated firms. Then, increase σ_ν^2 until the model reproduces the treated β_C, leaving β_I for the treated firms as a non-targeted moment. Results of this exercise are shown in Figure 4 and Panel A of Appendix Table D.8.

The first row of Figure 4 reports the results of the exercise for the overconfidence model.[15] Note that the value of β_C reflects how aggressively agents update their forecasts in response to new information: when updating is stronger, β_C is lower. In the overconfidence model, agents underreact to the public signal relative to rational agents. Consequently, in order to match the empirically observed drop in β_C, the public signal has to be made very precise. At the same time, in all models, increasing the precision of the public signal reduces the Kalman gain on the private signal (see Appendix Figure D.2). In the overconfidence model, individual-level overreaction arises precisely because forecasters overweight their private signal. Therefore, a precise public signal leads to a significantly smaller weight on the private signal, which in turn reduces individual-level overreaction.

The second row of Figure 4 shows the results of this exercise for the internal cognitive constraints model. Because agents retain only a fraction λ of past information, their priors are weak and they place relatively high weight on current signals. A small increase in the precision of the public signal therefore raises the public signal Kalman gain substantially, producing a *large* decline in the consensus coefficient β_C. Turning to β_I, better information

[15] Figure 4 displays the three models in separate rows (Overconfidence, Noisy Memory, Over-extrapolation), with β_C on the left and β_I on the right as functions of public signal precision τ_ν. Appendix Figure D.1 overlays the models: the left panel plots β_C against τ_ν for all three models, and the right panel does the same for β_I.

today reduces reliance on distorted priors and nudges β_I upward (toward zero). However, the increase in precision required to match the decline in β_C is modest, so the impact on β_I is limited. Intuitively, greater signal precision today does not repair memory: only a fraction λ of today's information is carried into the next period. Entering the next period with a weakened prior, agents place too much weight on the new signal and therefore overreact. These mechanisms account for the patterns in Table 9: to match the decline in forecast dispersion that we observe in the data, the public signal must be relatively precise. Under the Noisy Memory model this precision yields an outsized drop in β_C but only a modest increase in β_I, so the model's ability to fit the treated coefficients is weak.

What would it take for the model with Noisy Memory to match both β_I moments? We run a second just-identified exercise in which we keep σ_ϵ^2 and the model-specific behavioral parameter calibrated to match the non-treated coefficients, and then choose σ_ν^2 to match the treated β_I, leaving the treated β_C as a non-targeted moment. Panel B of Appendix Table D.8 reports the outcome of this exercise.

In the Noisy Memory model, β_I is only weakly affected by the precision of the public signal, so to raise β_I from its non-treated level to the treated value requires an extremely precise public signal. This has two consequences: First, because treated firms load heavily on the same precise public signal, the cross-sectional dispersion of forecasts collapses too much relative to what we observe in the data. Second, the common-noise term in equation (5.10) that depresses the consensus coefficient becomes large, and this leads to a negative β_C for the treated firms, contrary to what we observe in the data.

Finally, the Over-extrapolation model fails on two dimensions. First, it cannot match the coefficient β_I for non-treated firms (as discussed above). Second, β_I barely changes after treatment. The reason is that, in this model, the covariance between forecast errors and revisions is driven by agents' misperceived persistence of x_t (with $\hat{\rho} > \rho$); increasing the precision of the public signal σ_ν^2 does not correct this belief. As a result, the effect of reducing σ_ν^2 on β_I is small.

Taken together, our model selection exercise identifies *Overconfidence in private information* as providing the best overall fit: it jointly matches the empirical (β_C, β_I) for both treated and non-treated firms at realistic values of σ_ν^2 disciplined by the forecast-dispersion ratio of treated firms relative to non-treated firms. Therefore, the Overconfidence model attains a very small $J = 0.19$ with a high $p = 0.91$, so we do not reject the overidentifying restrictions. By contrast, the Noisy Memory model can reproduce the nontreated (β_C, β_I), but for values of σ_ν^2 disciplined by the forecast dispersion ratio, it implies a counterfactually large decline in β_C and only a small increase in β_I. The over-extrapolation model performs worst: it cannot deliver the non-treated β_I at $h = 1$ under admissible $\hat{\rho} < 1$ and is decisively rejected ($p < 0.01$).

6 Conclusion

Using a unique randomized information experiment embedded in the Survey of Inflation and Growth Expectations (SIGE), we study the extent to which targeted information provision can mitigate systematic deviations from rationality in firms' inflation forecasts. We find that while non-treated firms exhibit underreaction at the consensus level and overreaction at the individual level - in line with findings from studies using professional forecaster data - information treatments can significantly reduce rigidity at the consensus level (Coibion & Gorodnichenko, 2012, 2015) as well as overreaction at the individual level (Afrouzi et al., 2023; Angeletos et al., 2021; Bordalo et al., 2020; Kohlhas & Walther, 2021). These results also connect to the broader evidence on the inertial response of macroeconomic variables to shocks (Christiano et al., 2005) and to the subsequent literature that has tried to rationalize this through information frictions (Mackowiak & Wiederholt, 2009; Woodford, 2001). In particular, the coefficient β_C can be used to discipline the extent of information rigidity in macroeconomic models, as shown by Angeletos and Huo (2021). Our evidence on the reduction in β_C in response to the information treatment can therefore be used as a new calibration target to study how targeted information provision in the form of central bank communication shapes the propagation and persistence of shocks to the economy, which we view as a promising direction for future work.

We address standard methodological concerns as well as some considerations that are unique to our setting and find that our main results are quite robust. Turning to theory, we find that a simple model featuring overconfidence in private information (Broer & Kohlhas, 2024; Daniel et al., 1998) and an informative public signal provided to treated firms provides the best explanation for our observed empirical results. Our model selection exercise also reveals that many alternative models of expectation-formation struggle to rationalize the increase in β_I that we observe in our sample of treated firms. Finally, while our analysis is confined to Italy over a specific period and to inflation as the outcome, the growing availability of firm surveys with a long time-series coverage and similar information treatments in other countries will enable tests of external validity – assessing whether our results generalize across contexts and to other outcome variables.

References

Adam, K., Kuang, P., & Xie, S. (2025). Overconfidence in private information explains biases in professional forecasts. *Journal of Monetary Economics*, *155*, 103839.

Afrouzi, H., Kwon, S. Y., Landier, A., Ma, Y., & Thesmar, D. (2023). Overreaction in Expectations: Evidence and Theory. *The Quarterly Journal of Economics*, *138*(3), 1713–1764.

Andrade, P., & Le Bihan, H. (2013). Inattentive professional forecasters. *Journal of Monetary Economics*, *60*(8), 967–982.

Angeletos, G.-M., & Huo, Z. (2021). Myopia and anchoring. *American Economic Review*, *111*(4), 1166–1200.

Angeletos, G.-M., Huo, Z., & Sastry, K. A. (2021). Imperfect macroeconomic expectations: Evidence and theory. *NBER Macroeconomics Annual*, *35*, 1–86.

Bartiloro, L., Bottone, M., & Rosolia, A. (2019). The heterogeneity of the inflation expectations of italian firms along the business cycle. *International Journal of Central Banking*, *15*(5), 175–205.

Bianchi, F., Ilut, C. L., & Saijo, H. (2024, February). *Smooth diagnostic expectations* (Working Paper No. 32152). National Bureau of Economic Research.

Bordalo, P., Gennaioli, N., Ma, Y., & Shleifer, A. (2020). Overreaction in macroeconomic expectations. *American Economic Review*, *110*(9), 2748–82.

Bordalo, P., Gennaioli, N., Porta, R. L., & Shleifer, A. (2024). Belief overreaction and stock market puzzles. *Journal of Political Economy*, *132*(5), 1450–1484.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2018). Diagnostic expectations and credit cycles. *The Journal of Finance*, *73*(1), 199–227.

Born, B., Enders, Z., Menkhoff, M., Müller, G., & Niemann, K. (2022, December). *Firm expectations and news: Micro v macro* (CEPR Discussion Papers No. 17768). C.E.P.R. Discussion Papers.

Bottone, M., Tagliabracci, A., & Zevi, G. (2022). Inflation expectations and the ecb's perceived inflation objective: Novel evidence from firm-level data. *Journal of Monetary Economics*, *129*, S15–S34.

Bouchaud, J.-P., Krüger, P., Landier, A., & Thesmar, D. (2019). Sticky expectations and the profitability anomaly. *The Journal of Finance*, *74*(2), 639–674.

Broer, T., & Kohlhas, A. N. (2024). Forecaster (mis-)behavior. *The Review of Economics and Statistics*, *106*(5), 1334–1351.

Christiano, L. J., Eichenbaum, M., & Evans, C. L. (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy*, *113*(1), 1–45.

Coibion, O., Georgarakos, D., Gorodnichenko, Y., & van Rooij, M. (2023). How does consumption respond to news about inflation? field evidence from a randomized control trial. *American Economic Journal: Macroeconomics*, *15*(3), 109–52.

Coibion, O., & Gorodnichenko, Y. (2012). What can survey forecasts tell us about information rigidities? *Journal of Political Economy*, *120*(1), 116–159.

Coibion, O., & Gorodnichenko, Y. (2015). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review*, *105*(8), 2644–78.

Coibion, O., Gorodnichenko, Y., & Ropele, T. (2019). Inflation Expectations and Firm Decisions: New Causal Evidence. *The Quarterly Journal of Economics*, *135*(1), 165–219.

Coibion, O., Gorodnichenko, Y., & Weber, M. (2022). Monetary policy communications and their effects on household inflation expectations. *Journal of Political Economy*, *130*(6), 1537–1584.

Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor psychology and security market under- and overreactions. *The Journal of Finance*, *53*(6), 1839–1885.

Farmer, L. E., Nakamura, E., & Steinsson, J. (2024). Learning about the long run. *Journal of Political Economy*, *132*(10), 3334–3377.

Greenwood, R., & Shleifer, A. (2014). Expectations of returns and expected returns. *The Review of Financial Studies*, *27*(3), 714–746.

Kohlhas, A. N., & Walther, A. (2021). Asymmetric attention. *American Economic Review*, *111*(9), 2879–2925.

Link, S., Peichl, A., Pfäuti, O., Roth, C., & Wohlfart, J. (2023, September). *Attention to the macroeconomy* (ECONtribute Discussion Papers Series No. 256). University of Bonn and University of Cologne, Germany.

Mackowiak, B., & Wiederholt, M. (2009). Optimal sticky prices under rational inattention. *American Economic Review*, *99*(3), 769–803.

Mackowiak, B., & Wiederholt, M. (2024, March). *Rational Inattention during an RCT* (CEPR Discussion Papers No. 18875). C.E.P.R. Discussion Papers.

Mangiante, G., & Tagliabracci, A. (2025). Inflation regimes, information treatments, and firms expectation formation. *mimeo*.

Rosolia, A. (2024). Do Firms Act on Their Inflation Expectations? Another Look at Italian Firms. *Journal of Political Economy Macroeconomics*, *2*(4), 651–686.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, *50*(3), 665–690.

Sung, Y. (2025). Macroeconomic expectations and cognitive noise. *Federal Reserve Bank of San Francisco, Working Paper Series*, *2024*, 01–78.

Weber, M., Candia, B., Afrouzi, H., Ropele, T., Lluberas, R., Frache, S., Meyer, B., Kumar, S., Gorodnichenko, Y., Georgarakos, D., Coibion, O., Kenny, G., & Ponce, J. (2025). Tell me something i don't already know: Learning in low- and high-inflation settings. *Econometrica*, *93*(1), 229–264.

Woodford, M. (2001, December). *Imperfect common knowledge and the effects of monetary policy* (Working Paper No. 8673). National Bureau of Economic Research.

