Overconfidence in Private Information Explains Biases in Professional Forecasts

Klaus Adam[*]
University of Mannheim & CEPR

Pei Kuang[†]
University of Birmingham

Shihan Xie[‡]
University of Illinois Urbana-Champaign

April 13, 2024

——— preliminary version ———

Abstract

We observe a rich set of public information signals available to participants in the Survey of Professional Forecasters (SPF) and decompose individual forecast revisions into those due to public information and a remainder due to residual information. We robustly find that SPF forecasters overreact to residual information for almost all forecast horizons and variables. Likewise, forecasts are overly anchored to prior beliefs for all variables and forecast horizons. We show that overconfidence in private information explains both of these features; it also explains why forecast errors correlate positively with past forecast revisions at the consensus level, but negatively at the individual level. Overconfidence in private information also delivers a new approach for improving upon consensus forecasts: it reduces mean-squared forecast errors for SPF variables by 35% compared to the approach in Coibion and Gorodnichenko (2015).

JEL classification: C53, D83, D84, E31
Keywords: Expectation, Information, Belief Updating

[*]Department of Economics, University of Mannheim, Germany. Email: adam@uni-mannheim.de
[†]Department of Economics, University of Birmingham, UK. Email: p.kuang@bham.ac.uk
[‡]Department of Economics, University of Illinois Urbana-Champaign, U.S.A. Email: shihanx@illinois.edu

1 Introduction

Expectations play a central role in dynamic economic decisions and full-information rational expectations (FIRE) have become the dominant assumption on expectation formation in macroeconomics. In a seminal paper, Lucas (1972) relaxed the FIRE assumption, studying expectation formation in a setup with incomplete information. Subsequently, macroeconomists continued studying deviations from FIRE in settings with private information (Woodford 2002; Mankiw et al. 2003; Sims 2003).

A key difficulty with testing the forecast implications of private and incomplete information models is that the information set available to forecasters can typically not be observed. This makes it difficult to study the efficiency properties of individual forecasts and provides challenges for using private information models in quantitative applications. To deal with this issue, Coibion and Gorodnichenko (2015) propose using past forecasts as measures of the information available to agents. Using this approach, they show that professional forecasts underreact to past information at the consensus level. Applying the same approach to individual forecasts, Bordalo et al. (2020) document that forecasts overreact to past forecast revisions.

While these findings point towards deviations from FIRE, they provide only indirect evidence about the economic mechanisms giving rise to these deviations. In particular, it remains unclear which sources of information agents may or may not use optimally. Understanding this requires knowledge about the information set available to forecasters at the time of forecasting. The present paper makes progress on this front.

Going back to the survey forms that get administered when collecting SPF forecasts, we find that SPF forecasters are provided with the most recent data release of the 14 variables they are requested to forecast in every forecasting round. With forecasters being aware that all forecasters obtain the same survey questionnaire, the latest data release represents public information to forecasters. And since we measure agents' prior expectations about the newly released variables from the previous forecasting round, we can construct a high-dimensional measure of public news every forecaster received from one survey round to the next. Due to heterogeneity in prior expectations, the public news component differs in the cross-section of forecasters at any given point in time.

We then use our forecaster-specific measure of public news to estimate how individual forecast revisions depend on (i) public news, (ii) prior beliefs, and (iii) a residual capturing everything that is not spanned by the prior and the high-dimensional public signal. We show that the residual component contains - at any given point in time - a common component that is useful for forecasting and an idiosyncratic forecaster-specific component that is not. This suggests that the residual component reflects noisy private information available to forecasters.[1]

In a second step, we regress individual ex-post forecast errors on (i) the forecast revisions explained by public news, (ii) the prior beliefs, and (iii) the residual component. Doing so, we robustly find that the residual component (iii) is negatively associated with future forecast errors. This holds true for virtually all forecast horizons and forecast variables in the survey and suggests that forecasters' expectations overreact to private information.

We also show that agents' expectations are overly anchored to prior expectations because prior expectations negatively predict future forecast errors. Again, this is true for virtually all forecast variables and forecast horizons in the survey.

In a final step, we show that public news (i) tends to be mostly positively associated with forecast errors, indicating that forecast revisions tend to underreact to public news. However, the empirical evidence for public information is more ambiguous, as forecasters also appear to sometimes overreact to public signals. This finding is broadly in line with evidence previously provided in Broer and Kohlhas (2022).

To understand these new facts, as well as the over/underreaction patterns previously documented in Coibion and Gorodnichenko (2015) and Bordalo et al. (2020), we construct a simple Bayesian updating model in which agents receive public and private information, but are overconfident about the information content of the private signal. We show that this very simple model can replicate all of the documented deviations from FIRE mentioned above, suggesting that overconfidence in private information could be the sole source of the observed deviations from FIRE.

The simple theoretical model suggests that one can improve upon consensus forecasts by correcting for forecasters' overreliance on private information. Using the implied approach,

[1]This also appears plausible in light of the fact that we observe a rather large number of public signals available to forecasters.

we show that we get a 35% mean-squared error reduction (on average across all variables and forecast horizons), compared to the consensus forecast corrections implied by the regressions in Coibion and Gorodnichenko (2015).

While we do not seek to explain why forecasters overly rely on private information, several existing theories provide potential explanations. This includes models with strategic diversification motives (e.g., Gemmi and Valchev 2023) and models with behavioral overconfidence (e.g., Angeletos et al. 2021; Broer and Kohlhas 2022).

In particular, Broer and Kohlhas (2022) document overreaction and underreaction to public information and Gemmi and Valchev (2023) study the response of forecast errors to public signals, proposing a model with strategic diversification to explain the observed expectations patterns. These papers assume that public information consists of past consensus forecasts, while the present paper treats the most recent data release as public information, in line with the information that is provided to forecasters on the SPF survey questionnaire.

Angeletos et al. (2021) provides interesting conditional evidence on forecasting behavior, including delayed overshooting patterns for expectations in response to economic shocks. The present paper is not concerned with conditional evidence instead provides unconditional evidence on deviations from FIRE. However, our finding that agent's expectations are overly anchored to past beliefs, implies - in the impact period - underreaction to economic disturbances, in line with the findings in Angeletos et al. (2021).

More broadly, the paper is related to a large body of literature that adopts different approaches to deviate from FIRE and model the formation of beliefs and expectations. Prominent examples include sticky information (Mankiw and Reis, 2002), noisy information (Woodford, 2002), rational inattention (Sims, 2003), diagnostic expectations (Bordalo et al., 2020; Bianchi et al., 2023), internal rationality (Adam and Marcet, 2011; Adam et al., 2017), overconfidence (Broer and Kohlhas, 2022; Angeletos et al., 2021), cognitive discounting (Gabaix, 2020), level-K thinking (García-Schmidt and Woodford, 2019; Farhi and Werning, 2019), and narrow thinking (Lian, 2021). Our survey evidence has implications for rational and behavioral models of expectation formation and we discuss these in detail later.

The remainder of the paper is organized as follows: Section 2 documents the evidence that we aim to explain, including a rich set of new empirical facts. Section 3 presents a simple model with noisy information that can replicate all these facts. In section 4, we document that our findings deliver a new approach for improving upon consensus forecasts. Section 5 concludes.

2 New evidence on the source of forecast errors

This section explains how we identify the public information available to forecasters in the U.S. Survey of Professional Forecasters (SPF) at the time of forecasting. Using the information on available public information, we then decompose macroeconomic forecast revisions of individual forecasters into those due to public information, private information, and prior expectations. We then show how individual ex-post forecast errors depend on these components.

2.1 SPF forecasts and outcome variables

We use data on forecasts from the Survey of Professional Forecasters (SPF), provided by the Federal Reserve Bank of Philadelphia. Every quarter, around 40 professional forecasters contribute to the SPF with forecasts for outcomes in the current quarter and the next four quarters. Individual forecasts are collected at the end of the second month of each quarter and cover macroeconomic and financial variables. Individual forecasters can be identified by forecaster IDs.

In our analysis, we consider the same variables and time period as studied in Bordalo et al. (2020). This includes nominal GDP (NGDP), real GDP (RGDP), GDP price deflator (PGDP), housing starts (Housing), and the unemployment rate (UNEMP), all of which are available from 1968 Q4 to 2016 Q4, the index for industrial production (INPROD), the consumer price index (CPI), real consumption (RCONSUM), real nonresidential investment (RNRESIN), real residential investment (RRESINV), federal government consumption (RGF), and state and local government consumption (RGSL), available from 1981 Q3 to 2016 Q4, the three-month treasury rate (TB3M), available from 1981 Q3 to 2016 Q4, and the ten-year treasury rate (TN10Y), available from 1992 Q1 to 2016 Q4.

We use forecasts over multiple horizons. We transform growing variables, such as GDP and CPI, into growth rates, studying in quarter t the growth rate from quarter $t-1$ to quarter $t+h$ for $h = 1,2,3,4$. For stationary variables, such as the unemployment rate or interest rates, we consider the variable in levels in quarter $t+h$. We winsorize outliers that are more than 5 interquartile ranges away from the median for each forecast horizon in a given quarter.

As outcome variables, we use the initial releases from the Federal Reserve Bank of Philadelphia's Real-Time Dataset for Macroeconomists. For example, for actual GDP growth from quarter $t-1$ to quarter $t+h$, we use the *initial* release of GDP_{t+h} in quarter $t+h+1$ divided by the most recent update of GDP_{t-1} in period $t+h$.

2.2 Existing evidence on SPF forecast errors

In important work, Coibion and Gorodnichenko (2015) show that ex-post forecast errors are positively associated with past forecast revisions at the consensus level. Specifically, they consider regressions of the form

$$\pi_{t+h} - \pi^c_{t+h|t} = \delta_h + \beta^c_h(\pi^c_{t+h|t} - \pi^c_{t+h|t-1}) + \epsilon_{t,h}, \tag{2.1}$$

where π_{t+h} denotes the outcome of variable π in period $t+h$ and $\pi^c_{t+h|t}$ the consensus forecast of variable π_{t+h} in period t, where consensus forecasts are simply the average of individual forecasters' predictions. The orange dots in Figure 1 report β^c_h for $h=1,2,3$ and show that future consensus forecast errors are positively predicted by past consensus forecast revisions. This holds true for almost all forecast variables and forecast horizons, in line with evidence provided in Coibion and Gorodnichenko (2015).

Bordalo et al. (2020) considered the same regression at the level of individual forecasters:

$$\pi_{t+h} - \pi^i_{t+h|t} = \delta^i_h + \beta^p_h(\pi^i_{t+h|t} - \pi^i_{t+h|t-1}) + \epsilon^i_{t,h}, \tag{2.2}$$

where $\pi^i_{t+h|t}$ denotes forecaster i's forecast of π_{t+h} as of time t. The blue dots in Figure 1 report the coefficient β^p_h for different forecast horizons ($h=1,2,3$). The coefficient β^p_h are often statistically significantly negative, with only the unemployment rate and the three-month treasury rate displaying significantly positive coefficients. This shows that individual forecasts tend to overreact to individual past forecast revisions.

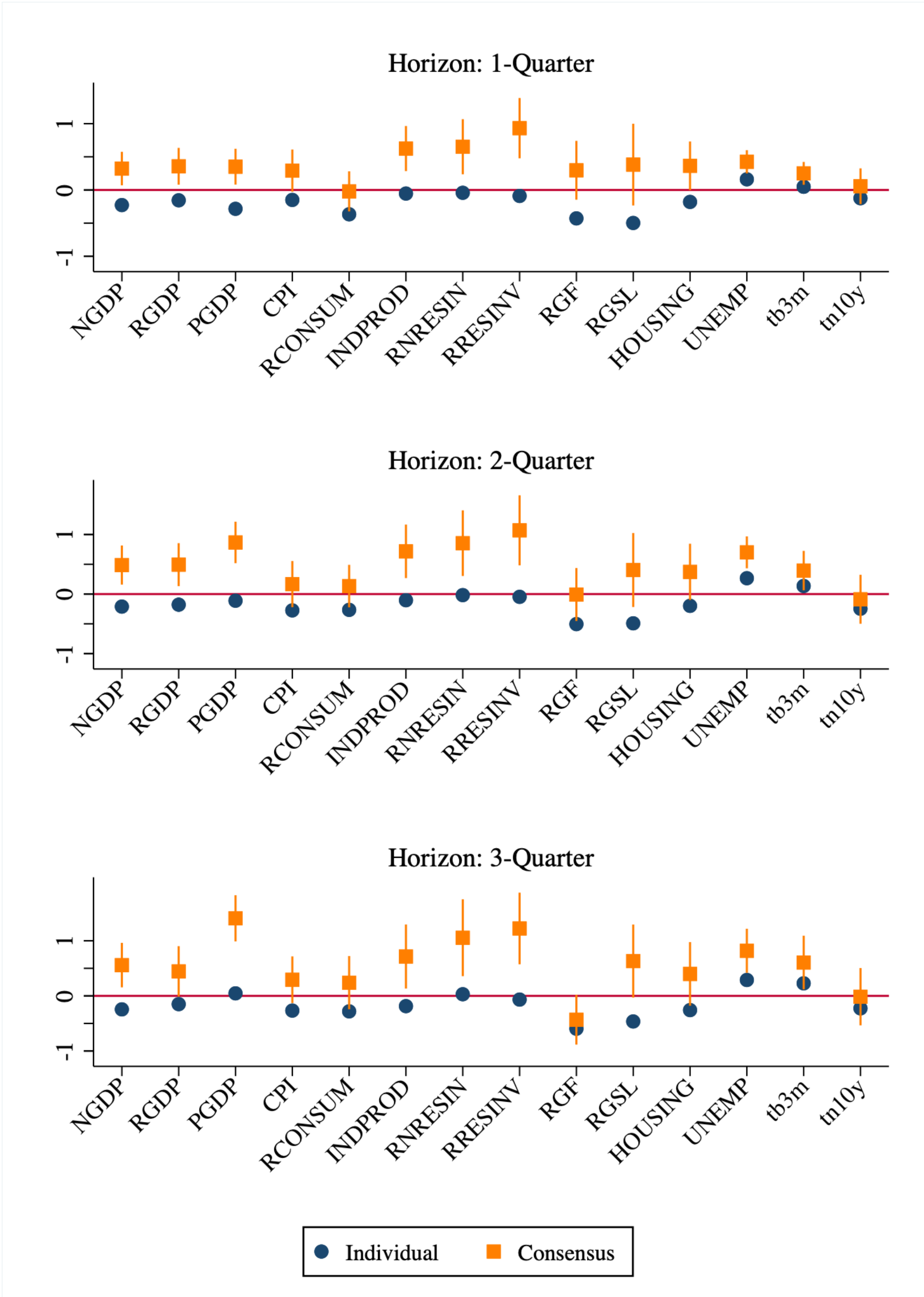


Figure 1: RESPONSES OF FORECAST ERRORS TO FORECAST REVISIONS AT THE CONSENSUS AND INDIVIDUAL LEVEL

Notes: This figure plots the coefficients of β_h^c (in orange) and β_h^p (in blue) from Eqn. (2.1) and (2.2). 95% confidence intervals based on clustered standard errors are reported.

2.3 Public information available to SPF forecasters

At the end of the first month in every quarter, the Bureau of Economic Analysis (BEA) releases its advance report of the national income and product accounts (NIPA) for the previous quarter. In the second month of the quarter, the SPF survey questionnaires are sent out to forecast participants. These questionnaires report - *in front of the response fields where survey respondents enter their forecasts* - the most recent data release from the BEA's advance report, and for non-NIPA data the latest release of other government statistical agencies.

Figure A.1 provides a sample questionnaire sent to SPF panelists: the column on the left in the table contains the most recent quarterly data release and to the right of these, the forecasts are entered. Given this, panelists can hardly avoid seeing the last data release when submitting their forecasts.

The SPF survey management team confirmed to us that they have been providing the most recent data release to panelists in every survey round since the 1990 Q2 survey, i.e., from the time they took over the administration of the surveys. From 1968:Q4 to 1990:Q2, the survey was conducted by the American Statistical Association (ASA) and the National Bureau of Economic Research (NBER). A few sample ASA-NBER survey forms are available on the SPF webpage. The survey forms state that "Recently reported figures are given on an attached sheet", which strongly suggests that forecasters have been provided with the most recent data release also during this earlier period.

Together with the survey form, forecasters also receive a historical data sheet from the SPF survey management team. Figure A.2 shows such a sample data sheet. For quarterly variables, the data sheet contains the realized values for the last four quarters and the annual value for the most recent year. For monthly variables, the data sheet contains their realized values for the last six months.