Table 1: Summary Statistics: Treated vs. Non-Treated Firms

Statistic	N. Obs		Mean		SD	
Treatment Status	No	Yes	No	Yes	No	Yes
Variable:						
Inflation Forecast (6-Months)	8,580	21,217	0.99	1.11	0.88	1.07
Inflation Forecast (1-Year)	8,580	21,217	1.14	1.22	0.89	1.05
Inflation Forecast (2-Year)	8,580	21,217	1.34	1.37	0.93	1.02
Log Number of Employees	8,580	21,217	5.34	5.42	1.32	1.41
Beliefs about current macroeconomic conditions	8,462	20,946	1.85	1.83	0.60	0.62
Improvement of macroeconomic conditions in 3 months	8,549	21,141	1.88	1.88	0.90	0.86
Share of turnover from exports	8,580	21,217	2.02	2.01	1.13	1.11
Expected price change next 12 months	8,580	21,217	0.72	0.78	3.70	4.20
Beliefs about own business conditions	8,418	20,859	3.43	3.46	0.79	0.78
Effect of Changes in Demand on Business Conditions	6,853	17,441	0.34	0.38	1.59	1.57
Effect of Credit Availability and Cost of Credit on Business Conditions	8,439	20,775	0.01	-0.02	1.37	1.40

Table 1 reports summary statistics for treated and non-treated firms in the SIGE dataset across a range of variables, such as inflation expectations at different horizons, number of employees (log), share of turnover from exports, changes in prices they expect to make over the next 12 months, optimism about future business conditions and factors affecting their business conditions over the next 3 months.

Table 2: Forecast Errors and Revisions: Treated vs. Non-Treated Firms

Panel A: Consensus Errors and Revisions

Variable	Mean (Treated)	Mean (Non-treated)	SD (Treated)	SD (Non-treated)	SE (Treated)	SE (Non-treated)
Forecast Error (6 Months)	-0.23	-0.49	0.82	0.71	0.21	0.21
Forecast Revision (6 Months - lagged 1 Year)	-0.19	-0.26	0.51	0.24	0.12	0.07

Panel B: Individual-Level Forecast Statistics

Variable	Treated	Non-treated
Forecast Dispersion	0.47	0.67
Non-revision share	0.31	0.35
Avg no. of obs per firm	18.4	13.24
Standard Deviation- Forecast Errors	0.88	0.99
Standard Deviation- Forecast Revisions	0.73	0.72

Note: Panel A of Table 2 reports summary statistics on consensus forecast errors and revisions for the 6-month-ahead inflation forecast. Forecast errors are computed as the difference between realized inflation and the consensus forecast, while forecast revisions are defined as the difference between the current 6-month-ahead forecast and the 1-year-ahead forecast made two quarters prior. Columns (1) and (2) report the time-series average of these statistics for treated and non-treated firms, respectively. Columns (3) and (4) report the time-series standard deviation, and columns (5) and (6) report Newey-West standard errors. Panel B of Table 2 reports summary statistics for individual-level forecasts. Forecast Dispersion is the average within-quarter standard deviation of individual 6-month-ahead forecasts. Non-revision share is the fraction of firms whose current 6-month-ahead forecast is within 0.01% of their 1-year-ahead forecast from two quarters earlier. Avg. no. of obs. per firm indicates the average number of periods a firm appears in the treated and non-treated groups, respectively. The final two rows report the standard deviations of individual forecast errors and forecast revisions over time.

Table 3: Full Sample Estimates of β_C and β_I

Dependent Variables:	Consensus Error (6 Months)		Individual Error (6 Months)	
Sample:	Treated	Not-treated	Treated	Not-treated
β_C	0.59**	1.09***		
	(0.23)	(0.27)		
β_I			-0.18**	-0.40***
			(0.07)	(0.05)
Fixed-effects				
firm			Yes	Yes
Fit statistics				
Observations	41	32	13,861	4,797
R^2	0.1335	0.1795	0.1855	0.4657
Within R^2			0.0263	0.1575

Table 3 presents the results from estimating equations (3.1) and (3.2). Columns (1) and (2) estimate the consensus-level error-on-revision regression (3.1) for treated and non-treated firms, respectively. Standard errors are Newey-West with automatic bandwidth selection. Columns (3) and (4) report the individual-level error-on-revision regressions (Equation (3.2)), controlling for firm fixed effects. Standard errors in columns (3) and (4) are clustered by firm and by time. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4: β_I Estimates: With and Without Fixed Effects

Dependent Variable:	Individual Forecast Error (6 Months)			
Sample:	Treated	Non-treated	Treated	Non-treated
β_I	-0.11	-0.39***	-0.18**	-0.39***
	(0.07)	(0.06)	(0.07)	(0.05)
Fixed-effects				
sector of activity	Yes	Yes		
geographical area	Yes	Yes		
firm size class	Yes	Yes		
firm			Yes	Yes
Fit statistics				
Observations	13,936	4,850	13,861	4,797
R^2	0.0128	0.1145	0.1855	0.4657
Within R^2			0.0263	0.1575

Table 4 reports estimates of equation (3.2), with and without firm fixed effects, for both treated and non-treated firms. Columns (1) and (2) estimate the regression without fixed effects, while columns (3) and (4) include firm fixed effects. All specifications use clustered standard errors at the firm and quarter level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5: Individual-Level Regressions: Robustness to Measurement Error

Dependent Variable:	Individual Forecast Error 6 (Months)			
Sample:	Treated	Non-treated	Treated	Non-treated
Variables				
Forecast Revision (1 Year - 2 Year)	-0.04	-0.30***	-0.09	-0.31***
	(0.06)	(0.06)	(0.05)	(0.05)
Fixed-effects				
sector of activity	Yes	Yes		
geographical area	Yes	Yes		
firm size class	Yes	Yes		
firm			Yes	Yes
Fit statistics				
Observations	12,769	4,068	12,648	3,976
R^2	0.0045	0.0961	0.1641	0.4343
Within R^2			0.0097	0.12

Table 5 addresses concerns about mechanical correlation by estimating Equation (3.2) using an alternative forecast revision measure. Specifically, it reports regressions of the 6-month-ahead forecast error on forecast revisions at a different horizon-defined as the difference between the 1-year-ahead forecast and the 2-year-ahead forecast from four quarters earlier. Columns (1) and (2) exclude firm fixed effects; columns (3) and (4) include them. Standard errors are clustered by firm and quarter. Significance levels: *** $\Rightarrow$ p $<$ 0.01, ** $\Rightarrow$ p $<$ 0.05, * $\Rightarrow$ p $<$ 0.1.

Table 6: β_I Estimates: Robustness to Learning Dynamics

Dependent Variable:	Individual Forecast Error (6 Months)			
Sample:	Treated	Non-treated	Treated	Non-treated
Variables				
β_I	-0.18*	-0.40***	-0.25*	-0.40***
	(0.07)	(0.05)	(0.06)	(0.02)
Sample Years	Full Sample	Full Sample	2012 Q4-2015 Q4	2012 Q4-2015 Q4
Fixed-effects				
firm	Yes	Yes	Yes	Yes
Fit statistics				
Observations	13,861	4,797	4,392	1,942
R^2	0.1855	0.4657	0.3239	0.6906
Within R^2	0.0263	0.1575	0.1012	0.2975

Table 6 reports estimates of Equation (3.2) for both the full sample and an early treatment window, separately for treated and non-treated firms. All specifications include firm fixed effects, and standard errors are clustered by firm and quarter. Significance levels: *** $\Rightarrow$ p $<$ 0.01, ** $\Rightarrow$ p $<$ 0.05, * $\Rightarrow$ p $<$ 0.1.

Table 7: β_I Estimates: Long Horizon

Dependent Variables:	Forecast Error (6M)		Forecast Error (12M)	
Sample:	Treated	Non-treated	Treated	Non-treated
β_I	-0.18**	-0.28***		
	(0.07)	(0.05)		
β_I (12 Months)			-0.21**	-0.36***
			(0.09)	(0.05)
Fixed-effects				
firm	Yes	Yes	Yes	Yes
Fit statistics				
Observations	13,861	4,797	12,198	3,857
R^2	0.1855	0.4657	0.1541	0.3379
Within R^2	0.0263	0.1575	0.0287	0.0899

Table 7 presents estimates of Equation (3.2) using both 6-month-ahead (columns (1) and (2)) and 12-month-ahead inflation forecast errors (columns (3) and (4)) as dependent variables, for treated and non-treated firms. In columns (3) and (4), the independent variable is the 1-year ahead inflation forecast in quarter t *minus* the 2-year ahead inflation forecast made in quarter $t-4$. All regressions include firm fixed effects, and standard errors are clustered by firm and quarter. This table tests whether the effects of information treatments persist at longer forecast horizons, where recent inflation realizations are less likely to anchor expectations. Significance levels: *** ⇒ $p < 0.01$, ** ⇒ $p < 0.05$, * ⇒ $p < 0.1$.

Table 8: Calibration Details: Over-Identification Test

	Model		
Parameter	Overconfidence	Noisy Memory	Extrapolation
ρ (Fix)	0.85	0.85	0.85
σ_u^2 (Fix)	0.42	0.42	0.42
σ_ϵ^2	1.84	1.21	2.40
$\hat{\sigma}_\epsilon^2$	0.85	–	–
λ	–	0.50	–
$\hat{\rho}$	–	–	0.97
σ_ν^2 (Treated)	1.31	1.75	3.96

Table 8 reports the values of our parameters used in the Overidentified GMM Test

Table 9: Results: Over-Identification Test

Moment	Data	Model		
		Overconfidence	Noisy Memory	Extrapolation
β_C (T)	0.43	0.33	0.03	0.25
β_I (T)	-0.15	-0.18	-0.25	-0.12
β_C (NT)	1.25	1.21	1.08	1.65
β_I (NT)	-0.28	-0.28	-0.28	-0.15
SD ratio (T/NT)	0.70	0.71	0.73	0.71
J-statistic		0.19	3.78	14.63
p-value		0.91	0.15	<0.01

Table 9 reports the fit of all three models from the Over-identified GMM Test.

Figure 1: Distribution Of 6-Months Ahead Inflation Forecasts

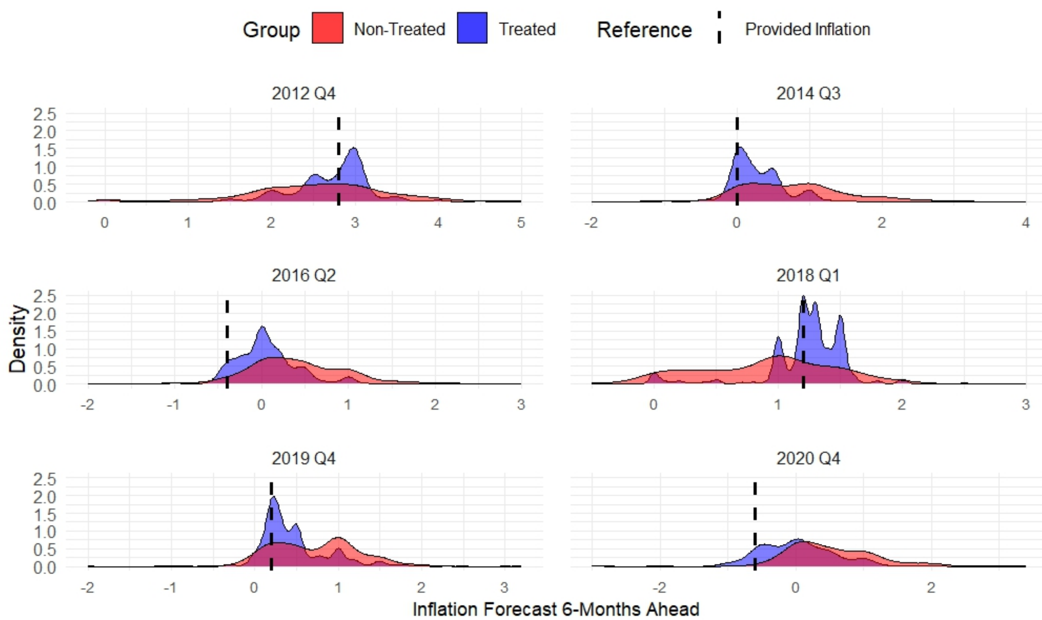


Figure 1 plots kernel density estimates of 6-month-ahead inflation expectations for treated and non-treated firms. The vertical dashed line indicates the most recent 12-month inflation rate provided to treated firms at the time of the survey. Each panel corresponds to a different survey quarter.

Figure 2: Inflation Rate and Average Inflation Forecasts: Treated Firms

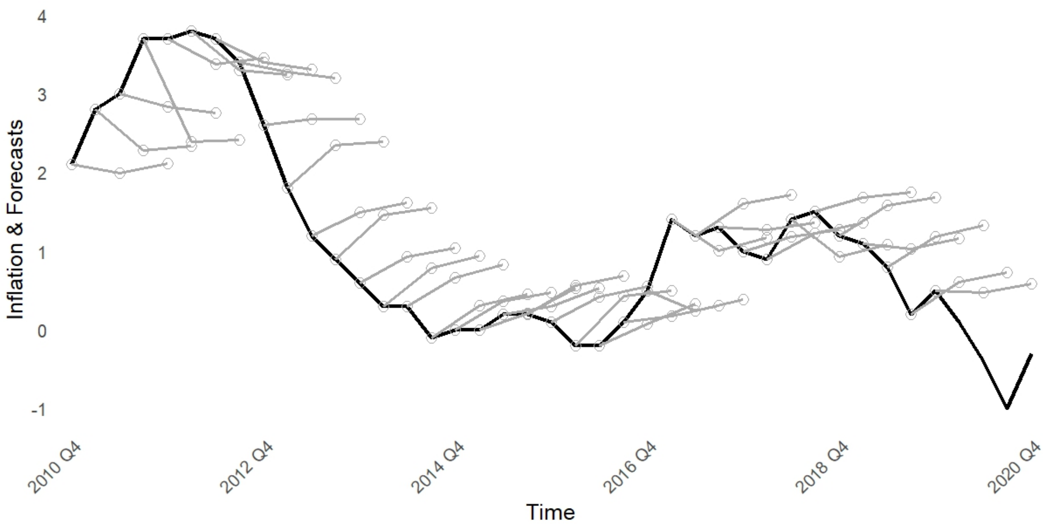


Figure 2 plots quarterly realized inflation (solid black) alongside treated firms mean 6- and 12-month-ahead inflation forecasts (gray lines with circle markers) across the sample period.

Figure 3: Inflation Rate and Average Inflation Forecasts: Non-Treated Firms

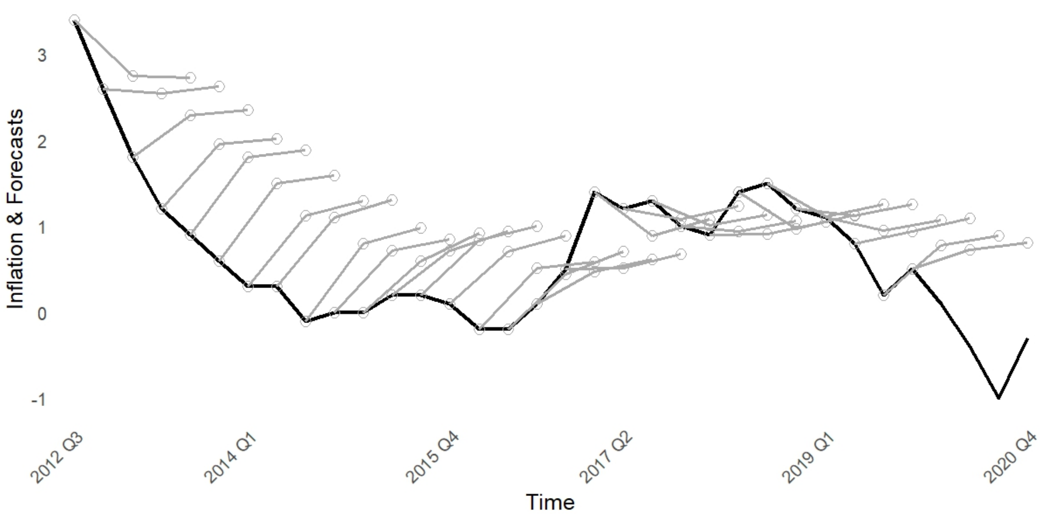
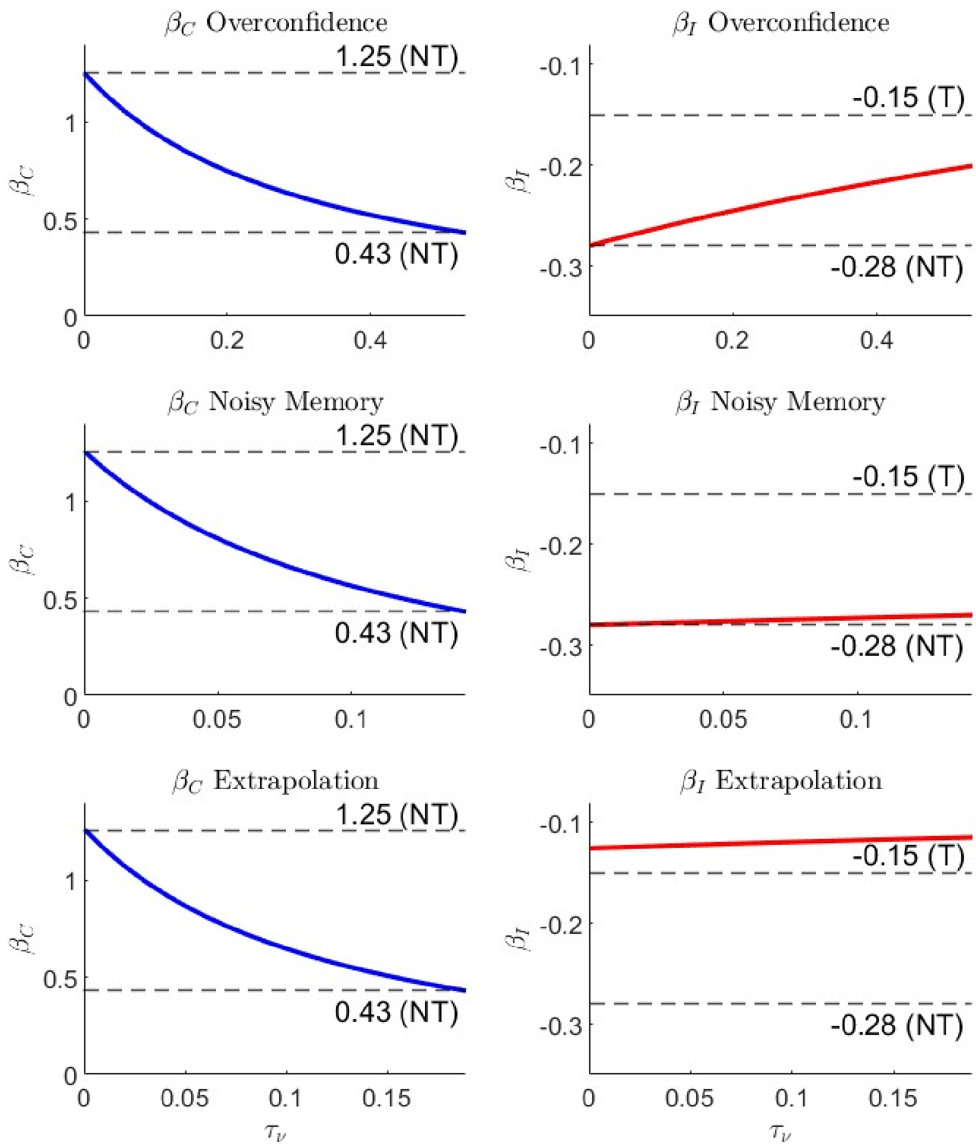


Figure 3 plots quarterly realized inflation (solid black) alongside non-treated firms mean 6- and 12-month-ahead inflation forecasts (gray lines with circle markers) across the sample period.

Figure 4: Effect of Higher Public Signal Precision on β_C and β_I under different Models



This figure shows the effect of increasing public signal precision on β_C and β_I under the three different model specifications

APPENDIX

A Descriptive Statistics

Figure A.1: Distribution Of 12-Months Ahead Inflation Forecasts

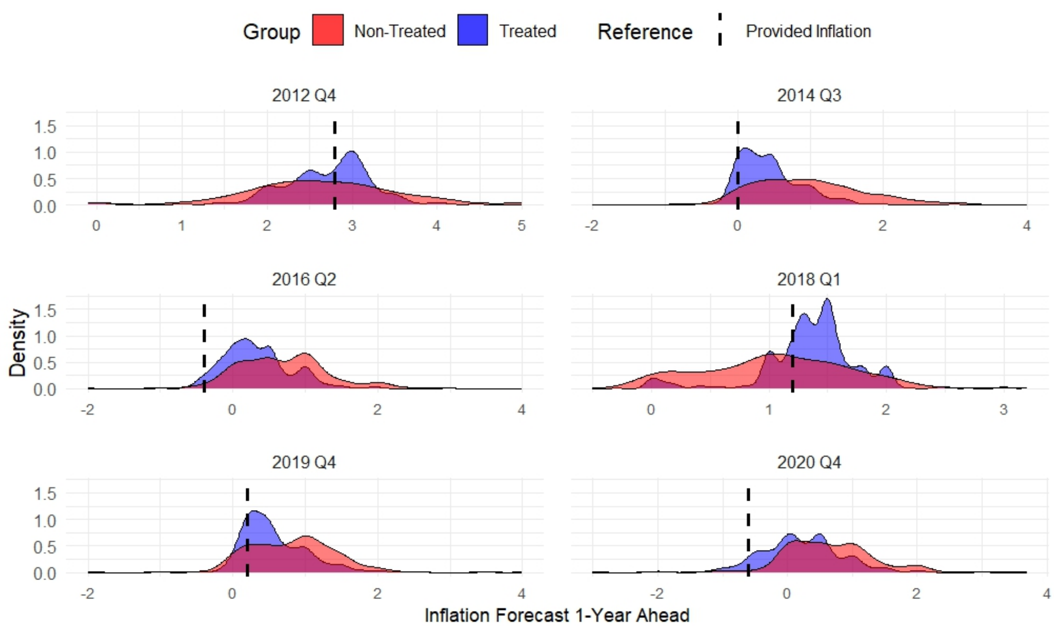


Figure A.1 plots kernel density estimates of 12-month-ahead inflation expectations for treated and non-treated firms. The vertical dashed line indicates the most recent 12-month inflation rate provided to treated firms at the time of the survey. Each panel corresponds to a different survey quarter.

B Additional Robustness Checks

B.1 Measurement Error

Here, we estimate another equation to demonstrate that our main results are robust to measurement error concerns. Specifically, Table B.1 contains the results of a regression where the dependent variable is the 1-year ahead inflation forecast error and the independent variable is the 6-month ahead inflation forecast at quarter t minus the 1-year ahead inflation forecast formed at quarter $t-2$. We estimate this specification with and without fixed effects and show that our results are qualitatively similar to our regressions at the individual level from Tables 3 and 4.

Table B.1: Individual-level Regressions: Measurement Error Robustness 2

Dependent Variable:	Individual Forecast Error 12 (Months)			
Sample:	Treated	Non-treated	Treated	Non-treated
Variables				
Forecast Revision (6 Months - 1 Year)	-0.08	-0.14**	-0.14	-0.27***
	(0.13)	(0.06)	(0.11)	(0.05)
Fixed-effects				
firm			Yes	Yes
Fit statistics				
Observations	15,246	5,792	15,246	5,792
R^2	0.00193	0.00522	0.24640	0.33571
Within R^2			0.00697	0.02414

Table B.1 addresses concerns about mechanical correlation by estimating Equation (3.2) using an alternative forecast revision measure. Specifically, it reports regressions of the 1-year ahead forecast error on forecast revisions computed as the difference between 6 months ahead inflation forecast minus the 1-year ahead inflation forecast from 2 quarters ago. Columns (1) and (2) exclude firm fixed effects; columns (3) and (4) include them. Standard errors are clustered by firm and quarter. Significance levels: *** $\Rightarrow$ p $<$ 0.01, ** $\Rightarrow$ p $<$ 0.05, * $\Rightarrow$ p $<$ 0.1.