We note that it is common practice to supply professional forecasters with the latest data release when conducting surveys. For instance, this is also the case for the Livingston survey, the survey run by Consensus Economics, and the European Central Bank's Survey of Professional Forecasters. Appendix A provides a detailed discussion on the information sets of forecasters participating in these surveys. Importantly, however, the decomposition exercise we implement below can only be performed using the SPF forecast. The SPF is the only survey that includes in every round forecasts for four consecutive quarters. In contrast, other surveys only ask for forecasts over much longer horizons (usually over one year) or ask forecasters to forecast a fixed calendar year. As we explain later, the availability of past forecast for the current quarter is key for our analysis, as it allows observing individual-specific news associated with newly-released public information.

2.4 Decomposing forecast revisions and their effects on forecast errors

This section decomposes individual forecast revisions into revisions associated with private and public news. Specifically, we exploit the fact that we observe - from the previous forecasting round - forecasters' prior expectations about the latest data release that gets presented to them on the survey questionnaire. This allows the construction of an individual-specific news measure for each newly released variable. We then collect these news measures across variables into an individual-specific vector of public news. This is possible because we know that the latest data release is public information, as explained in the previous section.

Consider the second month of quarter t in which forecasts are collected and let $s_t \in R^{14}$ denote the vector of public information presented to the forecasters, which contains the latest data release for outcomes in the preceding quarter. Letting $s^i_{t|t-1}$ denote the forecast for these variables in the preceding quarter, the individual-specific public news is given by $s_t - s^i_{t|t-1}$. Since agents hold heterogeneous prior expectations, e.g., due to the availability of private information, the news revealed by the data release s_t will vary across forecasters at any given point in time.

Next, let π_{t+h} denote the vector of variables agents are asked to forecast for quarter $t+h$ and $\pi^i_{t+h|t-1}$ forecaster i's forecast of π_{t+h} as of quarter $t-1$. We can analyze how this forecast is revised from quarter $t-1$ to t, i.e., $\pi^i_{t+h|t} - \pi^i_{t+h|t-1}$.

Linear normal Bayesian updating implies that the forecast revision is going to be a linear function of public information, $s_t - s^i_{t|t-1}$, prior beliefs $\pi^i_{t+h|t-1}$, and private information, which remains unobserved as of now. In particular, we can regress (for $h = 1, 2, 3$) the observed forecast revision on observed public information and observed prior expectations:

$$\pi^i_{t+h|t} - \pi^i_{t+h|t-1} = \delta^i_h + \gamma_h (s_t - s^i_{t|t-1}) + \eta_h \pi^i_{t+h|t-1} + \epsilon^i_{t,h}, \tag{2.3}$$

where δ_h^i is an individual-horizon fixed effect and the coefficient matrix $\gamma_h \in R^{14\times 14}$ captures how forecasters respond to public news. The coefficient vector η_h captures the speed at which past information becomes obsolete. When agents follow Bayesian updating, we have $\eta_h \in [-1, 0]$, with $\eta_h = 0$ indicating no decay of old information (π_t follows a random walk) and $\eta_h = -1$ indicating full decay (π_t is an i.i.d. process).

Equation (2.3) allows decomposing forecast revisions into those due to (i) the vector of public information, (ii) prior information becoming less relevant and (iii) the residual component $\epsilon_{t,h}^i$.

If the public information signal s_t exhausts the set of public information, then the residual vector $\epsilon_{t,h}^i$ in equation (2.3) captures forecasts revisions that are due to forecasters' private information. Given that we have a high-dimensional public information signal, the notion that $\epsilon_{t,h}^i$ in fact reflects private information appears warranted, especially in light of the fact that the dynamics of macroeconomic variables can typically be described as being driven only by a small number of common factors, see Stock and Watson (2016). Below, we provide further evidence corroborating the view that $\epsilon_{t,h}^i$ represents private information.

Given our decomposition, we can define two components driving forecast revision: (i) the one generated by the public signal and prior information, and (ii) the one generated by private information, i.e., the regression residual:

$$\text{Predicted}_{i,t}^h \equiv \hat{\gamma}_h(s_t - s_{t|t-1}^i) + \hat{\eta}_h \pi_{t+h|t-1}^i, \tag{2.4}$$

$$\text{Residual}_{i,t}^h \equiv \hat{\epsilon}_{t,h}^i. \tag{2.5}$$

We are now in a position to analyze how these components affect individual forecast errors by regressing forecast errors on these two components:

$$\pi_{t+h} - \pi_{t+h|t}^i = \alpha_i^h + \beta_1^h \times \text{Predicted}_{i,t}^h + \beta_2^h \times \text{Residual}_{i,t}^h + \nu_{t,h}^i. \tag{2.6}$$

Figure 2 reports the estimated coefficients of β_1^h (in green) and β_2^h (in orange). For almost all variables and forecasting horizons, the results robustly show that macroeconomic expectations underreact to forecast revisions induced by the prior and the public news ($\beta_1^h > 0$) but overreact to the residual component, i.e., the private news component ($\beta_1^h < 0$). The following results summarize our empirical findings:

Fact 1: At the individual level, forecasters' expectations underreact to forecast revisions induced by public news and prior beliefs ($\beta_1^h > 0$).

Fact 2: At the individual level, forecasters' expectations overreact to the residual component of forecast revisions ($\beta_2^h < 0$).

We now explore further the causes giving rise to Fact 1, and thereafter consider Fact 2. To better understand what generates Fact 1, we decompose the predicted component of forecast revisions constructed above into its two components, namely the one explained by public news and the one explained by prior expectations.

We can then regress individual ex-post forecast errors on (i) the forecast revisions explained by public news, (ii) the prior beliefs, and (iii) our measure of private information from the regression (2.3). To do so, we define a new variable that summarizes the revisions due to public information

$$\text{Public}_{i,t}^h \equiv \hat{\gamma}_h (s_t - s_{t|t-1}^i),$$

which uses the estimated coefficient $\hat{\gamma}_h$ from equation (2.3). We then consider the forecast-error regression:

$$\pi_{t+h} - \pi_{t+h|t}^i = \delta_h^i + \alpha_1^h \times \text{Public}_{i,t}^h + \alpha_2^h \pi_{t+h|t-1}^i + \beta_2^h \times \text{Residual}_{i,t}^h + \nu_{t,h}^i. \tag{2.7}$$

Figure 3 plots the estimated coefficients of α_1^h (in blue) and α_2^h (in maroon) from Eqn. (2.7).[2] The results indicate a negative coefficient on the prior beliefs ($\alpha_2^h < 0$) for all forecast variables and all forecast horizons. This shows that forecasters' expectations are too strongly anchored to prior expectations. Also, forecast errors covary mostly positively with public news ($\alpha_1^h > 0$), which suggests underreaction of expectations to public news. This feature is, however, less consistent across variables and forecast horizons. Overall, however, both sub-components tend to contribute to the underreaction result to $\text{Predicted}_{i,t}^h$ in equation (2.6). We summarize our empirical findings as follows:

Fact 3: At the individual level, forecasters' expectations mostly underreact to public news

[2]By construction, the regressor $\text{Residual}_{i,t}^h$ is orthogonal to the news component $(s_t - s_{t|t-1}^i)$ and the prior $(\pi_{t+h|t-1}^i)$, so that the estimate of β_2^h in equation (2.7) will be identical to the one in equation (2.6).

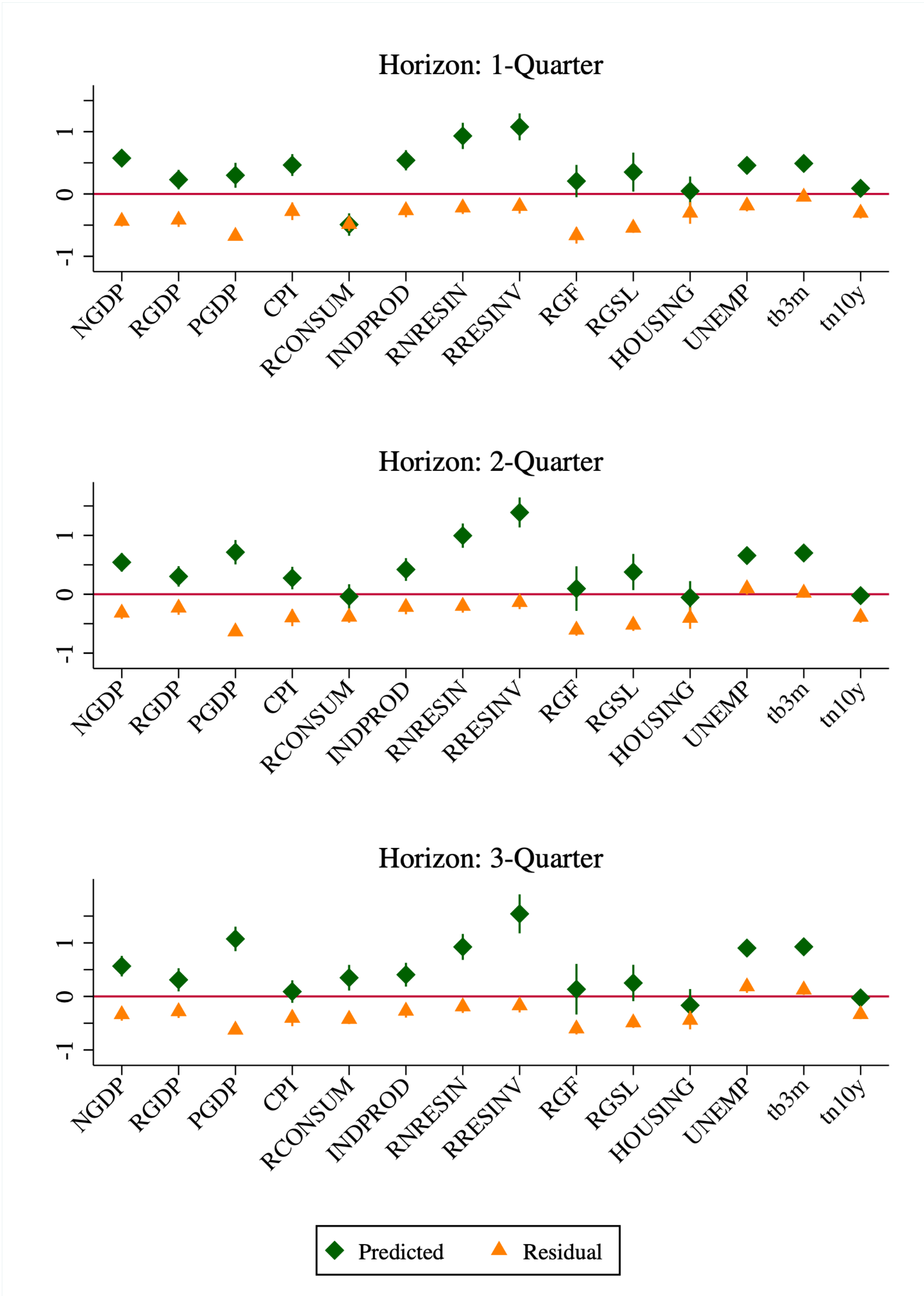


Figure 2: RESPONSES OF FORECAST ERRORS TO PREDICTED AND RESIDUAL COMPONENTS OF FORECAST REVISION

Notes: This figure plots the coefficients of β_1^h (in green) on the predicted component of forecast revisions and β_2^h on the residual component (in orange) from Eqn. (2.6). 95% confidence intervals based on clustered standard errors are reported.

($\alpha_1^h > 0$), although there are exceptions.

Fact 4: At the individual level, forecasters' expectations are overly anchored to prior expectations ($\alpha_2^h < 0$).

In a final step, we seek to better understand Fact 2 mentioned above. In particular, we seek to corroborate our interpretation of the estimated residual $\hat{\epsilon}^i_{t,h}$ in equation (2.3) as representing (noisy) private information. To this end, we decompose private information (at a given point in time) into a common and an idiosyncratic component

$$\text{Common}_{t,h} \equiv \frac{1}{N_t}\sum_i \hat{\epsilon}^i_{t,h}, \tag{2.8}$$

$$\text{Idiosync}^h_{i,t} \equiv \hat{\epsilon}^i_{t,h} - \text{Common}^h_t, \tag{2.9}$$

where I_t denotes the number of forecasters in quarter t. We can then consider another forecast error regression of the form:

$$\begin{aligned} \pi_{t+h} - \pi^i_{t+h|t} =& \delta^i_h + \alpha^h_1 \times \text{Public}^h_{i,t} + \alpha^h_2 \pi^i_{t+h|t-1} \\ &+ \theta^h_1 \times \text{Common}^h_t + \theta^h_2 \times \text{Idiosync}^h_{i,t} + \nu^i_{t,h}. \end{aligned} \tag{2.10}$$

Figure 4 plots the estimated coefficients of θ^h_1 (in green) and θ^h_2 (in orange) from Eqn. (2.10). It shows that the idiosyncratic component of the residual has a negative coefficient ($\theta^h_2 < 0$) for all variables and all horizons. This suggests that forecasters overreact to the noise component in private information. The coefficient on the common component of private information is generally positive ($\theta^h_1 > 0$), suggesting mostly underreaction of expectations to the common component. This is again consistent with the fact that the residuals $\epsilon^i_{t,h}$ in equation (2.3) capture updating due to noisy private information.

In Appendix B.1 we repeat the analysis carried out in the present section using as public signal only the information contained in the last release of the variable that gets forecasted. This leads to the same findings as the ones derived above.

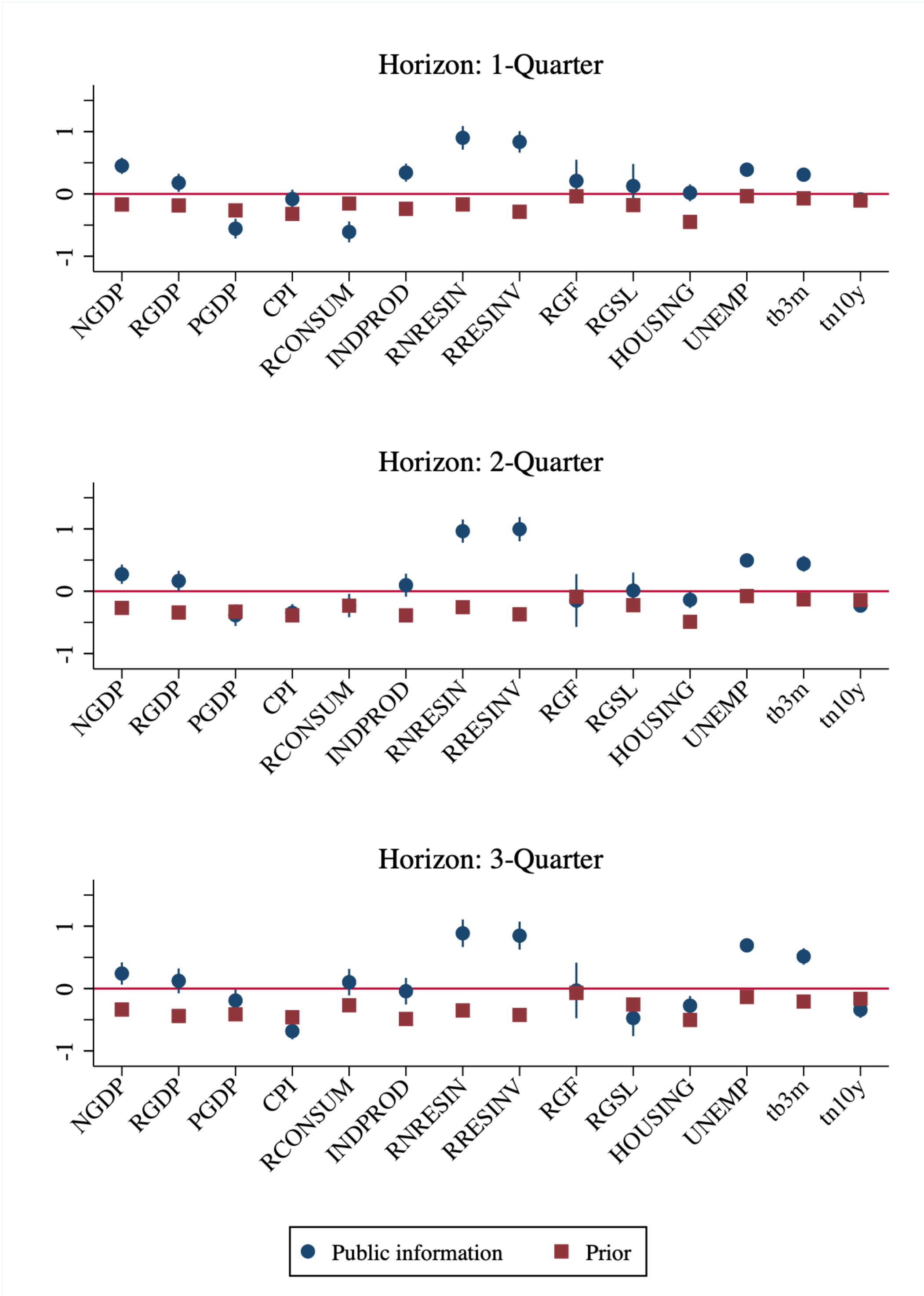


Figure 3: RESPONSES OF FORECAST ERRORS TO PUBLIC NEWS AND PRIOR EXPECTATIONS

Notes: This figure plots the estimated coefficients of α_1^h (in blue) and α_2^h (in maroon) from Eqn. (2.7). 95% confidence intervals based on clustered standard errors are reported.