B.2 Pre-COVID Sample Robustness

To ensure that our findings are not driven by the unusual macroeconomic environment during the COVID-19 period, we restrict the sample to the years 2010-2018. The results continue to hold, confirming that they are not specific to pandemic-era dynamics.

B.3 Drop Inattentive Firms

To ensure that our findings are not driven by treated firms reporting as their inflation forecast the same value that was given to them as treatment, we conduct another robustness check,

Table B.2: β_I Estimates: Pre-COVID Sample

Dependent Variable:	Individual Forecast Error (6 Months)			
Sample:	Treated	Non-Treated	Treated	Non-Treated
β_I	-0.15*	-0.28***	-0.15*	-0.25***
	(0.08)	(0.06)	(0.08)	(0.07)
Sample Years	Full Sample	Full Sample	Until 2019	Until 2019
Fixed-effects				
firm	Yes	Yes	Yes	Yes
Fit statistics				
Observations	15,266	5,805	11,961	4,423
R^2	0.18723	0.32974	0.21586	0.39279
Within R^2	0.01734	0.05621	0.01957	0.04752

Table B.2 reports estimates of Equation (3.2) for the full sample period and for a restricted pre-COVID sample (not including years 2019 onwards), for both treated and non-treated firms. All regressions include firm fixed effects and cluster standard errors at the firm and quarter level. This table assesses whether our main results are sensitive to the inclusion of the COVID-19 period. Significance levels: *** $\Rightarrow$ $p < 0.01$, ** $\Rightarrow$ $p < 0.05$, * $\Rightarrow$ $p < 0.1$.

where for the treated firms, we drop firm-quarter observations in the treated subsample where both the 6-month and 1-year ahead inflation forecasts are exactly equal to the inflation release that was provided to firms as part of the treatment.

Table B.3: β_I Regressions: Drop inattentive firms

Dependent Variable:	Individual Forecast Error (6 Months)		
Sample:	Treat	Non-Treated	Treat-Drop
β_I	-0.15*	-0.28***	-0.14*
	(0.08)	(0.06)	(0.08)
Fixed-effects			
firm	Yes	Yes	Yes
Fit statistics			
Observations	15,266	5,805	14,382
R^2	0.18723	0.32974	0.19303
Within R^2	0.01734	0.05621	0.01590

Table B.3 reports estimates of Equation (3.2) after excluding firms identified as inattentive : those firms in the treated sample that report 6-month and 1-year ahead inflation forecasts as exactly equal to the inflation release provided to them as treatment. All regressions include firm fixed effects, and standard errors are clustered by firm and quarter. This robustness check ensures that our results are not driven by firms that mechanically repeat past forecasts rather than responding to new information. Significance levels: *** ⇒ $p < 0.01$, ** ⇒ $p < 0.05$, * ⇒ $p < 0.1$.

B.4 Heterogeneity Analysis by Size

We conduct another robustness check to our individual level results where (3.2) is estimated separately on a sample of relatively small and large firms. Small (large) firms in any given quarter are defined as having fewer (greater) than 300 employees in that particular quarter. We estimate this specification with and without fixed effects. The results are given in Tables B.4 and B.5 and show that the results on the information treatment's effect on β_I also hold in these subsamples.

Table B.4: β_I Estimates: Small Firms

Dependent Variable:	Individual Forecast Error (6 Months)			
Sample:	Treated	Non-treated	Treated	Non-treated
β_I	-0.14	-0.21***	-0.19**	-0.29***
	(0.09)	(0.06)	(0.08)	(0.05)
Fixed-effects				
firm			Yes	Yes
Fit statistics				
Observations	9,502	3,802	9,502	3,802
R^2	0.01506	0.02537	0.22215	0.35438
Within R^2			0.03098	0.06305

Table B.4 reports estimates of equation (3.2), with and without firm fixed effects, for both treated and non-treated firms with less than 300 employees. Columns (1) and (2) estimate the regression without fixed effects, while columns (3) and (4) include firm fixed effects. All specifications use clustered standard errors at the firm and quarter level. Significance levels: *** ⇒ $p < 0.01$, ** ⇒ $p < 0.05$, * ⇒ $p < 0.1$.

Table B.5: β_I Estimates: Large Firms

Dependent Variable:	Individual Forecast Error (6 Months)			
Sample:	Treated	Non-treated	Treated	Non-treated
β_I	-0.02	-0.19***	-0.07	-0.26***
	(0.09)	(0.07)	(0.08)	(0.06)
Fixed-effects				
firm			Yes	Yes
Fit statistics				
Observations	5,487	1,901	5,487	1,901
R^2	0.00044	0.01903	0.16483	0.31884
Within R^2			0.00425	0.04344

Table B.5 reports estimates of equation (3.2), with and without firm fixed effects, for both treated and non-treated firms with more than 300 employees. Columns (1) and (2) estimate the regression without fixed effects, while columns (3) and (4) include firm fixed effects. All specifications use clustered standard errors at the firm and quarter level. Significance levels: *** ⇒ $p < 0.01$, ** ⇒ $p < 0.05$, * ⇒ $p < 0.1$.

C Proofs

This section of the appendix presents detailed derivations of the consensus and individual-level error-on-revision regression coefficients and the analytical formula for the cross sectional variance of forecasts under the assumptions detailed in the theory section. Throughout this section, we maintain the following assumptions about the data-generating process and the signal structure.

Specifically, $x_t = \rho x_{t-1} + u_t$; $u_t \sim N(0, \sigma_u^2)$. Forecasters observe an idiosyncratic noisy signal $s_t^i = x_t + \epsilon_{it}$; $\epsilon_{it} \sim N(0, \sigma_\epsilon^2)$. We model the survey's information intervention as giving treated firms access to a public signal about recent inflation at the time forecasts are formed. This public signal is denoted by: $s_t = x_t + \nu_t$; $\nu_t \sim N(0, \sigma_\nu^2)$. The non-treated case corresponds to shutting down this public signal (equivalently, letting its variance go to infinity, $\sigma_\nu^2 \to \infty$, i.e., its precision goes to zero.)[16].

C.1 Noisy Memory

Here, we assume that forecasters face constraints in their ability to recall past information. Limited ability to recall past information causes forecasters to place too much weight on current news, causing individual-level overreaction. This section adapts the model in Section 2 of Sung (2025), replacing the assumption of endogenous processing of external news with exogenous noisy signals, as in equations (2) and (3). Let m_{it} denote the agent's memory state at the start of period t, while $M_{it} = (m_{it-1}, s_{it-1}, s_{t-1})$ denotes the prior knowledge that is available to the agent before she observes the noisy signals at time t. Following Sung (2025), we assume that the mental representation of existing knowledge takes the following form:

$$m_{it} = \Lambda_t \; M_{it} + \tilde{\omega_{it}}; \quad \tilde{\omega_{it}} \sim N(0, \Sigma_{\omega t}) \tag{C.1}$$

The agent also faces a constraint on the mutual information between m_{it} and M_{it}:

$$\mathcal{I}(m_{it}, M_{it}) \leq -\frac{\log \phi_m}{2} \tag{C.2}$$

[16] This public-signal representation is a simple way to model the fact that the survey makes recent inflation information salient at the time of forecasting. In a costly information-processing or rational-inattention interpretation, the intervention can be viewed as lowering the cost of acquiring/processing publicly available inflation data.

In this case, the optimal memory representation takes the following form:

$$m_{it} = \lambda_t \, \mathbb{E}[x_t | M_{it}] + \omega_{it}; \quad \omega_{it} \sim N(0, \sigma_{\omega_t}) \tag{C.3}$$

where $\lambda_t \in [0, 1]$. Consistent with our earlier notation, $\mathbb{E}[x_t|M_{it}] = x^i_{t|t-1}$ In order to ensure that $\sigma_{\omega t}$ is fully determined by λ_t, we assume the following normalization:

$$\mathrm{Cov}[\mathbb{E}[x_t|M_{it} - m_{it} \ , \ m_{it}]] = 0 \tag{C.4}$$

which implies that $\sigma^2_{\omega t} = \lambda_t(1 - \lambda_t) \, \mathrm{Var}(\mathbb{E}[x_t|M_{it}])$. It can then be shown that:

$$\mathcal{I}(m_{it}, M_{it}) \leq -\frac{\log(1 - \lambda_t)}{2} \Rightarrow \lambda_t = \lambda = 1 - \phi_m \tag{C.5}$$

Following Proposition 2 of Sung (2025), the prior distribution of x_t conditional on memory state m_{it} (before observing noisy signals)

$$x_t | m_{it} \sim N(x^{i,m}_{t|t}, \Sigma^{i,m}_{t|t}) \tag{C.6}$$

$$x^{i,m}_{t|t} = \lambda x^i_{t|t-1} + \omega_{it} \ ; \tag{C.7}$$

$$\Sigma^{i,m}_{t|t} = (1 - \lambda) \frac{\sigma^2_u}{1 - \rho^2} + \lambda \Sigma^i_{t|t-1} \tag{C.8}$$

Standard Kalman filtering with normal priors and normal shocks implies that the posterior distribution of x_t conditional on memory state m_{it}, private signal s_{it} and public signal s_t is:

$$x_t | (m_{it}, s_{it}, s_t) \sim N(x^i_{t|t}, \Sigma^i_{t|t}) \tag{C.9}$$

$$x^i_{t|t} = x^{i,m}_{t|t} + K_{1t}(s_{it} - x^{i,m}_{t|t}) + K_{2t}(s_t - x^{i,m}_{t|t}) \tag{C.10}$$

$$\Sigma^i_{t|t} = (1 - K_t) \Sigma^{i,m}_{t|t} \tag{C.11}$$

$$K_{1t} = \frac{\Sigma^{i,m}_{t|t} \sigma^2_\nu}{\Sigma^{i,m}_{t|t} \, (\sigma^2_\nu + \sigma^2_\epsilon)} \tag{C.12}$$

$$K_{2t} = \frac{\Sigma^{i,m}_{t|t} \sigma^2_\epsilon}{\Sigma^{i,m}_{t|t} \, (\sigma^2_\nu + \sigma^2_\epsilon)} \tag{C.13}$$

where K_{1t} is the Kalman gain loading on the private signal, K_{2t} is the Kalman gain on the public signal and $K_t = K_{1t} + K_{2t}$. The usual Riccatti equation is given by:

$$\Sigma^i_{t+1|t} = \rho^2 \Sigma^i_{t|t} + \sigma^2_u = \rho^2(1 - K_t)\Sigma^{i,m}_{t|t} + \sigma^2_u \tag{C.14}$$

We solve for the steady-state variance by substituting the expression for expression for $\Sigma^{i,m}_{t|t}$ from equation (C.8) into the above equation and imposing $\Sigma^i_{t+1|t} = \Sigma^i_{t|t-1} = \Sigma$. The steady-state variance is then given by the following expression:

$$\Sigma = \frac{-B + \sqrt{B^2 - 4AC}}{2A}, \tag{C.15}$$

$$A = \lambda \left(\sigma^2_\epsilon + \sigma^2_\nu\right), \tag{C.16}$$

$$B = \left(\sigma^2_\epsilon + \sigma^2_\nu\right)\left[(1-\lambda)\sigma^2_u - \lambda\sigma^2_x\right] + \sigma^2_\epsilon\sigma^2_\nu\left[1 - \lambda\rho^2\right], \tag{C.17}$$

$$C = -\left[(1-\lambda)\sigma^2_u\sigma^2_x\left(\sigma^2_\epsilon + \sigma^2_\nu\right) + \sigma^2_\epsilon\sigma^2_\nu\left(\sigma^2_u + (1-\lambda)\sigma^2_x\right)\right], \tag{C.18}$$

where $\sigma^2_x = \frac{\sigma^2_u}{1-\rho^2}$. Plugging the steady state variance $\Sigma^i_{t|t-1} = \Sigma$ into the expression for $\Sigma^{i,m}_{t|t}$ and subsequently into K_{1t} and K_{2t} allows us to derive the steady-state Kalman gains K_1 and K_2. Given the steady state Kalman gains, we can write $F_{i,t}(x_t) = x^i_{t|t}$ using the standard Kalman updating equation:

$$\begin{aligned} x^i_{t|t} &= (1-K)\lambda x^i_{t|t-1} + Kx_t + K_1\epsilon_{it} + K_2\nu_t + (1-K)\omega_{it} &\text{(C.19)}\\ &= \left[(1-K)\lambda + K\right]x^i_{t|t-1} + K\left(x_t - x^i_{t|t-1}\right) + K_1\epsilon_{it} + K_2\nu_t + (1-K)\omega_{it}. &\text{(C.20)}\end{aligned}$$

The variance of the forecast revision can be expressed as follows:

$$\operatorname{Var}\left(x^i_{t|t} - x^i_{t|t-1}\right) = \operatorname{Var}(x^i_{t|t}) + \operatorname{Var}(x^i_{t|t-1}) - 2\operatorname{Cov}\left(x^i_{t|t}, x^i_{t|t-1}\right) \tag{C.21}$$

Using the expression for $x^i_{t|t}$:

$$\operatorname{Cov}\left(x^i_{t|t}, x^i_{t|t-1}\right) = \left[(1-K)\lambda + K\right]\operatorname{Var}\left(x^i_{t|t-1}\right). \tag{C.22}$$

Using $\text{Var}(x^i_{t+1|t}) = \text{Var}(x_{t|t-1})^i = \rho^2 \, \text{Var}(x^i_{t|t})$, it then follows that:

$$\text{Var}\left(x^i_{t|t} - x^i_{t-1|t-1}\right) = \text{Var}\left(x^i_{t|t-1}\right) \, \left[\rho^{-2} + 1 - 2\big((1-K)\lambda + K\big)\right]. \tag{C.23}$$

Next, note that contemporaneous forecast error is given by:

$$x_t - x^i_{t|t} = (1-K)x_t - (1-K)\lambda x^i_{t|t-1} - K_1\epsilon_{it} - K_2\nu_t - (1-K)\omega_{it}. \tag{C.24}$$

Taking the covariance between the forecast error $(x_t - x^i_{t|t})$ and the forecast revision $(x^i_{t|t} - x^i_{t|t-1})$ and using the expression for $x_t - x^i_{t|t}$ gives:

$$\begin{aligned} \text{Cov}\left(x_t - x^i_{t|t}, \, x^i_{t|t} - x^i_{t|t-1}\right) &= -\,\text{Cov}\left(x_t - x^i_{t|t}, x^i_{t|t-1}\right) && \text{(C.25)} \\ &= -(1-K)(1-\lambda) \; \text{Var}\left(x^i_{t|t-1}\right). && \text{(C.26)} \end{aligned}$$

Combining the above results, the individual-level error-on-revision coefficient is:

$$\begin{aligned} \beta_I &= \frac{\text{Cov}\left(x_{t+h} - x^i_{t+h|t}, \, x^i_{t+h|t} - x^i_{t+h|t-1}\right)}{\text{Var}\left(x^i_{t+h|t} - x^i_{t+h|t-1}\right)} && \text{(C.27)} \\ &= \frac{\text{Cov}\left(x_t - x^i_{t|t}, \, x^i_{t|t} - x^i_{t-1|t-1}\right)}{\text{Var}\left(x^i_{t|t} - x^i_{t-1|t-1}\right)} && \text{(C.28)} \\ &= -\,\frac{(1-K)(1-\lambda)}{\rho^{-2} + 1 - 2\big[(1-K)\lambda + K\big]}. && \text{(C.29)} \end{aligned}$$

Integrating (C.24) across forecasters i, we can write down the consensus forecast error:

$$x_t - x_{t|t} = (1-K)(x_t - x_{t|t-1}) + (1-\lambda)(1-K)x_{t|t-1} - K_2\nu_t \tag{C.30}$$

Subtracting $x^i_{t|t-1}$ from equation (C.19) and integrating over i, we can derive the consensus forecast revision:

$$x_{t|t} - x_{t|t-1} = (1-K)(\lambda-1)x_{t|t-1} + K(x_t - x_{t|t-1}) + K_2\ \nu_t \tag{C.31}$$

$$\Rightarrow \frac{x_{t|t} - x_{t|t-1}}{K} = \frac{(1-K)(\lambda-1)x_{t|t-1}}{K} + \left(x_t - x_{t|t-1}\right) + \frac{K_2\nu_t}{K} \tag{C.32}$$

$$\Rightarrow (x_{t|t} - x_{t|t-1})\left(\frac{1-K}{K}\right) = x_t - x_{t|t} - \frac{(1-\lambda)(1-K)}{K}x_{t|t-1} + \frac{K_2}{K}\nu_t \tag{C.33}$$

$$\Rightarrow x_t - x_{t|t} = \frac{1-K}{K}\left(x_{t|t} - x_{t|t-1} + (1-\lambda)x_{t|t-1}\right) - \frac{K_2}{K}\ \nu_t \tag{C.34}$$

where the third line follows from adding and subtracting $x_{t|t}$ on the RHS and rearranging. Then, β_C is given by the following expression:

$$\beta_C = \frac{\text{Cov}\left(x_{t+h} - x_{t+h|t},\, x_{t+h|t} - x_{t+h|t-1}\right)}{\text{Var}\left(x_{t+h|t} - x_{t+h|t-1}\right)} \tag{C.35}$$

$$= \frac{\text{Cov}\left(x_t - x_{t|t},\, x_{t|t} - x_{t|t-1}\right)}{\text{Var}\left(x_{t|t} - x_{t|t-1}\right)} \tag{C.36}$$

$$= \frac{\frac{1-K}{K}\text{Var}\left(x_{t|t} - x_{t|t-1}\right) + \frac{1-K}{K}(1-\lambda)\,\text{Cov}\left(x_{t|t} - x_{t|t-1}, x_{t|t-1}\right) - \frac{K_2^2}{K}\sigma_\nu^2}{\text{Var}\left(x_{t|t} - x_{t|t-1}\right)} \tag{C.37}$$

$$= \frac{1-K}{K} + \frac{1-K}{K}(1-\lambda)\frac{\text{Cov}\left(x_{t|t} - x_{t|t-1}, x_{t|t-1}\right)}{\text{Var}\left(x_{t|t} - x_{t|t-1}\right)} - \frac{K_2^2}{K}\frac{\sigma_\nu^2}{\text{Var}\left(x_{t|t} - x_{t|t-1}\right)} \tag{C.38}$$

which is the expression given in equation (5.10), with $\text{Cov}\left(x_{t|t} - x_{t|t-1}, x_{t|t-1}\right) = C_{x_{t|t-1},agg}$ and $\text{Var}\left(x_{t|t} - x_{t|t-1}\right) = V_{agg}$ (also used is the fact that $\text{Cov}\left(\nu_t\ ,\ x_{t|t} - x_{t|t-1}\right) = K_2\ \sigma_\nu^2$). Furthermore, using equation (C.31) we can write:

$$\text{Cov}\left(x_{t|t} - x_{t|t-1}, x_{t|t-1}\right) \tag{C.39}$$

$$= \text{Cov}\left((1-K)(\lambda-1)x_{t|t-1} + K(x_t - x_{t|t-1}) + K_2\ \nu_t, x_{t|t-1}\right) \tag{C.40}$$

$$= K\,\text{Cov}\left(x_t - x_{t|t-1}, x_{t|t-1}\right) + (\lambda-1)(1-K)\,\text{Var}(x_{t|t-1}) \tag{C.41}$$

Again, using equation (C.31), we can write:

$$\text{Var}\left(x_{t|t} - x_{t|t-1}\right) \tag{C.42}$$

$$= K^2 \text{Var}(x_t - x_{t|t-1}) + (1-K)^2(\lambda-1)^2 \text{Var}(x_{t|t-1}) + K_2^2\sigma_\nu^2 + 2\,\text{Cov}\left(x_t - x_{t|t-1}, x_{t|t-1}\right) \tag{C.43}$$

The three expressions we need to obtain expressions to calculate β_C are: $\text{Var}(x_t - x_{t|t-1})$, $\text{Cov}\left(x_t - x_{t|t-1}, x_{t|t-1}\right)$ and $\text{Var}(x_{t|t-1})$. Let $x_t - x_{t|t-1} = a_t$ and $x_{t|t-1} = b_t$. It can then be shown that:

$$a_{t+1} = \rho(1-K)a_t + \rho(1-K)(1-\lambda)b_t + u_{t+1} - \rho K_2\ \nu_t \tag{C.44}$$

$$b_{t+1} = \rho\ K\ a_t + \rho\left(K + \lambda(1-K)\right)b_t + \rho\ K_2\ \nu_t \tag{C.45}$$

Writing this system as a VAR:

$$\begin{bmatrix} a_{t+1} \\ b_{t+1} \end{bmatrix} = \begin{bmatrix} \rho(1-K) & \rho(1-K)(1-\lambda) \\ \rho K & \rho\left[K + \lambda(1-K)\right] \end{bmatrix} \begin{bmatrix} a_t \\ b_t \end{bmatrix} + \begin{bmatrix} 1 & -\rho k_2 \\ 0 & \rho k_2 \end{bmatrix} \begin{bmatrix} u_{t+1} \\ v_t \end{bmatrix} \tag{C.46}$$

Solving for the steady-state variance of this 2x2 system allows us to obtain the expressions needed to calculate β_C, namely, $\text{Cov}(a_t, b_t)$, $\text{Var}(a_t)$ and $\text{Var}(b_t)$[17]. Finally, we also show how to calculate the forecast dispersion. Integrating equation (C.19) across i and subtracting it from $x_{t|t}$, we have:

$$x^i_{t|t} - x_{t|t} = (1-K)\lambda(x^i_{t|t-1} - x_{t|t-1}) + K_1\epsilon_{it} + (1-K)\omega_{it} \tag{C.47}$$

$$= (1-K)\lambda\ \rho\ (x^i_{t-1|t-1} - x_{t-1|t-1}) + K_1\epsilon_{it} + (1-K)\omega_{it} \tag{C.48}$$

$$\text{Var}(x^i_{t|t} - x_{t|t}) = \frac{K_1^2\sigma_\epsilon^2 + (1-K)^2\sigma_\omega^2}{1 - (1-K)^2\lambda^2\rho^2} \tag{C.49}$$

$$\Rightarrow \text{Var}(x^i_{t+h|t} - x_{t+h|t}) = \text{Var}\left(\rho^h(x^i_{t|t} - x_{t|t})\right) \tag{C.50}$$

$$= \rho^{2h}\,\text{Var}(x^i_{t|t} - x_{t|t}) \tag{C.51}$$

[17]We call a matlab solver to solve this 2x2 system

C.2 Overconfidence

We now derive the consensus and individual error-on-revision regression coefficient in the presence of idiosyncratic private signals and a public signal, when forecasters perceive the variance of their idiosyncratic signal $\hat{\sigma}_\epsilon^2$ to be lower than its true variance σ_ϵ^2. Given equations (5.1), (5.2) and (5.3), we can write the measurement and state equations of the system as:

$$y_t^i = Hx_t + w_t, \tag{C.52}$$

$$x_t = \rho x_{t-1} + u_t \tag{C.53}$$

where $y_t^i = [s_t^i, s_t]', H = [1, \quad 1]', w_t = [\epsilon_{it}, \quad \nu_t]'$ Let $\Sigma_{t|t-1}^i = \mathbb{E}\left[\left(x_t - x_{t|t-1}^i\right)^2\right]$ be their prior variance belief and $K_t^v = [K_{1t}, K_{2t}]'$ denote the vector of Kalman gains. The Kalman recursion satisfies the following equations:

$$K_t^v = \Sigma_{t|t-1}^i H' \left(H\Sigma_{t|t-1}^i H' + R_{\text{perceived}}\right)^{-1}, \tag{C.54}$$

$$\Sigma_{t|t}^i = (1 - K_t^v \; H)\, \Sigma_{t|t-1}^i, \tag{C.55}$$

$$\Sigma_{t+1|t}^i = \rho^2 \Sigma_{t|t}^i + \sigma_u^2. \tag{C.56}$$

with the matrix $R_{\text{perceived}}$ given by the following expression:

$$\begin{bmatrix} \hat{\sigma}_\epsilon^2 & 0 \\ 0 & \sigma_\nu^2 \end{bmatrix} \tag{C.57}$$