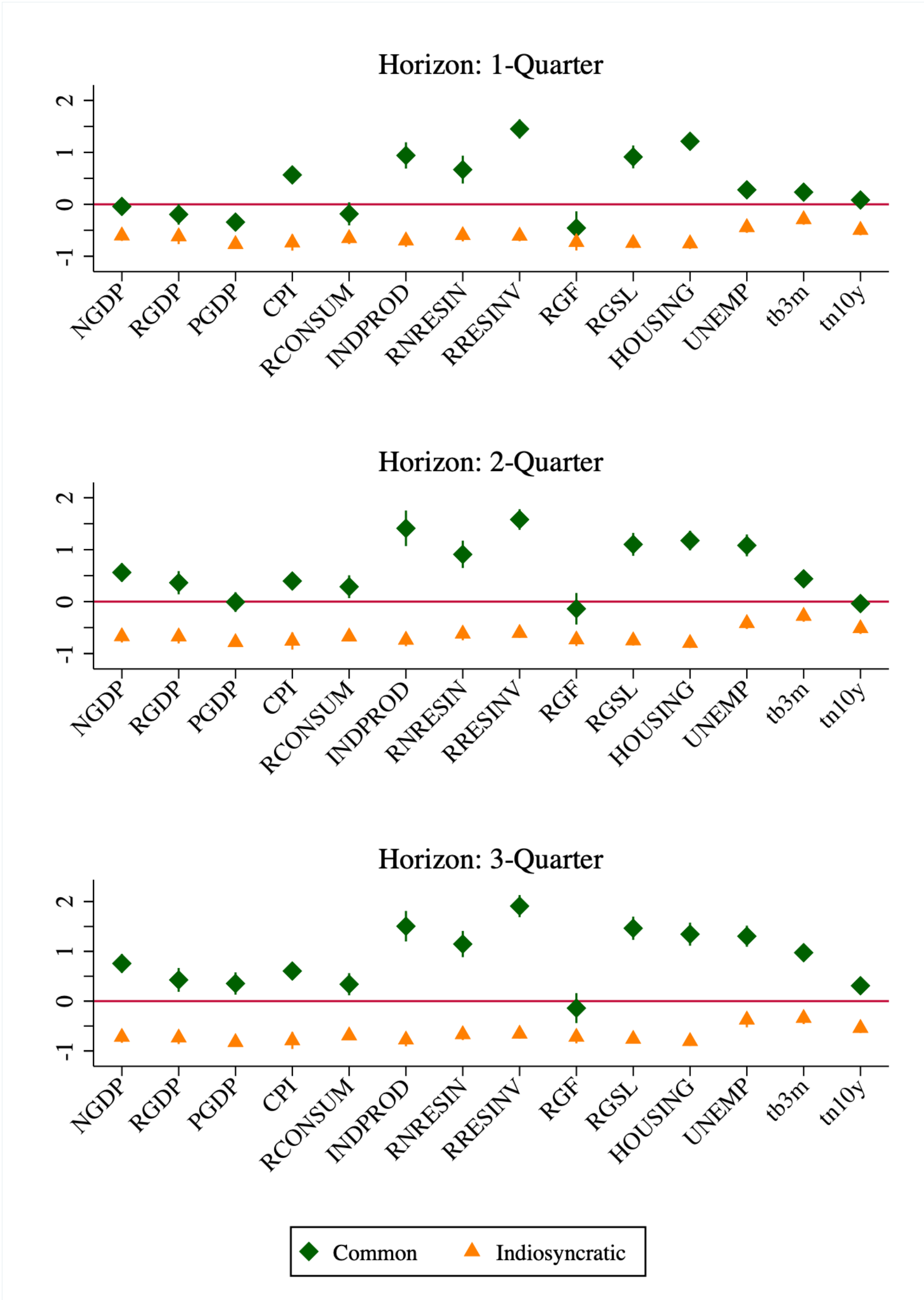


Figure 4: RESPONSES OF FORECAST ERRORS TO COMMON AND IDIOSYNCRATIC COMPONENTS OF PRIVATE INFORMATION

Notes: This figure plots the estimated coefficients of θ_1^h (in green) and θ_2^h (in orange) from Eqn. (2.10). 95% confidence intervals based on clustered standard errors are reported.

3 A simple noisy information model

This section presents a simple forecasting model that can replicate the newly documented Facts 1 to 4 from the previous section and the evidence from Coibion and Gorodnichenko (2015) and Bordalo et al. (2020) summarized in section 2.2.

Specifically, we consider a setup in which forecasters receive public and private information about an underlying state process that drives the realizations of observable variables. In line with the empirical analysis in the previous section, public information consists of the most recent data release from the previous quarter, while private signals provide noisy information about the underlying state in the current quarter.

The next subsection introduces the setup. Section 3.2 shows that the model misses almost all facts when assuming rational expectations. In contrast, section 3.3 shows that all facts are replicated when agents are overconfident in the private signal.

3.1 The setup

Consider a setting with a measure one of forecasters, indexed by $i \in [0,1]$. In period t, forecasters seek to forecast some variable s_{t+h} for $h \geq 1$, which depends on some underlying factor π that evolves according to

$$\pi_t = \rho \pi_{t-1} + u_t, \tag{3.1}$$

where $\rho \in (0,1)$ and $u_t \sim_{iid} N(0, \sigma_u^2)$. In period t, agents receive information s_t about the lagged outcome of the variable of interest, which is given by

$$s_t = \lambda \pi_{t-1} + \nu_t, \tag{3.2}$$

where $\lambda > 0$ denotes the factor loading and $\nu_t \sim_{iid} N(0, \sigma_\nu^2)$ variable specific noise. In the special case with $\sigma_\nu^2 = 0$, agents directly observe the lagged value of the factor and the factor is identical to the variable of interest. Without loss of generality, we set $\lambda = 1$ from now on.

In each period t, forecaster i receives an idiosyncratic private signal x_{it} about the value of the current factor

$$x_{it} = \pi_t + \epsilon^x_{it}, \tag{3.3}$$

where $\epsilon^x_{it} \sim_{iid} N(0, \sigma^2_\epsilon)$ is idiosyncratic observations noise.

In period t, forecaster i formulates subjective forecasts $\mathbb{E}^{\mathscr{P}}[s_{t+h}|\Omega^i_t]$ for $h > 0$, where $\mathscr{P}$ is a subjective probability measure, described further below, and the information is given by $\Omega^i_t = \{s_\tau, x_{i\tau}\}^t_{\tau=0}$. The subjective probability measure $\mathscr{P}$ allows forecasters to entertain subjective point beliefs about the value of the variances $(\sigma^2_u, \sigma^2_v, \sigma^2_\epsilon)$, denoted by $(\hat{\sigma}^2_u, \hat{\sigma}^2_v, \hat{\sigma}^2_\epsilon)$. In the initial period t=-1, forecasters hold normal prior beliefs about π_{-1}, with prior uncertainty given by the steady-state value implied by the subjective Kalman filter. Since we have

$$\mathbb{E}^{\mathscr{P}}[s_{t+h}|\Omega^i_t] = \mathbb{E}^{\mathscr{P}}[\pi_{t+h-1}|\Omega^i_t], \tag{3.4}$$

forecasting future realizations for s is the same as forecasting the underlying factor (one period lagged).

Given the subjective beliefs, forecaster i finds it optimal to use a prediction rule of the form

$$\pi^i_{t|t} \equiv \mathbb{E}^{\mathscr{P}}\left[\pi_t \mid \Omega^i_t\right] = (1 - \kappa_x - \kappa_y)\pi^i_{t|t-1} + \kappa_x x_{it} + \kappa_y \rho s_t, \tag{3.5}$$

where $\pi^i_{t|t-1} \equiv \mathbb{E}^{\mathscr{P}}\left[\pi_t \mid \Omega^i_{t-1}\right]$ is agent i's prior belief about π_t. The Kalman filter weights κ_x and κ_y are implied by the subjective point beliefs for $(\hat{\sigma}^2_u, \hat{\sigma}^2_v, \hat{\sigma}^2_\epsilon)$ and given below. Equivalently, we have the updating equation

$$\pi^i_{t|t} = \kappa_x x_{it} + (1 - \kappa_x)\rho \left[\omega s_t + (1-\omega)\pi^i_{t-1|t-1}\right],$$

where $\kappa_y = (1 - \kappa_x)\omega$ and with Kalman filter weights given by

$$\omega = \frac{(\hat{\sigma}^2_v)^{-1}}{(\hat{\sigma}^2_\tau)^{-1} + (\hat{\sigma}^2_v)^{-1}}, \tag{3.6}$$

$$\kappa_x = \frac{(\hat{\sigma}^2_\epsilon)^{-1}}{(\hat{\sigma}^2_\epsilon)^{-1} + \left[\rho^2\left(\omega^2\hat{\sigma}^2_v + (1-\omega)^2\hat{\sigma}^2_\tau\right) + \hat{\sigma}^2_u\right]^{-1}}. \tag{3.7}$$

where $\hat{\sigma}_\tau^2$ is the (stationary subjective) uncertainty about the prior mean $\pi^i_{t-1|t-1}$. For the case with rational beliefs $((\hat{\sigma}_u^2, \hat{\sigma}_\nu^2, \hat{\sigma}_\epsilon^2) = (\sigma_u^2, \sigma_\nu^2, \sigma_\epsilon^2))$, the previous equations deliver the rational Kalman filter weights that we denote by ω^* and κ_x^*.

3.2 Model predictions under rational expectations

We first explore the predictions of the model under rational expectations. The following proposition shows that the model then fails to replicate Facts 1 to 4:

Proposition 1. *Under rational expectations:*

1. *Forecasters' expectations do not over- or under-react to news-related forecast revision ($\beta_1^h = 0$), contrary to Fact 1.*
2. *Forecasters' expectations do not over- or under-react to the residual component of forecast revision ($\beta_2^h = 0$), contrary to Fact 2.*
3. *Forecasters' expectations do not over- or under-react to public news ($\alpha_1^h = 0$), contrary to Fact 3.*
4. *Forecasters' expectations are not over- or under-anchored to prior expectations ($\alpha_2^h = 0$), contrary to Fact 4.*
5. *Forecasters' expectations do not over- or under-react to forecast revisions ($\beta_h^p = 0$), contrary to the Fact in Figure 1.*
6. *Consensus forecasts underreact to consensus forecast revisions ($\beta_h^c > 0$), consistent with the Fact in Figure 1.*

The proof is in Appendix D.1. Not surprisingly, under rational expectations, forecast errors cannot be explained by information available to agents at the time of forecasting. For the same reason, the model fails to replicate the evidence of Bordalo et al. (2020) that $\beta_h^p < 0$ in regression (2.2), see figure 1. Instead it implies that $\beta_h^p = 0$, see Appendix C.2 for a proof.

There is one dimension along which the rational expectations model performs well. Since forecasters know that private information is contaminated by noise, they adjust beliefs only gradually to private information. This leads to predictability of mean forecast errors by ex-ante mean forecast revisions and hence a positive regression coefficient $\beta_h^c > 0$ in the consensus CG regression (2.1), see Appendix C.3 for a proof.

3.3 Over-confidence in private information

We now consider the case where agents are overly optimistic about the accuracy of their private information. In particular, individuals perceive the standard deviation of their private signal as $\widehat{\sigma}_\epsilon^x = \tau \sigma_\epsilon^x$ with $\tau < 1$. In this setting, agents will update their beliefs using the following weights:

$$\widehat{\omega} = \frac{(\sigma_\nu^2)^{-1}}{(\widehat{\sigma}_\tau^2)^{-1} + (\sigma_\nu^2)^{-1}}, \tag{3.8}$$

$$\widehat{\kappa}_x = \frac{(\widehat{\sigma}_\epsilon^2)^{-1}}{(\widehat{\sigma}_\epsilon^2)^{-1} + \left[\rho^2\left(\widehat{\omega}^2\sigma_\nu^2 + (1-\widehat{\omega})^2\widehat{\sigma}_\tau^2\right) + \sigma_u^2\right]^{-1}}, \tag{3.9}$$

where agents' prior uncertainty is

$$\widehat{\sigma}_\tau^2 = \frac{\widehat{\kappa}_x^2\widehat{\sigma}_\epsilon^2 + (1-\widehat{\kappa}_x)^2\sigma_u^2 + \rho^2(1-\widehat{\kappa}_x)^2\widehat{\omega}^2\sigma_\nu^2}{1-\rho^2(1-\widehat{\kappa}_x)^2(1-\widehat{\omega})^2}. \tag{3.10}$$

We show that this *overconfidence* about the private signal simultaneously accounts for Facts 1 to 4. Intuitively, when agents are overly optimistic about their private information $\tau < 1$, they overreact to private signals ($\widehat{\kappa}_x > \kappa_x^*$). Since agents perceive a higher signal quality, prior uncertainty decreases compared to the case with rational expectations:

$$\widehat{\sigma}_\tau^2 < \sigma_{\tau,RE}^2.$$

As a result, agents overly anchor beliefs to prior information ($\widehat{\omega} < \omega^*$) as they perceive the public signal to be less useful. The following proposition summarizes our results:

Proposition 2. *When agents are overconfident about their private information:*

1. *Forecasters' expectations underreact to news-related forecast revision ($\beta_1^h > 0$), consistent with Fact 1.*
2. *Forecasters' expectations overreact to the residual component of forecast revision ($\beta_2^h < 0$), consistent with Fact 2.*

3. *Forecasters' expectations underreact to public news ($\alpha_1^h > 0$), consistent with Fact 3.*
4. *Forecasters' expectations are overly anchored to prior expectations ($\alpha_2^h < 0$), consistent with Fact 4.*
5. *Forecasters' expectations overreact to forecast revisions ($\beta_h^p < 0$), consistent with the Fact in Figure 1.*
6. *Consensus forecasts underreact to consensus forecast revisions ($\beta_h^c > 0$), consistent with the Fact in Figure 1.*

The proof is in Appendix D.2. The proof shows that overconfidence in private information allows also replicating a negative individual-level CG coefficient. And appendix Appendix C.3 shows that overconfidence in private information implies positive consensus-level CG coefficient, provided over-confidence is not excessively strong (i.e., $1 > \tau > N^{-1}$, where N is the number of forecasters).

3.4 Further tests of the overconfidence model

The overconfidence model in the previous section replicates the empirical evidence by assuming that the residuals in the empirical forecast revision equation (3.5) are due to private information. This interpretation of residual information gives rise to further testable predictions. This section derives these predictions and shows that they are supported by the data.

Consider equation (3.5) which specifies how - according to the model - forecasts react to forecasters' private information x_{it}. We can decompose this reaction into a component that is common across forecasters, $\kappa_x \frac{1}{N_t} \sum_i x_{it}$, where N_t denotes the numbers of forecasters, and an idiosyncratic component.[3] When N_t is sufficiently large, then the common component represents very precise information about the variable that is to be forecasted, see equation (3.3), while the idiosyncratic component reflects observation noise that is detrimental to forecasting performance. This implies that the reaction to the common component should increase agents' forecasting accuracy, while the reaction to the idiosyncratic components should reduce forecast accuracy. This implication can be tested in the data.

[3] Note that forecasters cannot perform this decomposition at the time of forecasting because they do not observe other forecasters' private information.

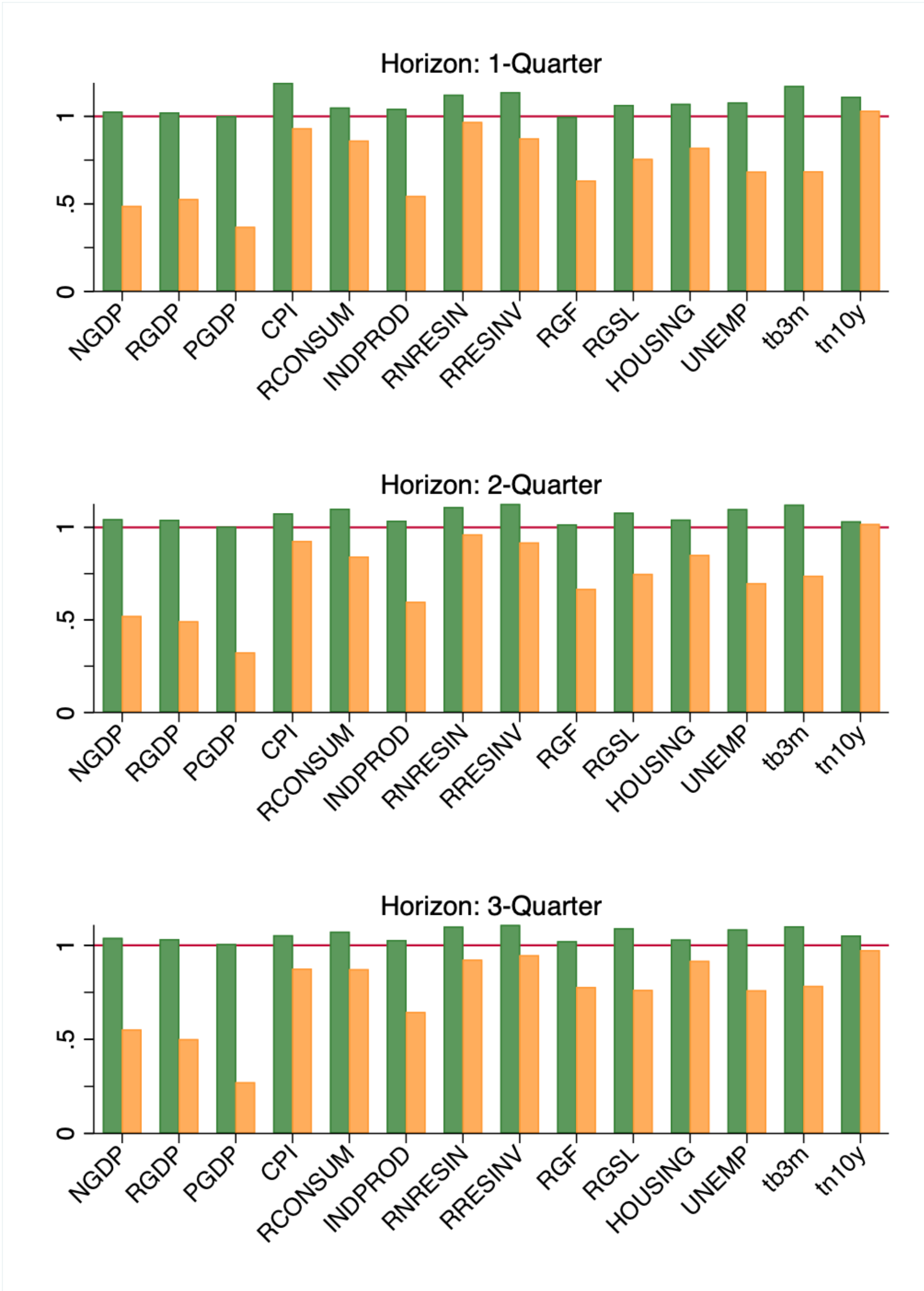


Figure 5: Individual Forecast Errors: Common vs. Idiosyncratic Components of Residual Information

Notes: This figure compares the individual forecast errors implied by equation (3.5) to those implied when replacing the residuals $\epsilon^h_{t,h}$ by the common component across forecasters (orange bars) or the idiosyncratic component (green bars). All forecast errors are expressed relative to those implied by equation (3.5), which uses both the common and idiosyncratic components.

Specifically, consider the common and idiosyncratic components (2.8)-(2.9) of the residual $\epsilon^i_{t,h}$ in equation (2.3). If residual information represents private information then individual forecast accuracy should increase, if we replace $\epsilon^i_{t,h}$ by the common component in equation (2.3). It should decrease when we replace it with the idiosyncratic component.

Figure 5 computes the resulting mean squared forecast errors (averaged across all forecasters) for each variable and forecast horizon, relative to the forecast errors implied by equation (2.3), which uses both the idiosyncratic and common component. The figure shows that using the common component instead of $\epsilon^i_{t,h}$ substantially reduces forecast errors. This holds true for virtually all variables and forecast horizons. Conversely, using the idiosyncratic components increases mean squared errors. These findings are in line with the predictions of the overconfidence model and support the notion that residual information captures private information.

4 Improving over consensus forecasts: a new approach

This section shows that the forecasting biases associated with overconfidence in private information imply that one can improve upon the forecasting performance of consensus forecasts, i.e., the average prediction across forecasters. Equation (2.1) from Coibion and Gorodnichenko (2015) already suggest this to be the case, as it implies one can improve upon consensus forecast using the forecasting equation:

$$\mathbb{E}_t \pi_{t+h} = \delta_h + \pi^c_{t+h|t} + \beta^c_h(\pi^c_{t+h|t} - \pi^c_{t+h|t-1}). \tag{4.1}$$

However, this forecasting correction neither exploits private information nor the structure of forecaster-specific public news, unlike the result we derive below.

In particular, to obtain consensus forecasts, we simply average equation (2.10) across forecasters i in period t and take condition expectations as of time t. This delivers:

$$\begin{aligned}\mathbb{E}_t \pi_{t+h} =& \frac{1}{N_t}\sum_i \delta^i_h + \pi^c_{t+h|t} + \frac{\alpha^h_1}{N_t}\sum_i \text{Public}_{i,t,h} \\ &+ \alpha^h_2 \pi^c_{t+h|t-1} + \theta^h_1 \times \text{Common}_{t,h}.\end{aligned} \tag{4.2}$$

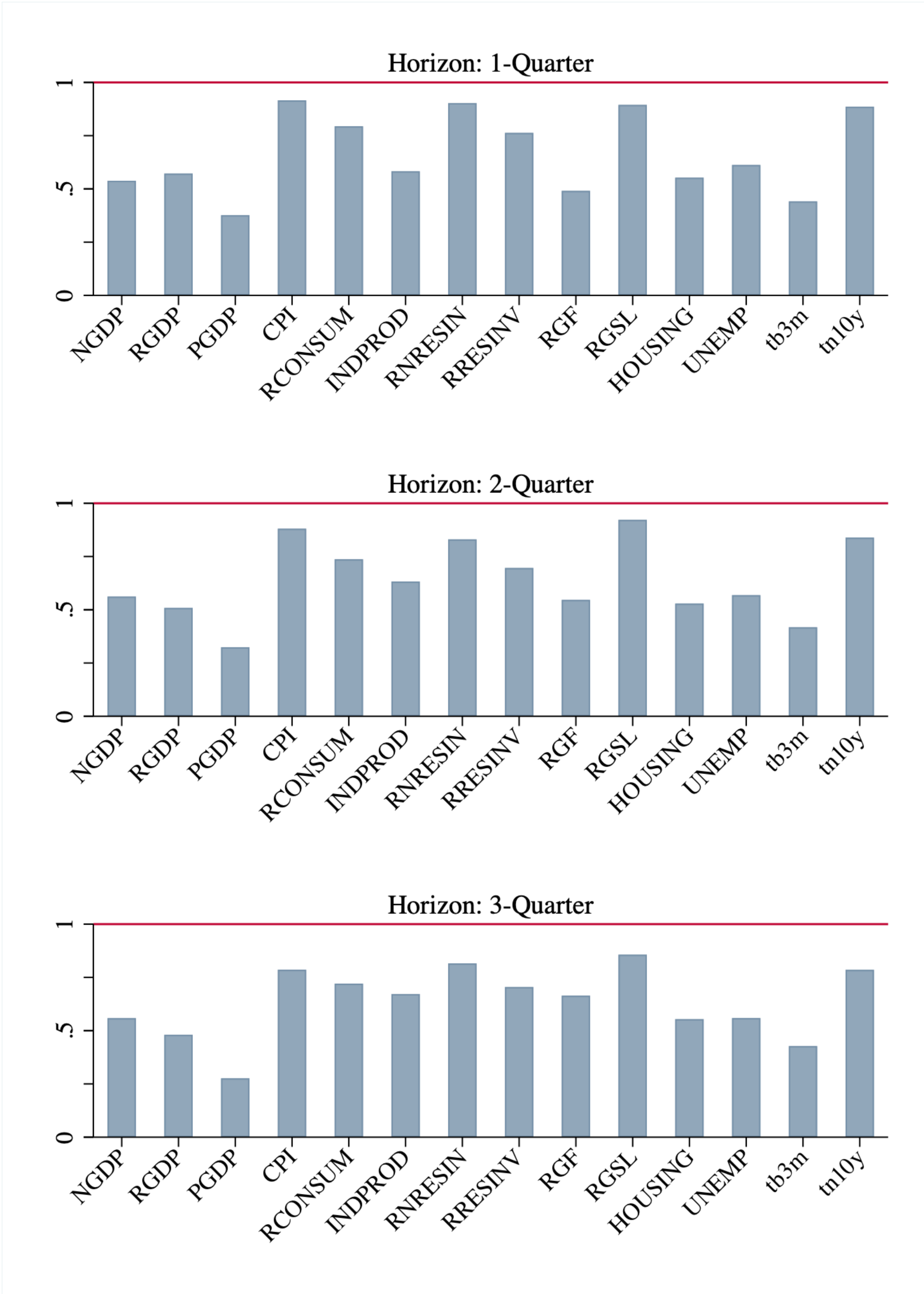


Figure 6: IMPROVING OVER CONSENSUS FORECASTS: MSE RELATIVE TO CG REGRESSION

Notes: This figure plots the mean-squared forecast error of (4.2) relative to the CG regression (4.1).

where the coefficients $(\delta_h^i, \alpha_1^h, \alpha_2^h, \theta_1^h)$ are determined by regression equation (2.10) and N_t denotes the number of forecasters in period t.

We now compute the mean-squared forecast error (MSE) implied by equation (4.2) and divide it by the MSE implied by the Coibion-Gorodnichenko regression (4.1), doing so for all 14 variables in the SPF and for all considered forecast horizons. Figure 6 plots the resulting MSE ratios. Our decomposition consistently and significantly improves upon Coibion-Gorodnichenko regression and often significantly so. On average across all variables and forecast horizons, relative MSE is about 0.65, i.e., 35% lower. For some variables, such as the GDP deflator or the 3-months treasury bill rate, improvements are even larger than 50%. Interestingly, a very similar improvement in forecast errors can be obtained when adding the predictors $\frac{\alpha_1^h}{N_t} \sum_i \text{Public}_{i,t,h}, \alpha_2^h \pi_{t+h|t-1}^c$ to equation (4.1) and then estimating (δ_h, β_h^c).[4] The improved forecast performance of equation (4.2) is thus to only to minor extent due to the fact that equations (4.1) and (4.2) impose different coefficient restrictions on consensus forecasts. Instead, the inclusion of the average public news and of the common component of private news drive the large reduction in MSE.

5 Conclusion

Observing forecasters' public information sets allows documenting a number of new facts about the behavior of forecasts in the Survey of Professional Forecasters. A simple model in which agents overweight a noisy private information signal when updating beliefs delivers all the new facts, but also previously established facts on how forecast errors relate to past forecast revisions at the consensus and individual levels. The results we document have important implications for the construction of empirically plausible private information models. They also raise the need to understand better the source of forecasters' over-reliance on private information.

[4] See figure B.12 in appendix Appendix B.2.

References

Adam, Klaus, Albert Marcet, and Johannes Beutel, "Stock Price Booms and Expected Capital Gains," *Amercian Economic Review*, 2017, *Vol. 107 No. 8*, 2352–2408.

_ **and** _ , "Internal Rationality, Imperfect Market Knowledge and Asset Prices," *Journal of Economic Theory*, 2011, *146*, 1224–1252.

Angeletos, George-Marios, Zhen Huo, and Karthik A. Sastry, "Imperfect Macroeconomic Expectations: Evidence and Theory," *NBER Macroeconomics Annual*, 2021, *35*, 1–86.

Bianchi, Francesco, Cosmin Ilut, and Hikaru Saijo, "Diagnostic Business Cycles," *The Review of Economic Studies*, 02 2023, *91* (1), 129–162.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer, "Overreaction in Macroeconomic Expectations," *American Economic Review*, September 2020, *110* (9), 2748–82.

Broer, Tobias and Alexandre N. Kohlhas, "Forecaster (Mis-)Behavior," *Review of Economics and Statistics, forthcoming*, 2022.

Carlson, John A., "A Study of Price Forecasts," in "Annals of Economic and Social Measurement, Volume 6, number 1" NBER Chapters, National Bureau of Economic Research, Inc, December 1977, pp. 27–56.

Coibion, Olivier and Yuriy Gorodnichenko, "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts," *American Economic Review*, August 2015, *105* (8), 2644–78.

Farhi, Emmanuel and Iván Werning, "Monetary Policy, Bounded Rationality, and Incomplete Markets," *American Economic Review*, November 2019, *109* (11), 3887–3928.

Gabaix, Xavier, "A Behavioral New Keynesian Model," *American Economic Review*, August 2020, *110* (8), 2271–2327.

García-Schmidt, Mariana and Michael Woodford, "Are Low Interest Rates Deflationary? A Paradox of Perfect-Foresight Analysis," *The American Economic Review*, 2019, *109* (1), 86–120.

Gemmi, Luca and Rosen Valchev, "Biased Surveys," Working Paper 31607, National Bureau of Economic Research August 2023.

Lian, Chen, "A Theory of Narrow Thinking," *Review of Economic Studies*, 2021, *88* (5), 2344–2374.

Lucas, Robert E, "Expectations and the neutrality of money," *Journal of Economic Theory*, 1972, *4* (2), 103–124.

Mankiw, N. Gregory and Ricardo Reis, "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *The Quarterly Journal of Economics*, 2002, *117* (4), 1295–1328.

_ **,** _ **, and Justin Wolfers**, "Disagreement about Inflation Expectations," *NBER Macroeconomics Annual*, 2003, *18*, 209–248.

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

Stock, James H and Mark W Watson, "Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics," in "Handbook of macroeconomics," Vol. 2, Elsevier, 2016, pp. 415–525.

Woodford, Michael, "Imperfect Common Knowledge and the Effects of Monetary Policy," in P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford, eds., *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honour of Edmund S. Phelps*, Princeton: Princeton University Press, 2002.

Appendix

Appendix A Information set of professional forecasters

Appendix A.1 SPF questionnaire

SPF 2014:Q1

Section 1. U.S. Business Indicators

Forecaster:
Date:

		Quarterly Data						Annual Data [a]				
	L / G	2013:Q4	2014:Q1	2014:Q2	2014:Q3	2014:Q4	2015:Q1	2013	2014	2015	2016	2017
1. Nominal GDP		17102.5						16802.9				
2. GDP Price Index (Chain)		107.02						106.47				
3. Corporate Prof After Tax		.						.				
4. Civilian Unemp Rate	L	7.0						7.4				
5. Nonfarm Payroll Employment [b]		136747						135927				
6. Industrial Prod Index		101.2						99.6				
7. Housing Starts		1.002						0.928				
8. T-Bill Rate, 3-month	L	0.06						0.06				
9. AAA Corp Bond Yield	L	4.59						4.24				
10. BAA Corp Bond Yield	L	5.36						5.10				
11. Treasury Bond Rate, 10-year	L	2.75						2.35				

[a] If you provide your forecasts in growth rates, your annual forecasts in Sections 1 and 2 should be computed as the growth in annual-average level.

[b] Please provide your forecasts for nonfarm payroll employment either in levels (thousands of jobs, seasonally adjusted) or annualized growth rates.

Do your forecasts for Nonfarm Payrolls include the February 7, 2014 benchmark revision?

Did you use (check one): Unrevised Data? Revised Data?

Section 2. Real GDP and Its Components

		Quarterly Data						Annual Data [a]				
Chain-weighted (2009$)	L / G	2013:Q4	2014:Q1	2014:Q2	2014:Q3	2014:Q4	2015:Q1	2013	2014	2015	2016	2017
12. Real GDP		15965.6						15767.1				
13. Real Personal Cons Expenditures		10832.8						10728.2				
14. Real Nonres Fixed Investment		2013.5						1982.1				
15. Real Res Fixed Investment		486.5						486.0				
16. Real Fed Government C & GI		1125.2						1157.5				
17. Real State & Local Govt C & GI		1745.4						1739.7				
18. Real Change in Private Inventories	L	127.2						85.4				
19. Real Net Exports of Goods & Services	L	-370.1						-409.1				

Section 3. CPI and PCE Inflation

	Quarterly Data (Q/Q)						Annual Data (Q4/Q4) [c]			
	2013:Q4	2014:Q1	2014:Q2	2014:Q3	2014:Q4	2015:Q1	2013	2014	2015	2016
20. CPI Inflation Rate	0.9						1.2			
21. Core CPI Inflation Rate	1.6						1.7			
22. PCE Inflation Rate	0.7						0.9			
23. Core PCE Inflation Rate	1.1						1.1			

[c] Annual growth rate forecasts in Section 3 should be computed as a fourth-quarter over fourth-quarter percent change.