Using equations from the Kalman recursion, the steady-state prior variance $\Sigma_{t|t-1}^i = \Sigma$ is given by the following expression:

$$\Sigma = \frac{\sigma_u^2\left(\hat{\sigma}_\epsilon^2 + \sigma_\nu^2\right) - (1-\rho^2)\sigma_\nu^2\hat{\sigma}_\epsilon^2 + \sqrt{\left[\sigma_u^2\left(\hat{\sigma}_\epsilon^2 + \sigma_\nu^2\right) - (1-\rho^2)\hat{\sigma}_\epsilon^2\sigma_\nu^2\right]^2 + 4\left(\hat{\sigma}_\epsilon^2 + \sigma_\nu^2\right)\sigma_\nu^2\sigma_u^2\hat{\sigma}_\epsilon^2}}{2\left(\hat{\sigma}_\epsilon^2 + \sigma_\nu^2\right)} \tag{C.58}$$

The steady state (distorted) Kalman gains on the public and private signals are then given by the following expressions:

$$K_1 = \frac{\Sigma\,\sigma_\nu^2}{(\hat{\sigma}_\epsilon^2 + \sigma_\nu^2)\,\Sigma + \hat{\sigma}_\epsilon^2 \sigma_\nu^2}, \tag{C.59}$$

$$K_2 = \frac{\Sigma\,\hat{\sigma}_\epsilon^2}{(\hat{\sigma}_\epsilon^2 + \sigma_\nu^2)\,\Sigma + \hat{\sigma}_\epsilon^2 \sigma_\nu^2}. \tag{C.60}$$

The Kalman updating formula at the steady-state Kalman gains gives us the following expression:

$$x^i_{t|t} = x^i_{t|t-1} + K_1\left(s^i_t - x^i_{t|t-1}\right) + K_2\left(s_t - x^i_{t|t-1}\right) \tag{C.61}$$

$$\Rightarrow x^i_{t|t} - x^i_{t|t-1} = K(x_t - x^i_{t|t-1}) + K_1 \epsilon_{it} + K_2\ \nu_t \tag{C.62}$$

$$\Rightarrow \frac{x^i_{t|t} - x^i_{t|t-1}}{K} - \frac{K_1}{K}\epsilon_{it} - \frac{K_2}{K}\nu_t = x_t - x^i_{t|t-1} \tag{C.63}$$

$$\Rightarrow x_t - x^i_{t|t} = (x^i_{t|t} - x^i_{t|t-1})\left(\frac{1-K}{K}\right) - \frac{K_1}{K}\epsilon_{it} - \frac{K_2}{K}\nu_t \tag{C.64}$$

We can now derive the expression for β_I using equation (C.64):

$$\beta_I = \frac{\text{Cov}\left(x_{t+h} - x^i_{t+h|t},\, x^i_{t+h|t} - x^i_{t+h|t-1}\right)}{\text{Var}\left(x^i_{t+h|t} - x^i_{t+h|t-1}\right)} \tag{C.65}$$

$$= \frac{\text{Cov}\left(x_t - x^i_{t|t},\, x^i_{t|t} - x^i_{t-1|t-1}\right)}{\text{Var}\left(x^i_{t|t} - x^i_{t-1|t-1}\right)} \tag{C.66}$$

$$= \frac{1-K}{K} - \frac{K_1^2 \sigma_\epsilon^2}{K\,V_{ind}} - \frac{K_2^2 \sigma_\nu^2}{K\,V_{ind}}. \tag{C.67}$$

where $V_{ind} = \text{Var}\left(x^i_{t|t} - x^i_{t-1|t-1}\right)$. We have also used: $\text{Cov}\left(\epsilon_{it}, x^i_{t|t} - x^i_{t|t-1}\right) = K_1 \sigma^2_\epsilon$ and $\text{Cov}\left(\nu_t, x^i_{t|t} - x^i_{t|t-1}\right) = K_2\ \sigma^2_\nu$. Finally, we can also show that:

$$x_{t+1} - x^i_{t+1|t} = \rho\left(x_t - x^i_{t|t}\right) + u_{t+1} \tag{C.68}$$

$$= \rho\left[(1-K)\left(x_t - x^i_{t|t-1}\right) - K_1\epsilon_{it} - K_2\nu_t\right] + u_{t+1} \tag{C.69}$$

$$\Rightarrow \quad \text{Var}\left(x_{t+1} - x^i_{t+1|t}\right) = \rho^2(1-K)^2\, \text{Var}\left(x_t - x^i_{t|t-1}\right) + \rho^2 K_1^2\sigma^2_\epsilon + \rho^2 K_2^2 \sigma^2_\nu + \sigma^2_u \tag{C.70}$$

$$\Rightarrow \quad \text{Var}\left(x_t - x^i_{t|t-1}\right) = \frac{\rho^2\left[K_1^2\sigma^2_\epsilon + K_2^2\sigma^2_\nu\right]}{1-\rho^2(1-K)^2} + \sigma^2_u \tag{C.71}$$

$$\text{Var}\left(x^i_{t|t} - x^i_{t|t-1}\right) = K^2\, \text{Var}\left(x_t - x^i_{t|t-1}\right) + K_1^2\sigma^2_\epsilon + K_2^2\sigma^2_\nu \tag{C.72}$$

$$= K^2\left[\frac{\rho^2\left(K_1^2\sigma^2_\epsilon + K_2^2\sigma^2_\nu\right) + \sigma^2_u}{1-\rho^2(1-K)^2}\right] + K_1^2\sigma^2_\epsilon + K_2^2\sigma^2_\nu \tag{C.73}$$

$$= V_{ind} \tag{C.74}$$

Integrating equation (C.64) over i and following very similar steps as above, we can write:

$$\beta_C = \frac{1-K}{K} - \frac{K_2^2\sigma^2_\nu}{K\ V_{agg}}. \tag{C.75}$$

$$V_{agg} = \text{Var}(x_{t|t} - x_{t|t-1}) \tag{C.76}$$

$$= K^2\left[\frac{\rho^2 K_2^2\sigma^2_\nu + \sigma^2_u}{1-\rho^2(1-K)^2}\right] + K_2^2\sigma^2_\nu \tag{C.77}$$

Finally, we also show how to calculate the cross-sectional variance of forecasts. From equation (C.61), we integrate across i to get $x_{t|t}$. Subtracting $x_{t|t}$ from $x^i_{t|t}$, we have:

$$x^i_{t|t} - x_{t|t} = (1-K)(x^i_{t|t-1} - x_{t|t-1}) + K_1\epsilon_{it} \tag{C.78}$$

$$= (1-K)\ \rho\ (x^i_{t-1|t-1} - x_{t-1|t-1}) + K_1\epsilon_{it} \tag{C.79}$$

$$\Rightarrow \text{Var}(x^i_{t|t} - x_{t|t}) = \frac{K_1^2\sigma^2_\epsilon}{1-(1-K)^2\ \rho^2} \tag{C.80}$$

$$\Rightarrow \text{Var}(x^i_{t+h|t} - x_{t+h|t}) = \text{Var}\left(\rho^h(x^i_{t|t} - x_{t|t})\right) \tag{C.81}$$

$$= \rho^{2h}\, \text{Var}(x^i_{t|t} - x_{t|t}) \tag{C.82}$$

In our GMM exercises we target the ratio of the standard deviation of forecasts between treated and non-treated firms to discipline the value of σ_ν^2.

C.3 Over-extrapolation

We now derive the consensus and individual error-on-revision regression coefficients when agents misperceive the persistence of the process to be $\hat{\rho}$, where actual persistence is larger than perceived persistence. In this case, the measurement and state equations of the system can be written as:

$$y_t^i = Hx_t + w_t, \tag{C.83}$$

$$x_t = \hat{\rho} x_{t-1} + u_t \tag{C.84}$$

where $y_t^i = [s_t^i, s_t]', H = [1, \quad 1]', w_t = [\epsilon_{it}, \quad \nu_t]'$ Let $\Sigma_{t|t-1}^i = \mathbb{E}\left[\left(x_t - x_{t|t-1}^i\right)^2\right]$ be their prior variance belief and $K_t^v = [K_{1t}, K_{2t}]'$ denote the vector of Kalman gains. The Kalman recursion satisfies the following equations:

$$K_t^v = \Sigma_{t|t-1}^i H' \left(H\Sigma_{t|t-1}^i H' + R\right)^{-1}, \tag{C.85}$$

$$\Sigma_{t|t}^i = (1 - K_t^v \; H)\, \Sigma_{t|t-1}^i, \tag{C.86}$$

$$\Sigma_{t+1|t}^i = \hat{\rho}^2 \Sigma_{t|t}^i + \sigma_u^2. \tag{C.87}$$

with the matrix R given by the following expression:

$$\begin{bmatrix} \sigma_\epsilon^2 & 0 \\ 0 & \sigma_\nu^2 \end{bmatrix} \tag{C.88}$$

Using equations from the Kalman recursion, the steady-state prior variance $\Sigma_{t|t-1}^i = \Sigma$ is given by the following expression:

$$\Sigma = \frac{\sigma_u^2\left(\sigma_\epsilon^2 + \sigma_\nu^2\right) - (1-\hat{\rho}^2)\sigma_\nu^2\sigma_\epsilon^2 + \sqrt{\left[\sigma_u^2\left(\sigma_\epsilon^2 + \sigma_\nu^2\right) - (1-\hat{\rho}^2)\sigma_\epsilon^2\sigma_\nu^2\right]^2 + 4\left(\sigma_\epsilon^2 + \sigma_\nu^2\right)\sigma_\nu^2\sigma_u^2\sigma_\epsilon^2}}{2\left(\sigma_\epsilon^2 + \sigma_\nu^2\right)} \tag{C.89}$$

which is the same as equation (C.58), which σ_ϵ^2 in place of $\hat{\sigma}_\epsilon^2$ and $\hat{\rho}$ in place of ρ The steady state (distorted) Kalman gains on the public and private signals are then given by the following expressions:

$$K_1 = \frac{\Sigma\,\sigma_\nu^2}{(\sigma_\epsilon^2 + \sigma_\nu^2)\,\Sigma + \sigma_\epsilon^2\sigma_\nu^2}, \tag{C.90}$$

$$K_2 = \frac{\Sigma\,\sigma_\epsilon^2}{(\sigma_\epsilon^2 + \sigma_\nu^2)\,\Sigma + \sigma_\epsilon^2\sigma_\nu^2}. \tag{C.91}$$

The Kalman updating formula at the steady-state Kalman gains gives us the following expression:

$$x_{t|t}^i = x_{t|t-1}^i + K_1(s_t^i - x_{t|t-1}^i) + K_2(s_t - x_{t|t-1}^i) \tag{C.92}$$

$$\Rightarrow \frac{x_{t|t}^i - x_{t|t-1}^i}{K} = x_t - x_{t|t-1}^i + \frac{K_1}{K}\,\epsilon_{it} + \frac{K_2}{K}\nu_t \tag{C.93}$$

$$\Rightarrow (x_{t|t}^i - x_{t|t-1}^i)\left(\frac{1-K}{K}\right) = x_t - x_{t|t}^i + \frac{K_1}{K}\epsilon_{it} + \frac{K_2}{K}\nu_t \tag{C.94}$$