Appendix Figure A.1: SAMPLE SPF SURVEY FORM

Historical Economic Data (as of July 26, 2019)
Survey of Professional Forecasters
Research Department, Federal Reserve Bank of Philadelphia

Section 1 - U.S. Business Indicators	**2018Q3**	**2018Q4**	**2019Q1**	**2019Q2**	**2018**
1. Nominal Gross Domestic Product	20749.8	20897.8	21098.8	21337.9	20580.3
2. GDP Chain-Weighted Price Index	110.77	111.21	111.50	112.16	110.38
3. Corporate Profits After Tax	1873.9	1867.1	1791.4	.	1854.9
4. Civilian Unemployment Rate	3.8	3.8	3.9	3.6	3.9
5. Nonfarm Payroll Employment	149409	150058	150675	151135	149064
6. Industrial Production Index	109.3	110.3	109.8	109.5	108.6
7. Housing Starts	1.233	1.185	1.213	1.263	1.250
8. Treasury Bill Rate, 3-month	2.04	2.32	2.39	2.30	1.94
9. Moody's AAA Corporate Bond Yield *	.	.	.	.	.
10. Moody's BAA Corporate Bond Yield *	.	.	.	.	.
11. Treasury Bond Rate, 10-year	2.93	3.03	2.65	2.33	2.91
Section 2 - Real GDP & Components (chain-weighted)	**2018Q3**	**2018Q4**	**2019Q1**	**2019Q2**	**2018**
12. Real Gross Domestic Product	18732.7	18783.5	18927.3	19023.8	18638.2
13. Real Personal Consumption Expenditures	13019.8	13066.3	13103.3	13241.1	12944.6
14. Real Nonresidential Fixed Investment	2703.9	2735.8	2765.6	2761.4	2692.3
15. Real Residential Fixed Investment	600.1	593.0	591.4	589.1	602.9
16. Real Federal Government C & GI	1238.7	1242.1	1248.8	1272.7	1232.2
17. Real State & Local Government C & GI	1997.7	1991.4	2007.9	2023.9	1990.1
18. Real Change in Private Inventories	87.2	93.0	116.0	71.7	48.2
19. Real Net Exports of Goods & Services	-962.4	-983.0	-944.0	-978.7	-920.0
Section 3 - CPI and PCE Inflation	**2018Q3**	**2018Q4**	**2019Q1**	**2019Q2**	**2018 (Q4/Q4)**
20. CPI Inflation	2.0	1.5	0.9	2.9	2.2
21. Core CPI Inflation	2.0	2.2	2.3	1.8	2.2
22. PCE Inflation	1.6	1.3	0.4	2.3	1.9
23. Core PCE Inflation	1.6	1.7	1.1	1.8	1.9

Selected Monthly Economic Data	**JAN2019**	**FEB2019**	**MAR2019**	**APR2019**	**MAY2019**	**JUN2019**
Civilian Unemployment Rate	4.0	3.8	3.8	3.6	3.6	3.7
Nonfarm Payroll Employment	150587	150643	150796	151012	151084	151308
Industrial Production Index	110.1	109.6	109.7	109.2	109.6	109.6
Housing Starts	1.291	1.149	1.199	1.270	1.265	1.253
Treasury Bill Rate, 3-month	2.37	2.39	2.40	2.38	2.35	2.17
Moody's AAA Corporate Bond Yield *	.	.	.	.	.	.
Moody's BAA Corporate Bond Yield *	.	.	.	.	.	.
Treasury Bond Rate, 10-year	2.71	2.68	2.57	2.53	2.40	2.07

* Moody's Aaa and Baa rates are proprietary. The Philadelphia Fed cannot provide the historical values, except upon a special request to Tom Stark. You must send an email to Tom.Stark@phil.frb.org to request the data and agree to limit usage of the data to the *Survey of Professional Forecasters*.

Appendix Figure A.2: SAMPLE SPF HISTORICAL DATA SHEET

Appendix A.2 Other important surveys of professional forecasters

Apart from the SPF data set, several survey forecast data sets are widely used in macroeconomics. The Livingston survey was started by American journalist Joseph Livingston and has been conducted since 1946 and is now managed by the Philadelphia Fed. It is the oldest continuous survey of economists' expectations for the US. As is explained in the Livingston survey documentation (p. 11), the survey forms contain the last historical values known at the time the survey questionnaires were mailed to panelists. Carlson (1977), a reference recommended by the survey documentation, also explained the survey design: "*Along with the questionnaire he [Joseph Livingston] provides the most current data when available on the economic variables to be forecast*" (see p. 28). Figures A.3 - A.5 provide a sample survey form and historical data sheet sent to panelists, both obtained from the survey team. The survey form and datasheet provide panelists with data on the most recent four quarters for quarterly variables, six months for monthly variables, and three years for annual variables.

Consensus Economics Inc. has been conducting surveys of professional forecasters since 1989. The surveys cover a large sample of countries including G7 countries and Western European economies. Figures A.6 and A.7 provide a sample survey form for Consensus Economics surveys. Another survey data set, the European Central Bank Survey of Professional Forecasters, is the longest-running survey of euro area macro expectations. Figure A.8, taken from the ECB SPF documentation, explains the information provided to survey participants for the ECB SPF survey. Like the SPF and Livingston surveys, both surveys provide the most recent data release to panelists in every survey round.[5]

[5]Steven Hubbard, Vice President of Consensus Economics Inc., confirmed that Consensus Economics surveys have been providing the most recent data release to panelists since 1989 (the start of the survey) and provided us with the sample survey form.

Page 1 of 3

Federal Reserve Bank of Philadelphia

Livingston Survey

December 2022

Please : 1. Update contact information below
2. See the worksheet: "Historical Data" for historical data values of the variables you will forecast
3. Provide your forecasts in the worksheet: "Livingston Questionnaire"
4. Send forecasts to : **phil.liv@phil.frb.org**

Name: ______________________________

Address: ______________________________

____

Phone Number: ______________________________

E-Mail: ______________________________

Date completed: ______________________________

Appendix Figure A.3: SAMPLE LIVINGSTON SURVEY FORM AND HISTORICAL DATA SHEET (PAGE 1)

Table A

FORECAST

Name & Address: ______________________________

____

____

Phone: ______________________________

Email: ______________________________

Date Completed: ______________________________

Please check the category (listed below) that best describes your principal area of business:

____ Academic Institution
____ Consulting
____ Government
____ Industry Trade Group
____ Labor
____ Commercial Banking
____ Federal Reserve System
____ Insurance
____ Investment Banking
____ Non-financial Business

QUARTERLY INDICATORS		Quarterly Data 2022 Q2	2022 Q3	2022 Q4	2023 Q2	2023 Q4	Annual Data 2021	2022	2023	2024
1. REAL GROSS DOMESTIC PRODUCT	2012 C-W $Bil, SAAR	19895.3	20021.7				19609.8			
2. GROSS DOMESTIC PRODUCT	$Billions, SAAR	25248.5	25663.3				23315.1			
3. REAL NONRESID. FIXED INVESTMENT	2012 C-W $Bil, SAAR	2915.5	2942.4				2835.4			
4. CORPORATE PROFITS AFTER TAXES	$Billions, SAAR	2522.6					2382.9			

MONTHLY INDICATORS		Monthly Data 2022 JUN	2022 OCT	2022 DEC	2023 JUN	2023 DEC	Annual Data 2021	2022	2023	2024
5. INDUSTRIAL PRODUCTION	2012=100, SA	104.1	104.7				100.0			
6. TOTAL PRIVATE HOUSING STARTS	Millions of Units, SAAR	1.575	1.425				1.605			
7. PRODUCER FINISHED GOODS PRICES	1982=100, SA	259.7	255.9				221.0			
8. CONSUMER PRICES	1982-84=100, SA	295.3	298.1				271.0			
9. UNEMPLOYMENT RATE	SA, %	3.6	3.7				5.4			
10. AVG. WEEKLY EARNINGS IN MFG.	$, Not SA	1018.0	1045.2				985.7			
11. RETAIL SALES	$Billions, SAMR	684.1	694.5				619.6			
12. AUTOMOBILE SALES (incl. foreign)	Millions of Units, SAAR	2.7	3.2				3.4			

INTEREST RATES & STOCK PRICES (End-of-Period)		Monthly Data, End-of-Period 2022 30-Jun	2022 31-Oct	2022 30-Dec	2023 30-Jun	2023 29-Dec	Annual, End-of-Period 2021	2022	2023	2024
13. PRIME INTEREST RATE	Commercial Banks, %	4.75	6.25				3.25	SAME AS DEC 30, 2022	SAME AS DEC 29 2023	
14. 10-YR U.S. TREASURY BOND	Secondary Market, %	2.98	4.10				1.52			
15. 90-DAY U.S. TREASURY BILL	Secondary Market, %	1.66	4.06				0.06			
16. STOCK PRICES (S&P500)	1941-43=10	3785.38	3871.98				4766.18			

What is your forecast of the average annual rate of change in the CPI-U for the next ten years? __________%

What is your forecast of the average annual rate of change in Real Gross Domestic Product for the next ten years? __________%

Appendix Figure A.4: SAMPLE LIVINGSTON SURVEY FORM AND HISTORICAL DATA SHEET (PAGE 2)

Table B

HISTORICAL DATA for DECEMBER SURVEY

QUARTERLY INDICATORS		Quarterly Data 2021 Q1	Q2	Q3	Q4	2022 Q1	Q2	Q3	Annual Data 2019	2020	2021
1. REAL GROSS DOMESTIC PRODUCT	2012 C-W $Bil, SAAR	19216.2	19544.2	19672.6	20006.2	19924.1	19895.3	20021.7	19036.1	18509.2	19609.8
2.GROSS DOMESTIC PRODUCT	$Billions, SAAR	22313.9	23046.9	23550.4	24349.1	24740.5	25248.5	25663.3	21381.0	21060.5	23315.1
3. REAL NONRESID. FIXED INVESTMENT	2012 C-W $Bil, SAAR	2781.4	2847.7	2852.2	2860.2	2915.0	2915.5	2942.4	2804.6	2666.0	2835.4
4. CORPORATE PROFITS AFTER TAXES	$Billions, SAAR	2237.4	2401.7	2456.4	2435.9	2374.6	2522.6		2104.8	1971.2	2382.9

MONTHLY INDICATORS		Monthly Data 2022 APRIL	MAY	JUNE	JULY	AUG	SEPT	OCT	Annual Data 2019	2020	2021
5. INDUSTRIAL PRODUCTION	2012=100, SA	104.3	104.2	104.1	104.8	104.7	104.8	104.7	102.5	95.3	100.0
6. TOTAL PRIVATE HOUSING STARTS	Millions of Units, SAAR	1.805	1.562	1.575	1.377	1.508	1.488	1.425	1.291	1.395	1.605
7. PRODUCER FINISHED GOODS PRICES	1982=100, SA	248.6	252.9	259.7	254.9	252.2	253.1	255.9	205.7	203.0	221.0
8. CONSUMER PRICES	1982-84=100, SA	288.7	291.5	295.3	295.3	295.6	296.8	298.1	255.6	258.8	271.0
9. UNEMPLOYMENT RATE	SA, %	3.6	3.6	3.6	3.5	3.7	3.5	3.7	3.7	8.1	5.4
10. AVG. WEEKLY EARNINGS IN MFG.	$, Not SA	1009.7	1024.6	1018.0	1021.6	1027.1	1043.2	1045.2	921.9	928.4	985.7
11. RETAIL SALES	$Billions, SAMR	674.7	677.1	684.1	681.1	685.7	685.8	694.5	514.6	517.5	619.6
12. AUTOMOBILE SALES (incl. foreign)	Millions of Units, SAAR	2.9	2.6	2.7	2.7	2.8	2.9	3.2	4.7	3.4	3.4

INTEREST RATES & STOCK PRICES (End-of-Period)		Monthly Data, End-of-Period 2022 29-Apr	31-May	30-Jun	29-Jul	31-Aug	30-Sep	31-Oct	Annual, End-of-Period 2019	2020	2021
13. PRIME INTEREST RATE	Commercial Banks, %	3.50	4.00	4.75	5.50	5.50	6.25	6.25	4.75	3.25	3.25
14. 10-YR U.S. TREASURY BOND	Secondary Market, %	2.89	2.85	2.98	2.67	3.15	3.83	4.10	1.92	0.93	1.52
15. 90-DAY U.S. TREASURY BILL	Secondary Market, %	0.83	1.13	1.66	2.34	2.87	3.22	4.06	1.52	0.09	0.06
16. STOCK PRICES (S&P500)	1941-43=10	4131.93	4132.15	3785.38	4130.29	3955.00	3585.62	3871.98	3230.78	3756.07	4766.18

Appendix Figure A.5: SAMPLE LIVINGSTON SURVEY FORM AND HISTORICAL DATA SHEET (PAGE 3)

UNITED STATES - ECONOMIC SURVEY - MAY 2023

RETURN TO: CONSENSUS ECONOMICS INC.

by e-mail: *cf@consensuseconomics.com*

ConsensusEconomics®

Please enter your details below:

Name: ***Company:*** ***Date:***

ECONOMIC FORECASTS (CALENDAR YEAR BASIS, unless otherwise stated)

Page 1 of2

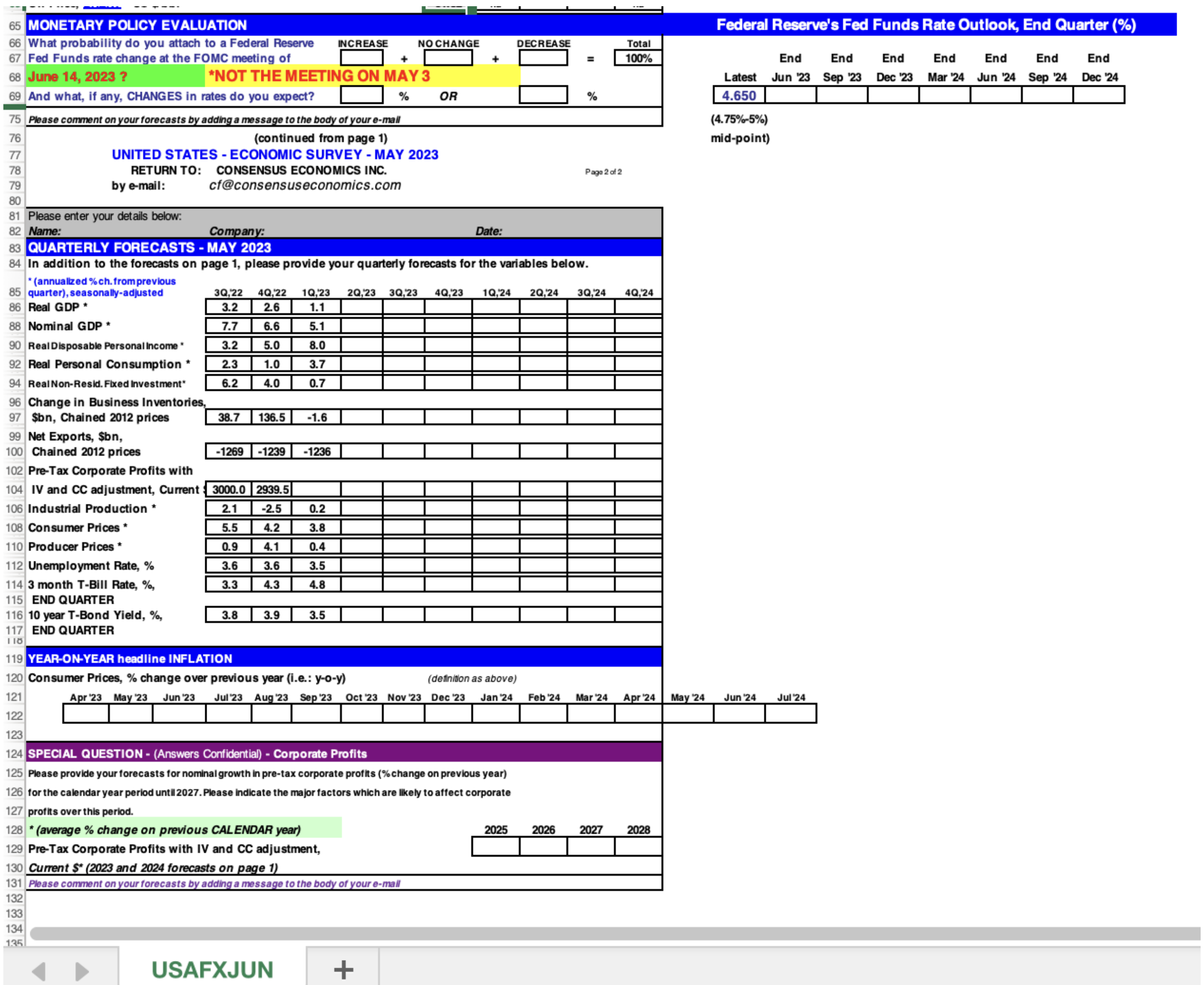
DEADLINE
May 9

* (average % change on previous CALENDAR year)	2022	2023	2024
Gross Domestic Product, Chained 2012 $ *	2.1		
Gross Domestic Product, Current $ *	9.2		
Disposable Personal Income, Chained 2012 $ *	-6.1		
Personal Consumption Expenditures, Chained 2012 $ *	2.7		
Government Consumption Expenditures and Gross Inv., Chained 2012 $ *	-0.6		
Private Non-Residential Fixed Investment, Chained 2012 $ *	3.9		
Pre-Tax Corporate Profits with IV and CC adjustments, Current $ *	6.5		
Change in Business Inventories, $bn, Chained 2012 Prices	125.0		
Net Exports of Goods and Services, $bn, Chained 2012 Prices	-1357		
Industrial Production - Total Index, 2017=100 *	3.4		
Consumer Price Index - All Urban Consumers, 1982/84=100 *	8.0		
NEW: Core PCE Prices (ex. food & energy), 2012=100 *	5.0		
Producer Price Index - Commodities, Finished Goods, 1982=100 *	13.5		
Employment Cost Index - Total Civilian Workers, December 2005=100 *	4.9		
New Auto and Light Truck Sales (including imports), Million Units	13.8		
New Privately Owned Housing Units Started, Million Units	1.56		
Unemployment Rate as a % of Civilian Labor Force, year average	3.6		
Current Account Balance (Balance of Payments), $bn	-944		
	FY21/22	FY22/23	FY23/24
Total Federal Budget Balance, FISCAL YEARS ending Sept 30th, $bn	-1375		
(i.e. FY 21/22 = October 1st, 2021 through to September 30th, 2022)			

INTEREST RATE FORECASTS	Latest	End Aug '23	End May '24
3 month US Treasury Bill Interest Rate (secondary market), % Yield Basis	5.0		
Yield on 10 Year Benchmark Treasury Bond (3.50%, February 2033), %	3.4		

EXCHANGE RATES AND OIL PRICES	Latest	End Jun '23	End Aug '23	End May '24	End May '25
Japanese Yen/US Dollar	134.1				
US Dollars/Euro	1.101				
US Dollars/UK Pound	1.247				
Canadian Dollars/US Dollar	1.361				
Oil Price, BRENT - US $/bbl	81.32	na			na

Appendix Figure A.6: Sample Consensus Economics survey form (page 1)



MONETARY POLICY EVALUATION

What probability do you attach to a Federal Reserve Fed Funds rate change at the FOMC meeting of **June 14, 2023 ?** ***NOT THE MEETING ON MAY 3**

INCREASE		NO CHANGE		DECREASE		Total
	+		+		=	100%

And what, if any, CHANGES in rates do you expect? ____ % *OR* ____ %

Please comment on your forecasts by adding a message to the body of your e-mail

Federal Reserve's Fed Funds Rate Outlook, End Quarter (%)

Latest	End Jun '23	End Sep '23	End Dec '23	End Mar '24	End Jun '24	End Sep '24	End Dec '24
4.650							

(4.75%-5%)
mid-point)

(continued from page 1)

UNITED STATES - ECONOMIC SURVEY - MAY 2023

RETURN TO: **CONSENSUS ECONOMICS INC.**

by e-mail: *cf@consensuseconomics.com*

Page 2 of 2

Please enter your details below:

Name: ***Company:*** ***Date:***

QUARTERLY FORECASTS - MAY 2023

In addition to the forecasts on page 1, please provide your quarterly forecasts for the variables below.

* (annualized % ch. from previous quarter), seasonally-adjusted	3Q,'22	4Q,'22	1Q,'23	2Q,'23	3Q,'23	4Q,'23	1Q,'24	2Q,'24	3Q,'24	4Q,'24
Real GDP *	3.2	2.6	1.1							
Nominal GDP *	7.7	6.6	5.1							
Real Disposable Personal Income *	3.2	5.0	8.0							
Real Personal Consumption *	2.3	1.0	3.7							
Real Non-Resid. Fixed Investment*	6.2	4.0	0.7							
Change in Business Inventories, $bn, Chained 2012 prices	38.7	136.5	-1.6							
Net Exports, $bn, Chained 2012 prices	-1269	-1239	-1236							
Pre-Tax Corporate Profits with IV and CC adjustment, Current $	3000.0	2939.5								
Industrial Production *	2.1	-2.5	0.2							
Consumer Prices *	5.5	4.2	3.8							
Producer Prices *	0.9	4.1	0.4							
Unemployment Rate, %	3.6	3.6	3.5							
3 month T-Bill Rate, %, END QUARTER	3.3	4.3	4.8							
10 year T-Bond Yield, %, END QUARTER	3.8	3.9	3.5							

YEAR-ON-YEAR headline INFLATION

Consumer Prices, % change over previous year (i.e.: y-o-y) *(definition as above)*

Apr '23	May '23	Jun '23	Jul '23	Aug '23	Sep '23	Oct '23	Nov '23	Dec '23	Jan '24	Feb '24	Mar '24	Apr '24	May '24	Jun '24	Jul '24

SPECIAL QUESTION - (Answers Confidential) - Corporate Profits

Please provide your forecasts for nominal growth in pre-tax corporate profits (% change on previous year) for the calendar year period until 2027. Please indicate the major factors which are likely to affect corporate profits over this period.

* (average % change on previous CALENDAR year)	2025	2026	2027	2028
Pre-Tax Corporate Profits with IV and CC adjustment, *Current $* (2023 and 2024 forecasts on page 1)*				

Please comment on your forecasts by adding a message to the body of your e-mail

USAFXJUN

Appendix Figure A.7: SAMPLE CONSENSUS ECONOMICS SURVEY FORM (PAGE 2)

Statistical definition of the variables included in the SPF questionnaire and basic information supplied to survey participants

Variables forecast

Forecasts are requested for the following euro area variables:

– ***Harmonised Index of Consumer Prices (HICP)*** **inflation** as published by Eurostat. Annual rates of growth.

– ***Real gross domestic product (GDP)*** according to the definition of the European System of National and Regional Accounts 1995 (ESA 95) as published by Eurostat. Annual rates of growth.

– ***Unemployment rate*** expressed as a percentage of the labour force.

Basic information supplied to participants

In each survey round, participants are supplied with the latest available data released for each of the variables requested. The basic information supplied in the 2003 Q2 SPF is given below as an example:

Basic reference data for the 2003 Q2 SPF

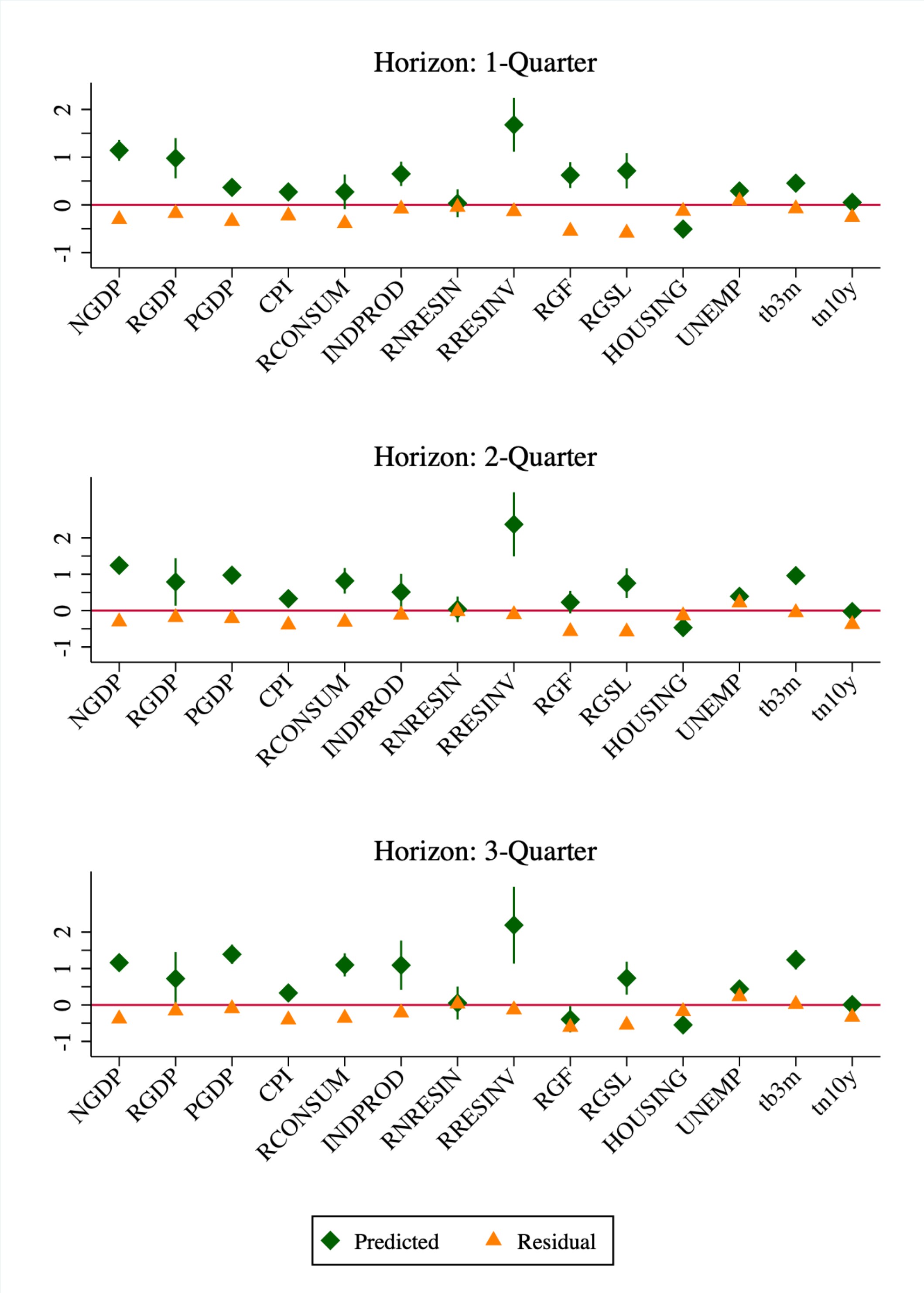
HICP inflation (March 2003)	2.4%
Annual GDP growth (2002 Q4)	1.3% (according to the ESA 95 definition)
Unemployment rate (February 2003)	8.7%

Appendix Figure A.8: ECB SPF SURVEY INFORMATION

Appendix B Additional results on empirical analyses

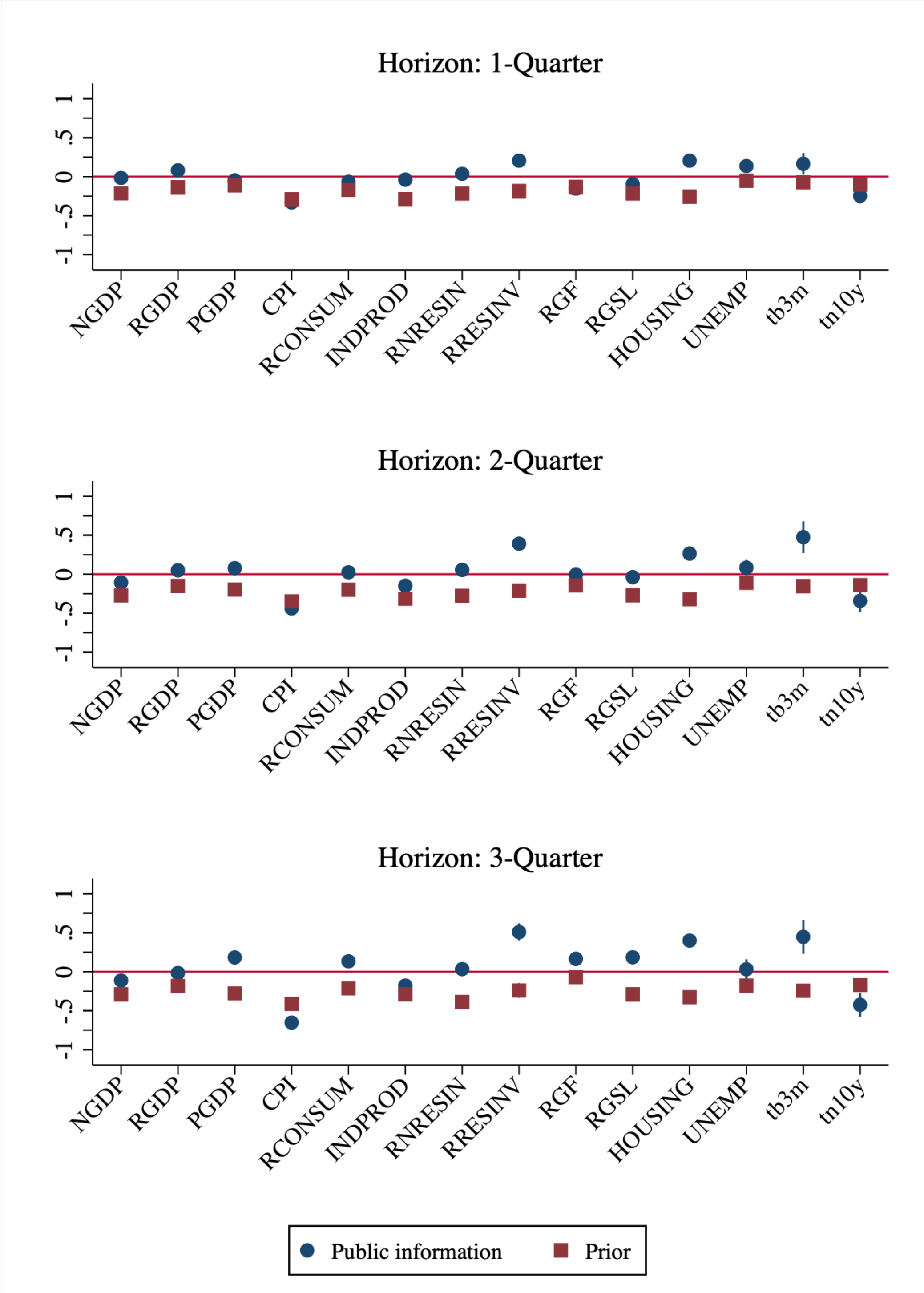
Appendix B.1 One-dimensional public information

In this appendix, we consider a special case where x_t in Eqn. (2.3) is one-dimensional. Specifically, x_t is the most recent release on the dependent variable y, the realized value of y in the previous period. We repeat the analysis in Section 2.4 and report the results in Figure B.9 - B.11.



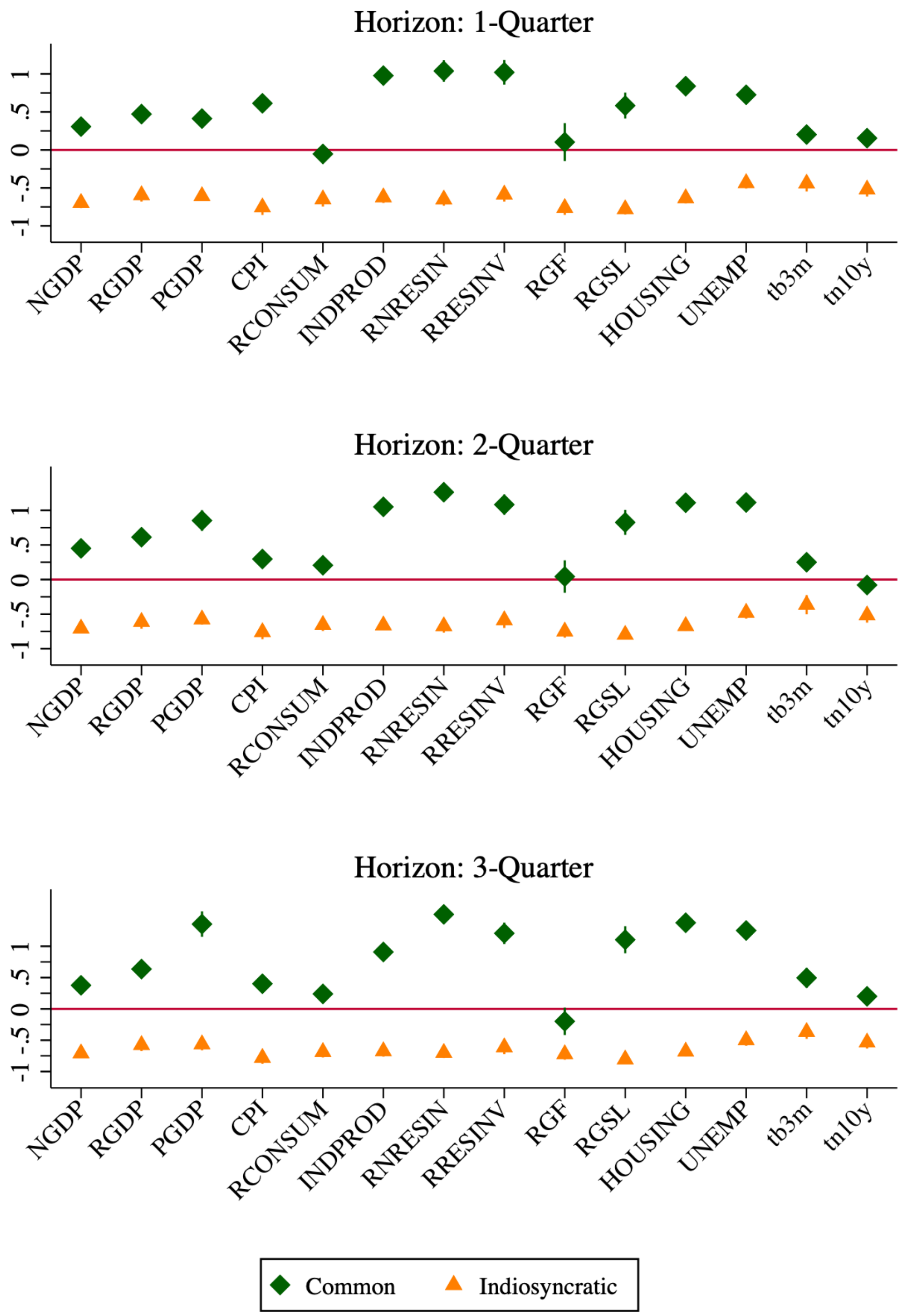
Appendix Figure B.9: RESPONSES OF FORECAST ERRORS TO FORECAST REVISION DECOMPOSITION: 1-DIMENSIONAL SIGNAL

Notes: This figure plots the coefficients of β_1^h (in green) and β_2^h (in orange) from Eqn. (2.6). The regressors of interest are FR predicted using the latest release of the dependent variable (in green) and FR residuals (in orange). 95% confidence intervals based on clustered standard errors are reported.