The expression for the individual level coefficient, β_I^h is given by:

$$\beta_I^h = \frac{\text{Cov}\left[\left(x^i_{t+h} - x^i_{t+h|t}\right), \left(x^i_{t+h|t} - x^i_{t+h|t-1}\right)\right]}{\text{Var}(x^i_{t+h|t} - x^i_{t+h|t-1})} \tag{C.95}$$

$$= \frac{\text{Cov}\left[\rho^h x_t - \hat{\rho}^h x^i_{t|t},\ \hat{\rho}^h(x^i_{t|t} - x^i_{t|t-1})\right]}{\hat{\rho}^{2h}\ \text{Var}\left(x^i_{t|t} - x^i_{t|t-1}\right)} \tag{C.96}$$

$$= \frac{\text{Cov}\left[\left(\rho^h - \hat{\rho}^h + \hat{\rho}^h\right) x_t -\ \hat{\rho}^h x^i_{t|t},\ \hat{\rho}^h \left(x^i_{t|t} - x^i_{t|t-1}\right)\right]}{\hat{\rho}^{2h}\ \text{Var}\left(x^i_{t|t} - x^i_{t|t-1}\right)} \tag{C.97}$$

$$= \frac{\text{Cov}\left[(\rho^h - \hat{\rho}^h)x_t,\ \hat{\rho}^h(x^i_{t|t} - x_{t|t-1})\right]}{\hat{\rho}^{2h}\ \text{Var}\left(x^i_{t|t} - x^i_{t|t-1}\right)} + \frac{\text{Cov}\left[\hat{\rho}^h(x_t - x^i_{t|t}),\ \hat{\rho}^h(x^i_{t|t} - x^i_{t|t-1})\right]}{\hat{\rho}^{2h}\ \text{Var}\left(x^i_{t|t} - x_{t|t-1}\right)} \tag{C.98}$$

$$= \frac{(\rho^h - \hat{\rho}^h)\ \text{Cov}\left[x_t,\ x^i_{t|t} - x^i_{t|t-1}\right]}{\hat{\rho}^h\ \text{Var}\left(x^i_{t|t} - x^i_{t|t-1}\right)} + \frac{\text{Cov}\left[(x_t - x^i_{t|t}),\ x^i_{t|t} - x_{t|t-1}\right]}{\text{Var}\left(x^i_{t|t} - x^i_{t|t-1}\right)} \tag{C.99}$$

Using equation (C.94), it can be shown that

$$= \frac{\text{Cov}\left(x_t - x^i_{t|t},\ x^i_{t|t} - x^i_{t-1|t-1}\right)}{\text{Var}\left(x^i_{t|t} - x^i_{t-1|t-1}\right)} \tag{C.100}$$

$$= \frac{1-K}{K} - \frac{K_1^2\sigma_\epsilon^2}{K\,V_{ind}} - \frac{K_2^2\sigma_\nu^2}{K\,V_{ind}}. \tag{C.101}$$

where $V_{ind} = \text{Var}\left(x^i_{t|t} - x^i_{t-1|t-1}\right)$. Substituting the above expression into equation (C.99)

$$\beta_I^h = \frac{1-K}{K} - \frac{K_1^2\sigma_\epsilon^2}{K\,V_{ind}} - \frac{K_2^2\sigma_\nu^2}{K\,V_{ind}} + \frac{\left(\rho^h - \hat{\rho}^h\right)}{\hat{\rho}^h}\frac{C_{x,Ind}}{V_{ind}}. \tag{C.102}$$

where $C_{x,ind} = \text{Cov}(x_t, x^i_{t|t} - x^i_{t|t-1})$. We now have to calculate V_{ind} and $C_{x,ind}$ Using equation (C.92), it can be shown that[18]

$$C_{x,ind} = \text{Cov}(x_t, x^i_{t|t} - x^i_{t|t-1}) = \text{Cov}\left(x_t, K(x_t - x^i_{t|t-1}) + K_1\epsilon_{it} + K_2\ \nu_t\right) \tag{C.103}$$

$$= \text{Cov}\left(x_t - x^i_{t|t-1} + x^i_{t|t-1}, K(x_t - x^i_{t|t-1})\right) \tag{C.104}$$

$$= K\,\text{Var}(x_t - x^i_{t|t-1}) + K\,\text{Cov}(x^i_{t|t-1}, x_t - x^i_{t|t-1}) \tag{C.105}$$

Furthermore, using equation (C.92), it can be shown that

$$V_{ind} = \text{Var}(x^i_{t|t} - x^i_{t|t-1}) = K^2\,\text{Var}(x_t - x^i_{t|t-1}) + \ K_1^2\ \sigma^2_\epsilon + K_2^2\ \sigma^2_\nu \tag{C.106}$$

In order to complete the calculation, we need expressions for $\text{Var}(x_t - x^i_{t|t-1})$ and $\text{Cov}(x^i_{t|t-1}, x_t - x^i_{t|t-1})$. Let $a^i_t = x_t - x^i_{t|t-1}$ and $b^i_t = x^i_{t|t-1}$. It can then be shown that:

$$a^i_{t+1} = (\rho - \hat{\rho}K)a^i_t + (\rho - \hat{\rho})b^i_t + u_{t+1} - \hat{\rho}\ (K_1\epsilon_{it} + K_2\nu_t) \tag{C.107}$$

$$b^i_{t+1} = \hat{\rho}b^i_t + \hat{\rho}Ka^i_t + \hat{\rho}(K_1\ \epsilon_{it} + K_2\nu_t) \tag{C.108}$$

Written as a VAR, we have that:

$$\begin{bmatrix} a^i_{t+1} \\ b^i_{t+1} \end{bmatrix} = \begin{bmatrix} \rho - \hat{\rho}K & \rho - \hat{\rho} \\ \hat{\rho}K & \hat{\rho} \end{bmatrix} \begin{bmatrix} a^i_t \\ b^i_t \end{bmatrix} + \begin{bmatrix} 1 & -\hat{\rho} \\ 0 & \hat{\rho} \end{bmatrix} \begin{bmatrix} u_{t+1} \\ K_1\varepsilon_{it} + K_2\nu_t \end{bmatrix} \tag{C.109}$$

Solving for the steady state variance of this system gives us expressions for $\text{Var}(x_t - x^i_{t|t-1})$ and $\text{Cov}(x^i_{t|t-1}, x_t - x^i_{t|t-1})$, which completes our calculation of β^h_I [19].

Integrating equation (C.94) across individuals and following similar steps as above allows us

[18] With extrapolative expectations, this calculation is more involved than under overconfidence because, in the true DGP, $\text{Cov}\left(x^i_{t|t-1}, x_t - x^i_{t|t-1}\right) \neq 0$. The reason is that forecasters believe the process is more persistent than it actually is, so they systematically overpredict after high values and underpredict after low values. This leads to forecast errors that are negatively related to the forecast itself, making $\text{Cov}\left(x^i_{t|t-1}, x_t - x^i_{t|t-1}\right) < 0$

[19] In order to solve this system for the steady state variance, we call a matlab solver

to write the following expression for β_C^h:

$$\beta_C^h = \frac{\text{Cov}\left[\left(x_{t+h} - x_{t+h|t}\right), \left(x_{t+h|t} - x_{t+h|t-1}\right)\right]}{\text{Var}(x_{t+h|t} - x_{t+h|t-1})} \tag{C.110}$$

$$= \frac{1-K}{K} - \frac{K_2^2\,\sigma_\nu^2}{K\,V_{agg}} + \frac{\left(\rho^h - \hat{\rho}^h\right)}{\hat{\rho}^h}\frac{C_{x,agg}}{V_{agg}}. \tag{C.111}$$

where $C_{x,agg} = \text{Cov}(x_t, x_{t|t} - x_{t|t-1})$ and $V_{agg} = \text{Var}(x_{t|t} - x_{t|t-1})$. Also, in a similar manner to equations (C.105) and (C.115), we can show that:

$$C_{x,agg} = \text{Cov}(x_t, x_{t|t} - x_{t|t-1}) \tag{C.112}$$

$$= K\,\text{Var}(x_t - x_{t|t-1}) + K\,\text{Cov}(x_{t|t-1}, x_t - x_{t|t-1}) \tag{C.113}$$

$$V_{agg} = \text{Var}(x_{t|t}^i - x_{t|t-1}^i) \tag{C.114}$$

$$= K^2\,\text{Var}(x_t - x_{t|t-1}) + \; K_1^2\;\sigma_\epsilon^2 + K_2^2\;\sigma_\nu^2 \tag{C.115}$$

In order to calculate $\text{Var}(x_t - x_{t|t-1})$ and $\text{Cov}(x_t - x_{t|t-1}, x_{t|t-1})$, we write down the 2x2 equation system similar to (C.109). Letting $x_t - x_{t|t-1} = a_t$ and $x_{t|t-1} = b_t$, one can show that:

$$\begin{bmatrix} a_{t+1} \\ b_{t+1} \end{bmatrix} = \begin{bmatrix} \rho - \hat{\rho}K & \rho - \hat{\rho} \\ \hat{\rho}K & \hat{\rho} \end{bmatrix} \begin{bmatrix} a_t \\ b_t \end{bmatrix} + \begin{bmatrix} 1 & -\hat{\rho} \\ 0 & \hat{\rho} \end{bmatrix} \begin{bmatrix} u_{t+1} \\ K_2\;\nu_t \end{bmatrix} \tag{C.116}$$

Solving for the steady state variance of this 2x2 system allows us to obtain the expressions needed for our calculation of β_C^h[20]. Finally, we also show how to calculate the cross-sectional variance of forecasts. From equation (C.92), we integrate across i to get $x_{t|t}$. Subtracting

[20] As before, in order to solve this system for the steady state variance, we call a matlab solver

$x_{t|t}$ from $x^i_{t|t}$, we have:

$$x^i_{t|t} - x_{t|t} = (1-K)(x^i_{t|t-1} - x_{t|t-1}) + K_1 \epsilon_{it} \tag{C.117}$$

$$= (1-K)\ \hat{\rho}\ (x^i_{t-1|t-1} - x_{t-1|t-1}) + K_1 \epsilon_{it} \tag{C.118}$$

$$\Rightarrow \text{Var}(x^i_{t|t} - x_{t|t}) = \frac{K_1^2 \sigma_\epsilon^2}{1-(1-K)^2\ \hat{\rho}^2} \tag{C.119}$$

$$\Rightarrow \text{Var}(x^i_{t+h|t} - x_{t+h|t}) = \text{Var}\left(\hat{\rho}^h (x^i_{t|t} - x_{t|t})\right) \tag{C.120}$$

$$= \hat{\rho}^{2h}\,\text{Var}(x^i_{t|t} - x_{t|t}) \tag{C.121}$$

In our GMM exercises we target the ratio of the standard deviation of forecasts between treated and non-treated firms to discipline the value of σ_ν^2.

D Robustness to Alternative DGP Specifications

We report overidentification tests under alternative data-generating-process parameters (ρ, σ_u^2). The results are quite similar to those from Table (9): the Overconfidence model provides the best fit; the Over-extrapolation model is strongly rejected because it cannot match the non-treated moments; and the Noisy Memory model lies in between-matching the non-treated coefficients reasonably well but failing to reproduce the magnitude of the treated coefficients for plausible values of σ_ν^2.

To build further intuition for why the Overconfidence model provides the best fit among the alternatives we consider, we also conduct two just-identified GMM exercises. In both exercises, we calibrate σ_ϵ^2 and the model-specific behavioral parameter to match β_C and β_I for non-treated firms. In the first exercise, σ_ν^2 is calibrated to match β_C for treated firms, with the treated β_I serving as a non-targeted moment. In the second exercise, σ_ν^2 is calibrated to match β_I for treated firms, with the treated β_C serving as a non-targeted moment. The calibrated parameter values from these exercises for each model are reported in Table D.7 and the results are reported in Table D.8.

Table D.1: Calibrated Parameters, Overidentification Test, $\rho = 0.8$

Parameter	Model Overconfidence	Noisy Memory	Extrapolation
ρ (Fix)	0.80	0.80	0.80
σ_u^2 (Fix)	0.42	0.42	0.42
σ_ϵ^2	1.65	1.11	2.27
$\hat{\sigma}_\epsilon^2$	0.78	–	–
λ	–	0.28	–
$\hat{\rho}$	–	–	0.96
σ_ν^2 (Treated)	1.19	1.54	3.80

Table (D.1) reports the calibrated parameters from our overidentification test when we set the value of $\rho = 0.8$.