Appendix Figure B.10: RESPONSES OF FORECAST ERRORS TO PRIOR AND REAL-TIME DATA RELEASE: 1-DIMENSIONAL SIGNAL

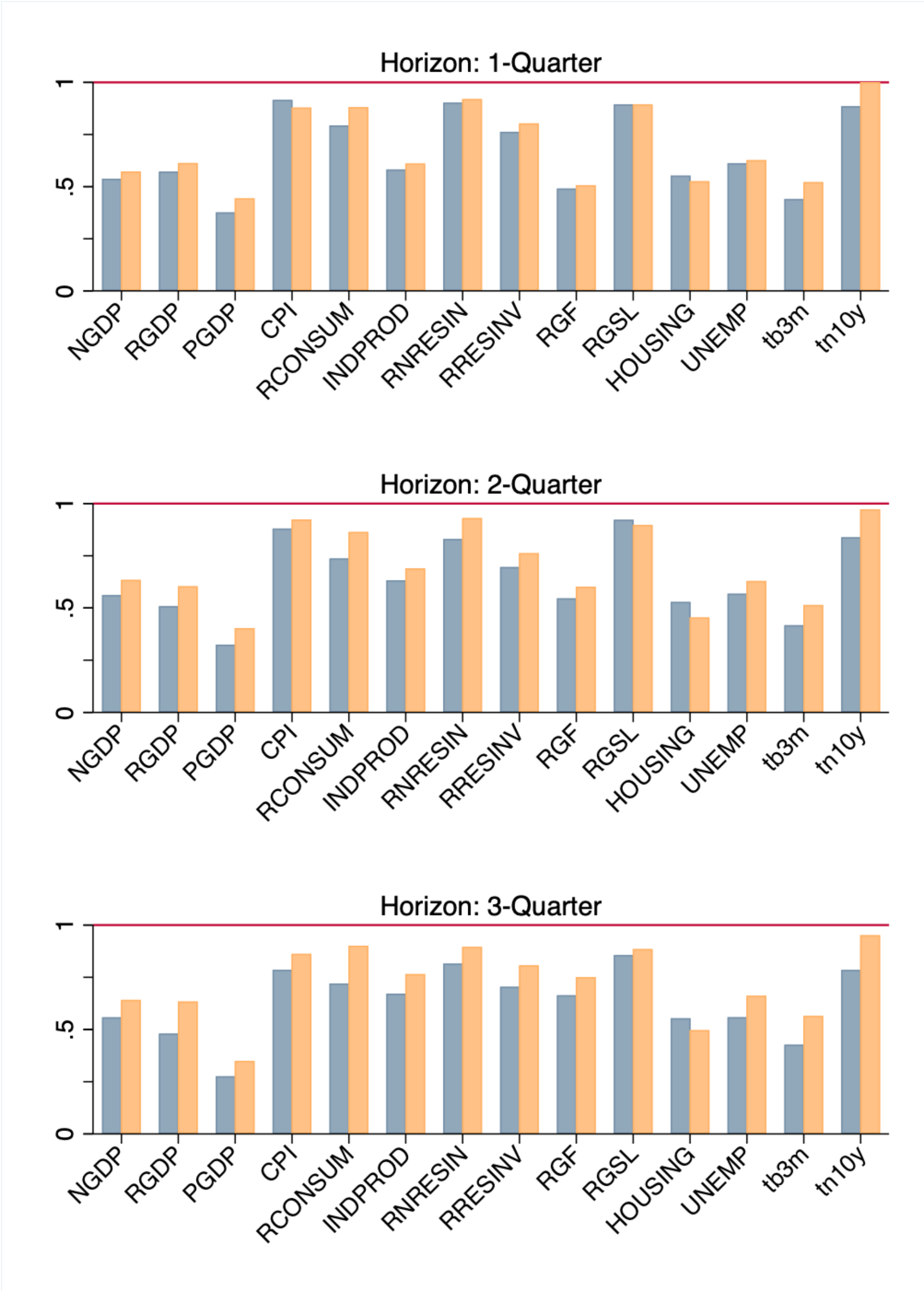
Notes: This figure plots the estimated coefficients of α_1^h (in blue) and α_2^h (in maroon) from Eqn. (2.7). 95% confidence intervals based on clustered standard errors are reported.



Appendix Figure B.11: RESPONSES OF FORECAST ERRORS TO PRIVATE INFORMATION DECOMPOSITION: 1-DIMENSIONAL SIGNAL

Notes: This figure plots the estimated coefficients of θ_1^h (in green) and θ_2^h (in orange) from Eqn. (2.10). 95% confidence intervals based on clustered standard errors are reported.

Appendix B.2 Decomposing forecast errors



Appendix Figure B.12: IMPROVING OVER CONSENSUS FORECASTS: FURTHER ANALYSIS

Notes: This figure analyzes the source of forecast error improvement in figure 6. The blue bars represent the forecast errors of equation (4.2), which are also depicted in figure 6; the orange bars represent the forecast errors of a modified consensus-CG regression in which we add the common component of residual information and the average public information to equation (4.2). All forecast errors are expressed relative to those implied by equation (4.1).

Appendix C Derivation of regression coefficients

Appendix C.1 Forecast error and forecast revision

We first derive the expression of forecast error and forecast revision under the general prediction rule given by Eqn. (3.5). The forecast error at time t is:

$$\begin{aligned}
FE_t^i \equiv \pi_t - \pi_{t|t}^i &= \rho\pi_{t-1} + u_t - \left[(1-\kappa_x-\kappa_y)\pi_{t|t-1}^i + \kappa_x x_{it} + \kappa_y \rho s_t\right] \\
&= (1-\kappa_x-\kappa_y)\rho(\pi_{t-1} - \pi_{t-1|t-1}^i) + (1-\kappa_x)u_t - \kappa_x\epsilon_{it} - \kappa_y\rho\nu_t. && \text{(C.1)} \\
\implies \mathbb{V}\mathrm{ar}(FE^i) &= \frac{(1-\kappa_x)^2\sigma_u^2 + \kappa_x^2\sigma_\epsilon^2 + \kappa_y^2\rho^2\sigma_\nu^2}{1-\rho^2(1-\kappa_x-\kappa_y)^2}. && \text{(C.2)}
\end{aligned}$$

The forecast revision at time t is:

$$\begin{aligned}
FR_t^i \equiv \pi_{t|t}^i - \pi_{t|t-1}^i &= \kappa_x(x_{it} - \pi_{t|t-1}^i) + \kappa_y(\rho s_t - \pi_{t|t-1}^i) \\
&= \kappa_x(\pi_t + \epsilon_{it}^x - \pi_{t|t-1}^i) + \kappa_y\rho(\pi_{t-1} - \pi_{t-1|t-1}^i + \nu_t) \\
&= (\kappa_x+\kappa_y)\rho(\pi_{t-1} - \pi_{t-1|t-1}^i) + \kappa_x(u_t + \epsilon_{it}^x) + \kappa_y\rho\nu_t, && \text{(C.3)} \\
FR_{t+h}^i \equiv \pi_{t+h|t}^i - \pi_{t+h|t-1}^i &= \rho^h(\pi_{t|t}^i - \pi_{t|t-1}^i) \\
&= \rho^h\Big((\kappa_x+\kappa_y)\rho(\pi_{t-1} - \pi_{t-1|t-1}^i) + \kappa_x(\epsilon_{it}^x + u_t) + \kappa_y\rho\nu_t\Big). && \text{(C.4)} \\
\implies \mathbb{V}\mathrm{ar}(FR^i) &= (\kappa_x+\kappa_y)^2\rho^2\mathbb{V}\mathrm{ar}(FE^i) + \kappa_x^2(\sigma_u^2+\sigma_\epsilon^2) + \kappa_y^2\rho^2\sigma_\nu^2. && \text{(C.5)}
\end{aligned}$$

Next, we derive the expression of $\mathbb{E}\left[(\pi_t - \pi_{t|t}^i)\pi_t\right]$:

$$\mathbb{E}\left[(\pi_t - \pi_{t|t}^i)\pi_t\right] = \mathbb{E}\left[(\pi_t - \pi_{t|t}^i)\pi_t\right] - \mathbb{E}\left[(\pi_t - \pi_{t|t}^i)^2\right],$$

From Eqn. (C.1), we get

$$\begin{aligned}
\mathbb{E}\left[(\pi_t - \pi_{t|t}^i)\pi_t\right] &= \mathbb{E}\left[\Big((1-\kappa_x)(1-\omega)\rho(\pi_{t-1} - \pi_{t-1|t-1}^i) + (1-\kappa_x)u_t - \kappa_x\epsilon_{it} - (1-\kappa_x)\omega\rho\nu_t\Big)(\rho\pi_{t-1} + u_t)\right] \\
&= (1-\kappa_x)\sigma_u^2 + \rho^2(1-\kappa_x)(1-\omega)\mathbb{E}\left[(\pi_{t-1} - \pi_{t-1|t-1}^i)\pi_{t-1}^i\right].
\end{aligned}$$

Therefore,

$$\mathbb{E}\left[(\pi_t - \pi_{t|t}^i)\pi_t\right] = \frac{(1-\kappa_x)\sigma_u^2}{1-\rho^2(1-\kappa_x)(1-\omega)} \tag{C.6}$$

$$\mathbb{E}\left[(\pi_t - \pi_{t|t}^i)\pi_{t|t}^i\right] = \frac{(1-\kappa_x)\sigma_u^2}{1-\rho^2(1-\kappa_x)(1-\omega)} - \mathbb{V}\mathrm{ar}(FE^i) \tag{C.7}$$

Appendix C.2 Compute individual CG coefficients

The individual-level CG coefficient is

$$\beta^p = \frac{\mathbb{C}\mathrm{ov}\left(\pi_{t+h} - \pi^i_{t+h|t}, \pi^i_{t+h|t} - \pi^i_{t+h|t-1}\right)}{\mathbb{V}\mathrm{ar}\left(\pi^i_{t+h|t} - \pi^i_{t+h|t-1}\right)} \tag{C.8}$$

$$= \frac{\mathbb{C}\mathrm{ov}\left(\rho^h(\pi_t - \pi^i_{t|t}), \rho^h(\pi^i_{t|t} - \pi^i_{t|t-1})\right)}{\mathbb{V}\mathrm{ar}\left(\rho^h(\pi^i_{t|t} - \pi^i_{t|t-1})\right)} \tag{C.9}$$

$$= \frac{\mathbb{C}\mathrm{ov}(FE^i_t, FR^i_t)}{\mathbb{V}\mathrm{ar}(FR^i_t)} \tag{C.10}$$

In particular,

$$\begin{aligned}\mathbb{C}\mathrm{ov}\left(FE^i_t, FR^i_t\right) &= (1-\kappa_x-\kappa_y)(\kappa_x+\kappa_y)\rho^2\mathbb{V}\mathrm{ar}(FE^i) + (1-\kappa_x)\kappa_x\sigma_u^2 - \kappa_x^2\sigma_\epsilon^2 - \rho^2\kappa_y^2\sigma_v^2 \\ &= \left[1-(1-\kappa_x-\kappa_y)\rho^2\right]\mathbb{E}\left[(\pi_t - \pi^i_{t|t})\pi^i_{t|t}\right]\end{aligned} \tag{C.11}$$

When forecasts are optimal, $\mathbb{E}\left[(\pi_t - \pi^i_{t|t})\pi^i_{t|t}\right] = 0$ since forecast errors are $(\pi_t - \pi^i_{t|t})$ are not predictable, and are therefore, orthogonal to the forecasts $(\pi^i_{t|t})$. As a result, $\mathbb{C}\mathrm{ov}\left(FE^i_t, FR^i_t\right) = 0$, forecasters do not over- or under-react to forecast revisions.

Appendix C.3 Compute consensus level CG coefficients

The consensus-level belief is

$$\begin{aligned}\pi^c_{t|t} &= \kappa_x x_t + (1-\kappa_x)\omega\rho s_t + (1-\kappa_x)(1-\omega)\rho\pi^c_{t-1|t-1} \\ &= \kappa_x\left[x_t - \omega\rho s_t - (1-\omega)\rho\pi^c_{t-1|t-1}\right] + \left[\omega\rho s_t + (1-\omega)\rho\pi^c_{t-1|t-1}\right]\end{aligned} \tag{C.12}$$

$$\pi_t - \pi^c_{t|t} = (1-\kappa_x)\left[x_t - \omega\rho s_t - (1-\omega)\rho\pi^c_{t-1|t-1}\right] \tag{C.13}$$

where $x_t = \frac{\sum_i x_{it}}{N_t}$ with $\mathbb{V}\mathrm{ar}(x_t) = \frac{\sigma_\epsilon^2}{N_t}$, and N_t is the number of forecasters in period t.

The consensus CG coefficient is

$$\beta^c \propto \mathbb{C}\mathrm{ov}\left(\overline{FE}_t, \overline{FR}_t\right) \tag{C.14}$$

$$\begin{aligned}&= \left[1-(1-\kappa_x-\kappa_y)\rho^2\right]\mathbb{E}\left[(\pi_t - \pi^c_{t|t})\pi^c_{t|t}\right] \\ &\propto (\kappa^c_x - \kappa_x)\kappa_x\mathbb{V}\mathrm{ar}\left(\pi_t - \omega\rho s_t - (1-\omega)\rho\pi^c_{t-1|t-1}\right) > 0\end{aligned} \tag{C.15}$$

where κ_x^c represents the optimal weight on x_t such that

$$\mathbb{E}\left[\pi_t \mid x_t; \omega s_t + (1-\omega)\pi^c_{t-1|t-1}\right] \equiv \kappa_x^c x_t + (1-\kappa_x^c)\rho\left(\omega s_t + (1-\omega)\pi^c_{t-1|t-1}\right)$$

denote the optimal forecast of π_t based on the two signals x_t and $\omega s_t + (1-\omega)\rho\pi_{t-1|t-1}$.