Table D.2: Overidentification Test, Results: $\rho = 0.8$

Moment	Data	Model Overconfidence	Noisy Memory	Extrapolation
β_C (T)	0.43	0.33	0.02	0.20
β_I (T)	-0.15	-0.18	-0.25	-0.15
β_C (NT)	1.25	1.19	1.06	1.52
β_I (NT)	-0.28	-0.29	-0.28	-0.18
SD ratio (T/NT)	0.70	0.71	0.73	0.71
J-statistic		0.25	4.17	9.46
p-value		0.88	0.12	<0.01

Table (D.2) reports the results of our overidentification test when $\rho = 0.8$.

Table D.3: Calibrated Parameters, Overidentification Test, $\rho = 0.9$

Parameter	Model		
	Overconfidence	Noisy Memory	Extrapolation
ρ (Fix)	0.90	0.90	0.90
σ_u^2 (Fix)	0.42	0.42	0.42
σ_ϵ^2	2.15	1.39	2.54
$\hat{\sigma}_\epsilon^2$	0.94	–	–
λ	–	0.71	–
$\hat{\rho}$	–	–	0.99
σ_ν^2 (Treated)	1.50	2.09	4.09

Table (D.3) reports the calibrated parameters from our overidentification test when we set the value of $\rho = 0.9$.

Table D.4: Overidentification Test, Results: $\rho = 0.9$

Moment	Data	Model		
		Overconfidence	Noisy Memory	Extrapolation
β_C (T)	0.43	0.34	0.05	0.30
β_I (T)	-0.15	-0.19	-0.25	-0.08
β_C (NT)	1.25	1.22	1.10	1.80
β_I (NT)	-0.28	-0.28	-0.28	-0.10
SD ratio (T/NT)	0.70	0.71	0.73	0.70
J-statistic		0.18	3.30	23.28
p-value		0.92	0.19	<0.01

Table (D.4) reports the results of our overidentification test when $\rho = 0.9$.

Table D.5: Calibrated Parameters, Overidentification Test, $\sigma_u^2 = 0.7$

Parameter	Model: Overconfidence	Noisy Memory	Extrapolation
ρ (Fix)	0.85	0.85	0.85
σ_u^2 (Fix)	0.70	0.70	0.70
σ_ϵ^2	3.04	2.00	3.99
$\hat{\sigma}_\epsilon^2$	1.40	–	–
λ	–	0.50	–
$\hat{\rho}$	–	–	0.97
σ_ν^2 (Treated)	2.17	2.88	6.57

Table (D.5) reports the calibrated parameters from our overidentification test when we set the value of $\sigma_u^2 = 0.7$.

Table D.6: Overidentification Test, Results: $\sigma_u^2 = 0.7$

Moment	Data	Model: Overconfidence	Noisy Memory	Extrapolation
β_C (T)	0.43	0.33	0.03	0.25
β_I (T)	-0.15	-0.18	-0.25	-0.12
β_C (NT)	1.25	1.21	1.08	1.65
β_I (NT)	-0.28	-0.28	-0.28	-0.15
SD ratio (T/NT)	0.70	0.71	0.73	0.71
J-statistic		0.19	3.78	14.57
p-value		0.91	0.15	<0.01

Table (D.6) reports the results of our overidentification test when $\sigma_u^2 = 0.7$.

Table D.7: Calibration Details: Just Identified GMM Exercises

Parameter	Model: Overconfidence	Model: Noisy Memory	Model: Extrapolation
ρ (Fix)	0.85	0.85	0.85
σ_u^2 (Fix)	0.42	0.42	0.42
σ_ϵ^2	1.89	1.39	1.59
$\hat{\sigma}_\epsilon^2$	0.88	–	–
λ	–	0.52	–
$\hat{\rho}$	–	–	0.97
σ_ν^2 (Treated, Exercise 1)	1.88	7.00	5.30
σ_ν^2 (Treated, Exercise 2)	0.92	0.26	–

Table (D.7) reports the parameters calibrated in our just-identified GMM Exercises.

Table D.8: Just-Identified GMM Exercises: Results

Panel A: Just-identified GMM, Exercise 1 (σ_ν^2 targets β_C (T))

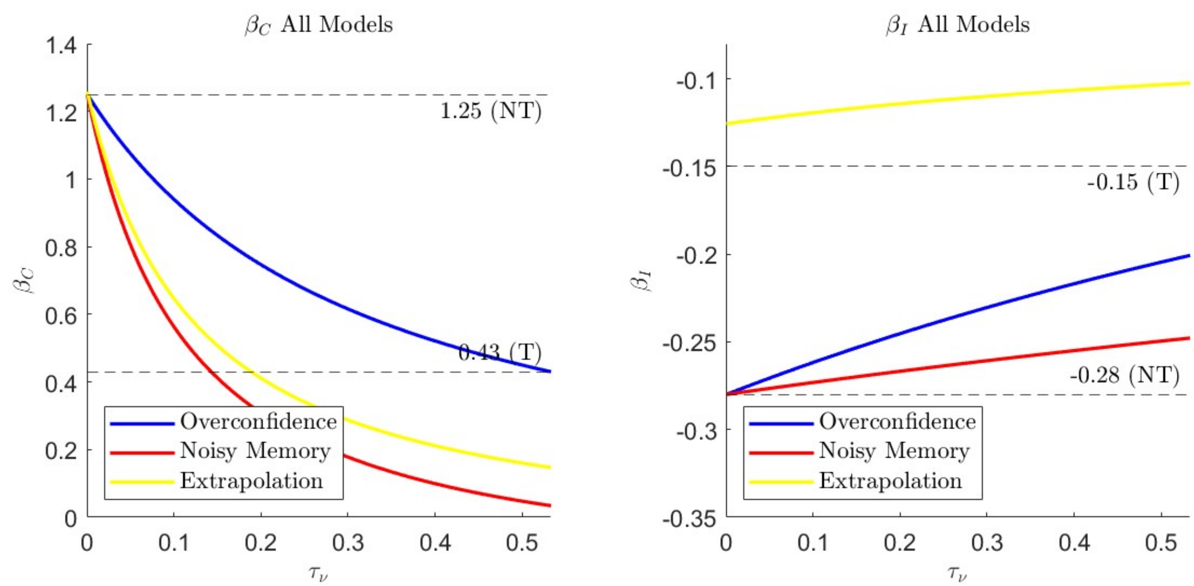
Moments	Targeted	Data	Noisy Memory	Overconfidence	Extrapolation
β_C (NT)	Yes	1.25	1.25	1.25	1.26
β_I (NT)	Yes	-0.28	-0.28	-0.28	-0.13
β_C (T)	Yes	0.43	0.43	0.43	0.43
β_I (T)	No	-0.15	-0.27	-0.20	-0.11
SD (T/NT)	No	0.70	0.91	0.77	0.82

Panel B: Just-identified GMM, Exercise 2 (σ_ν^2 targets β_I (T))

Moment	Targeted	Data	Noisy Memory	Overconfidence
β_C (NT)	Yes	1.25	1.25	1.25
β_C (T)	No	0.43	-0.11	0.24
β_I (NT)	Yes	-0.28	-0.28	-0.28
β_I (T)	Yes	-0.15	-0.15	-0.15
SD (T/NT)	No	0.70	0.29	0.62

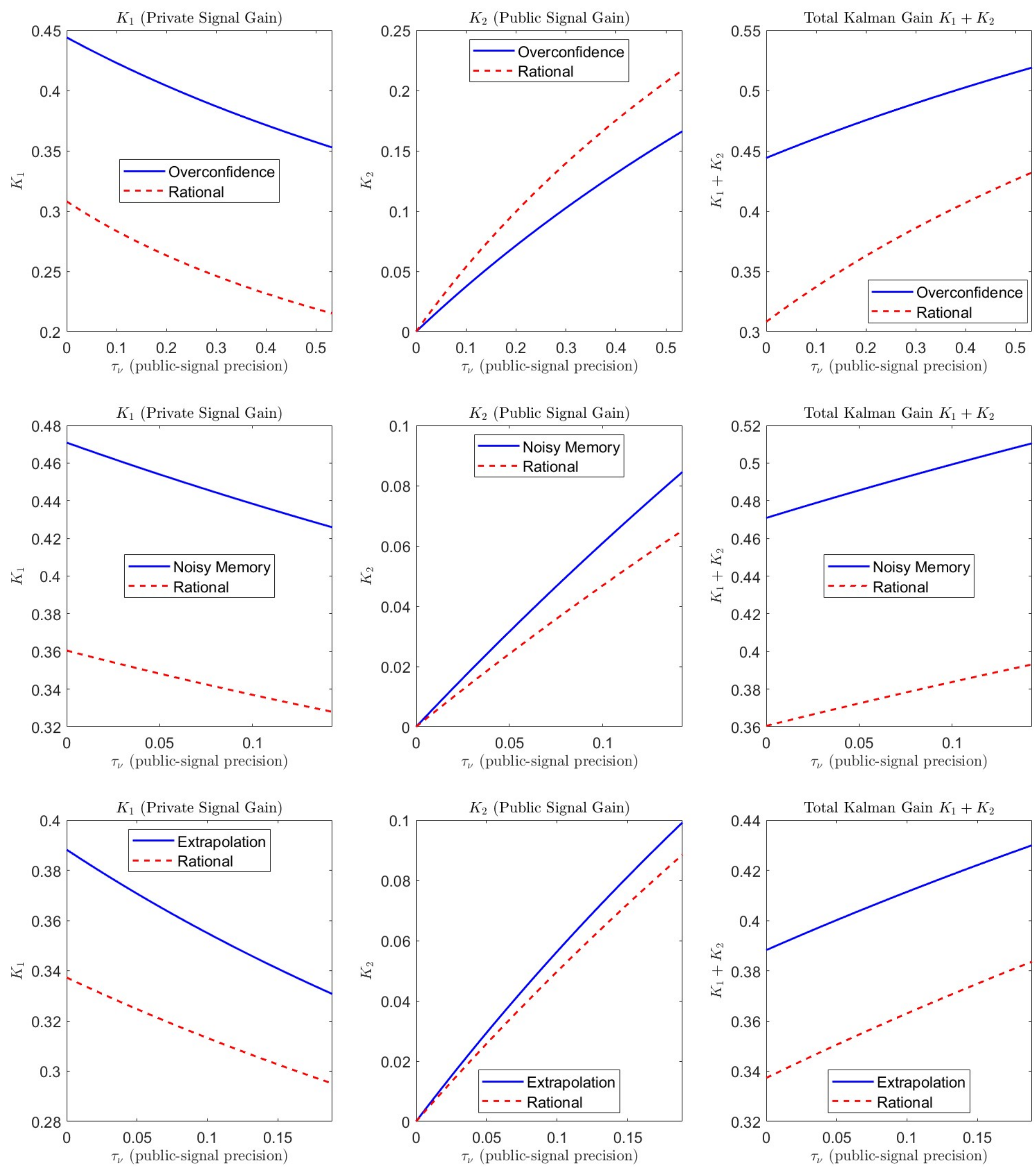
Panel A calibrates σ_ν^2 to match β_C (treated), leaving treated β_I non-targeted. Panel B calibrates σ_ν^2 to match β_I (treated), leaving treated β_C non-targeted. We exclude the Over-extrapolation model from the second exercise because, under our chosen DGP, there is no admissible parameter vector $(\sigma_\epsilon^2, \hat{\rho}, \sigma_\nu^2)$ that can simultaneously match $\beta_I\,(NT)$ and $\beta_I\,(T)$.

Figure D.1: β_C, β_I as a function of public signal precision τ_ν on under the three model specifications:



The figure overlays the three model specifications, plotting how increases in public-signal precision affect β_C and β_I on common axes.

Figure D.2: Kalman gain on Private and Public Signals Across Models.



The figure shows the Kalman gains on Private and Public Signals (K_1 and K_2 respectively) as a function of public signal precision