To see why the inequality in Eqn. (C.15) holds, first notice that $\kappa_x^c \geq \widehat{\kappa}_x$ when $\tau\sigma_\epsilon^2 \geq \frac{\sigma_\epsilon^2}{N_t}$. Given that optimal forecast errors are unforecastable, and therefore orthogonal to each element of the information set, we have

$$\mathbb{E}\left[\left(\pi_t - \mathbb{E}[\pi_t|x_t, \omega s_t + (1-\omega)\pi^c_{t-1|t-1}]\right)\pi^c_{t|t}\right] = 0 \tag{C.16}$$

$$\mathbb{E}\left[\left(x_t - \rho\omega s_t - \rho(1-\omega)\pi^c_{t-1|t-1}\right)\left(\rho\omega s_t + \rho(1-\omega)\pi^c_{t-1|t-1}\right)\right] = 0 \tag{C.17}$$

We get the following:

$$\begin{aligned}
\mathbb{E}\left[(\pi_t - \pi^c_{t|t})\pi^c_{t|t}\right] &= \mathbb{E}\Big[\Big(\pi_t - \mathbb{E}[\pi_t|x_t, \omega s_t + (1-\omega)\pi^c_{t-1|t-1}] \\
&\qquad + \mathbb{E}[\pi_t|x_t, \omega s_t + (1-\omega)\pi^c_{t-1|t-1}] - \pi^c_{t|t}\Big)\pi^c_{t|t}\Big] \\
&= \mathbb{E}\left[\left(\mathbb{E}[\pi_t|x_t, \omega s_t + (1-\omega)\pi^c_{t-1|t-1}] - \pi^c_{t|t}\right)\pi^c_{t|t}\right] \\
&= (\kappa_x^c - \kappa_x)\mathbb{E}\left[\left(x_t - \rho\omega s_t - \rho(1-\omega)\pi^c_{t-1|t-1}\right)\pi^c_{t|t}\right] \\
&= (\kappa_x^c - \kappa_x)\mathbb{E}\Big[\left(x_t - \rho\omega s_t - \rho(1-\omega)\pi^c_{t-1|t-1}\right) \\
&\qquad \left(\kappa_x(x_t - \rho\omega s_t - \rho(1-\omega)\pi^c_{t-1|t-1}) + (\rho\omega s_t + \rho(1-\omega)\pi^c_{t-1|t-1})\right)\Big] \\
&= (\kappa_x^c - \kappa_x)\kappa_x \mathbb{V}\mathrm{ar}\left(x_t - \rho\omega s_t - \rho(1-\omega)\pi^c_{t-1|t-1}\right) > 0
\end{aligned} \tag{C.18}$$

Therefore, β^c is always positive when $\tau > N_t^{-1}$, which holds under RE as $\tau^{RE} = 1$. In the limiting case where $x_t \to \pi_t$ as $N_t \to \infty$, $\kappa_x^c \to 1$ and the consensus-level CG coefficient is always positive when $\kappa_x < 1$.

Appendix C.4 Compute coefficients of regressing forecast revisions on news

Consider the regression model (2.3):

$$\pi^i_{t+h|t} - \pi^i_{t+h|t-1} = \gamma_h(s_t - \pi^i_{t-1|t-1}) + \eta_h \pi^i_{t+h|t-1} + \epsilon^i_{t,h} \tag{C.19}$$

We derive the OLS coefficient estimates as follows:

$$\begin{pmatrix}\gamma_h \\ \eta_h\end{pmatrix} = \begin{pmatrix} \mathbb{E}(s_t - \pi^i_{t-1|t-1})^2 & \mathbb{E}\left[(s_t - \pi^i_{t-1|t-1})\pi^i_{t+h|t-1}\right] \\ \mathbb{E}\left[(s_t - \pi^i_{t-1|t-1})\pi^i_{t+h|t-1}\right] & \mathbb{E}(\pi^i_{t+h|t-1})^2 \end{pmatrix}^{-1}$$

$$\mathbb{E}\left(\begin{pmatrix} s_t-\pi^i_{t-1|t-1} \\ \pi^i_{t+h|t-1}\end{pmatrix}(\pi^i_{t+h|t}-\pi^i_{t+h|t-1})\right)$$

$$=\left(\mathbb{E}(s_t-\pi^i_{t-1|t-1})^2\mathbb{E}(\pi^i_{t+h|t-1})^2-\left(\mathbb{E}\left[(s_t-\pi^i_{t-1|t-1})\pi^i_{t+h|t-1}\right]\right)^2\right)^{-1}$$

$$\begin{pmatrix} \mathbb{E}(\pi^i_{t+h|t-1})^2 & -\mathbb{E}\left[(s_t-\pi^i_{t-1|t-1})\pi^i_{t+h|t-1}\right] \\ -\mathbb{E}\left[(s_t-\pi^i_{t-1|t-1})\pi^i_{t+h|t-1}\right] & \mathbb{E}(s_t-\pi^i_{t-1|t-1})^2 \end{pmatrix}$$

$$\mathbb{E}\left(\begin{pmatrix} s_t-\pi^i_{t-1|t-1} \\ \pi^i_{t+h|t-1}\end{pmatrix}(\pi^i_{t+h|t}-\pi^i_{t+h|t-1})\right).$$

Denote the denominator as $\mathscr{D}_h$,

$$\begin{aligned}
\mathscr{D}_h &\equiv \mathbb{E}(s_t-\pi^i_{t-1|t-1})^2\mathbb{E}(\pi^i_{t+h|t-1})^2-\left(\mathbb{E}\left[(s_t-\pi^i_{t-1|t-1})\pi^i_{t+h|t-1}\right]\right)^2 \\
&= \rho^{2(h+1)}\left(\mathbb{E}(s_t-\pi^i_{t-1|t-1})^2\mathbb{E}(\pi^i_{t-1|t-1})^2-\left(\mathbb{E}\left[(s_t-\pi^i_{t-1|t-1})\pi^i_{t-1|t-1}\right]\right)^2\right) \\
&= \rho^{2(h+1)}\left(\mathbb{E}(\pi_{t-1}-\pi^i_{t-1|t-1})^2\mathbb{E}(\pi^i_{t-1|t-1})^2-\left(\mathbb{E}\left[(\pi_{t-1}-\pi^i_{t-1|t-1})\pi^i_{t-1|t-1}\right]\right)^2\right)+\rho^{2(h+1)}\mathbb{E}(\pi^i_{t-1|t-1})^2\sigma_\nu^2
\end{aligned} \tag{C.20}$$

Note $\mathscr{D}_h$ is always positive due to Cauchy–Schwarz inequality. Next, define the first and second elements of the numerator as $\mathscr{N}_h^\gamma$ and $\mathscr{N}_h^\eta$,

$$\begin{aligned}
\mathscr{N}_h^\gamma &\equiv \mathbb{E}(\pi^i_{t+h|t-1})^2\mathbb{E}\left[(s_t-\pi^i_{t-1|t-1})(\pi^i_{t+h|t}-\pi^i_{t+h|t-1})\right] \\
&\quad -\mathbb{E}\left[(s_t-\pi^i_{t-1|t-1})\pi^i_{t+h|t-1}\right]\mathbb{E}\left[\pi^i_{t+h|t-1}(\pi^i_{t+h|t}-\pi^i_{t+h|t-1})\right] \\
&= \rho^{3h+2}\mathbb{E}(\pi^i_{t-1|t-1})^2\mathbb{E}\left[(s_t-\pi^i_{t-1|t-1})(\pi^i_{t|t}-\pi^i_{t|t-1})\right] \\
&\quad -\rho^{3h+2}\mathbb{E}\left[(s_t-\pi^i_{t-1|t-1})\pi^i_{t-1|t-1}\right]\mathbb{E}\left[\pi^i_{t-1|t-1}(\pi^i_{t|t}-\pi^i_{t|t-1})\right] \\
&= \rho^{3(h+1)}(\kappa_x+\kappa_y)\left(\mathbb{E}(\pi^i_{t-1|t-1})^2\mathbb{E}\left[(\pi_{t-1}-\pi^i_{t-1|t-1})^2\right]-\left(\mathbb{E}\left[(\pi_{t-1}-\pi^i_{t-1|t-1})\pi^i_{t-1|t-1}\right]\right)^2\right) \\
&\quad +\rho^{3(h+1)}\mathbb{E}(\pi^i_{t-1|t-1})^2\kappa_y\sigma_\nu^2 \\
\mathscr{N}_h^\eta &\equiv -\mathbb{E}\left[(s_t-\pi^i_{t-1|t-1})\pi^i_{t+h|t-1}\right]\mathbb{E}\left[(s_t-\pi^i_{t-1|t-1})(\pi^i_{t+h|t}-\pi^i_{t+h|t-1})\right] \\
&\quad +\mathbb{E}\left[\pi^i_{t+h|t-1}(\pi^i_{t+h|t}-\pi^i_{t+h|t-1})\right]\mathbb{E}(s_t-\pi^i_{t-1|t-1})^2 \\
&= -\rho^{2h+1}\mathbb{E}\left[(\pi_{t-1}-\pi^i_{t-1|t-1})\pi^i_{t-1|t-1}\right]\mathbb{E}\left[(\pi_{t-1}-\pi^i_{t-1|t-1})(\pi^i_{t|t}-\pi^i_{t|t-1})\right] \\
&\quad -\rho^{2h+1}\mathbb{E}\left[(\pi_{t-1}-\pi^i_{t-1|t-1})\pi^i_{t-1|t-1}\right]\kappa_y\rho\sigma_\nu^2 \\
&\quad +\rho^{2h+1}\mathbb{E}\left[\pi^i_{t-1|t-1}(\pi^i_{t|t}-\pi^i_{t|t-1})\right]\mathbb{E}(\pi_{t-1}-\pi^i_{t-1|t-1})^2 \\
&\quad +\rho^{2h+1}\mathbb{E}\left[\pi^i_{t-1|t-1}(\pi^i_{t|t}-\pi^i_{t|t-1})\right]\sigma_\nu^2
\end{aligned}$$

$$
\begin{aligned}
&= -\rho^{2h+1}\mathbb{E}\left[(\pi_{t-1}-\pi^i_{t-1|t-1})\pi^i_{t-1|t-1}\right]\kappa_y\rho\sigma_v^2 \\
&\quad +\rho^{2h+1}\mathbb{E}\left[(\pi_{t-1}-\pi^i_{t-1|t-1})\pi^i_{t-1|t-1}\right](\kappa_x+\kappa_y)\rho\sigma_v^2 \\
&= \rho^{2(h+1)}\mathbb{E}\left[(\pi_{t-1}-\pi^i_{t-1|t-1})\pi^i_{t-1|t-1}\right]\kappa_x\sigma_v^2
\end{aligned}
$$

Thus, $\gamma_h = \frac{\mathcal{N}_h^\gamma}{\mathcal{D}_h}$ and $\eta_h = \frac{\mathcal{N}_h^\eta}{\mathcal{D}_h}$. In particular, $0 < \gamma_h \leq \rho^{h+1}(\kappa_x+\kappa_y)$ where the equality holds when $\sigma_v = 0$.

Appendix C.5 Compute coefficients of regressing forecast errors on predicted component and residual

Now we consider the regression model (2.6):

$$
\pi_{t+h}-\pi^i_{t+h|t} = \beta_1^h \times \text{Predicted}^h_{i,t} + \beta_2^h \times \text{Residual}^h_{i,t} + v^i_{t,h}. \tag{C.21}
$$

Given that $\text{Predicted}^h_{i,t}$ and $\text{Residual}^h_{i,t}$ are orthogonal by construction, the OLS coefficient estimates are as following

$$
\begin{aligned}
\beta_1 &= \frac{\mathbb{C}\mathrm{ov}\left(\pi_{t+h}-\pi^i_{t+h|t}, \gamma_h(s_t-\pi^i_{t-1|t-1})+\eta_h\pi^i_{t+h|t-1}\right)}{\mathbb{V}\mathrm{ar}\left(\gamma_h(s_t-\pi^i_{t-1|t-1})+\eta_h\pi^i_{t+h|t-1}\right)}, \\
\beta_2 &= \frac{\mathbb{C}\mathrm{ov}\left(\pi_{t+h}-\pi^i_{t+h|t}, \pi^i_{t+h|t}-\pi^i_{t+h|t-1}-\gamma_h(s_t-\pi^i_{t-1|t-1})-\eta_h\pi^i_{t+h|t-1}\right)}{\mathbb{V}\mathrm{ar}\left(\pi^i_{t+h|t}-\pi^i_{t+h|t-1}-\gamma_h(s_t-\pi^i_{t-1|t-1})-\eta_h\pi^i_{t+h|t-1}\right)},
\end{aligned}
$$

where the numerator of β_1 is

$$
\begin{aligned}
\mathcal{N}^\beta_{1,h} &\equiv \mathbb{C}\mathrm{ov}\left(\rho^h(\pi_t-\pi^i_{t|t}), \gamma_h(\pi_{t-1}+v_t-\pi^i_{t-1|t-1})+\eta_h\pi^i_{t+h|t-1}\right) \\
&= \rho^h\rho(1-\kappa_x-\kappa_y)\gamma_h\mathbb{E}\left[(\pi_{t-1}-\pi^i_{t-1|t-1})^2\right] - \rho^h\gamma_h\kappa_y\rho\sigma_v^2 + \rho^h\eta_h\mathbb{E}\left[(\pi_t-\pi^i_{t|t})\pi^i_{t+h|t-1}\right] \\
&= \rho^{h+1}(1-\kappa_x-\kappa_y)\gamma_h\mathbb{E}\left[(\pi_{t-1}-\pi^i_{t-1|t-1})^2\right] - \rho^{h+1}\gamma_h\kappa_y\sigma_v^2 \qquad \text{(C.22)} \\
&\quad + \rho^{2(h+1)}\eta_h(1-\kappa_x-\kappa_y)\mathbb{E}\left[(\pi_{t-1}-\pi^i_{t-1|t-1})\pi^i_{t-1|t-1}\right]. \qquad \text{(C.23)}
\end{aligned}
$$

Consider the first two terms in $\mathcal{N}^\beta_{1,h}$ as in line (C.22):

$$
\begin{aligned}
&\rho^{h+1}(1-\kappa_x-\kappa_y)\gamma_h\mathbb{E}\left[(\pi_{t-1}-\pi^i_{t-1|t-1})^2\right] - \rho^{h+1}\gamma_h\kappa_y\sigma_v^2 \\
&= \rho^{h+1}\gamma_h\left[(1-\kappa_x-\kappa_y)\mathbb{V}\mathrm{ar}(FE^i) - \kappa_y\sigma_v^2\right] \\
&= \rho^{h+1}\gamma_h\left[(1-\kappa_x)(1-\omega)\mathbb{V}\mathrm{ar}(FE^i) - (1-\kappa_x)\omega\sigma_v^2\right]
\end{aligned}
$$

$$
\begin{aligned}
&= \rho^{h+1}\gamma_h(1-\kappa_x)(1-\omega)\left[\mathbb{V}\mathrm{ar}(FE^i) - \frac{\omega}{1-\omega}\sigma_\nu^2\right] \\
&= \rho^{h+1}\gamma_h(1-\kappa_x)(1-\omega)\left[\mathbb{V}\mathrm{ar}(FE^i) - \frac{\sigma_\tau^2}{\sigma_\nu^2}\sigma_\nu^2\right] \\
&= \rho^{h+1}\gamma_h(1-\kappa_x)(1-\omega)\left[\mathbb{V}\mathrm{ar}(FE^i) - \sigma_\tau^2\right].
\end{aligned} \tag{C.24}
$$

The third term in $\mathcal{N}_{1,h}^{\beta}$ as in line (C.23) is always non-negative since

$$
\begin{aligned}
&\eta_h(1-\kappa_x-\kappa_y)\mathbb{E}\left[(\pi_{t-1}-\pi^i_{t-1|t-1})\pi^i_{t-1|t-1}\right] \\
&\propto \kappa_x(1-\kappa_x-\kappa_y)\left(\mathbb{E}\left[(\pi_{t-1}-\pi^i_{t-1|t-1})\pi^i_{t-1|t-1}\right]\right)^2 \geq 0.
\end{aligned} \tag{C.25}
$$

The numerator of β_2 is

$$
\begin{aligned}
\mathcal{N}_{2,h}^{\beta} &\equiv \mathbb{C}\mathrm{ov}\left(\pi_{t+h}-\pi^i_{t+h|t}, \pi^i_{t+h|t}-\pi^i_{t+h|t-1}-\gamma_h(s_t-\pi^i_{t-1|t-1})-\eta_h\pi^i_{t+h|t-1}\right) \\
&= \rho^{2h}\mathbb{C}\mathrm{ov}(FE^i_t, FR^i_t) - \rho^{h+1}\gamma_h(1-\kappa_x-\kappa_y)\mathbb{V}\mathrm{ar}(FE^i_{t-1}) + \rho^{h+1}\gamma_h\kappa_y\sigma_\nu^2 \\
&\quad - \rho^{2(h+1)}\eta_h(1-\kappa_x-\kappa_y)E\left[(\pi_{t-1}-\pi^i_{t-1|t-1})\pi^i_{t-1|t-1}\right].
\end{aligned} \tag{C.26}
$$

Note that the numerator of β_1^h and the numerator of β_2^h sum up to $\rho^{2h}\mathbb{C}\mathrm{ov}(FE^i_t, FR^i_t)$.

Appendix C.6 Compute regression coefficients on lagged belief and news

We compute coefficients of regressing forecast errors on lagged beliefs and news. Consider the regression model (2.7):

$$
\pi_{t+h}-\pi^i_{t+h|t} = \alpha_1^h(s_t-\pi^i_{t-1|t-1}) + \alpha_2^h\pi^i_{t+h|t-1} + \beta_2^h \times \mathrm{Residual}^i_{t,h} + \nu^i_{t,h}. \tag{C.27}
$$

Note that by construction, $\mathrm{Residual}^i_{t,h}$ is orthogonal to the new data-release information $(s_t - \pi^i_{t-1|t-1})$ and the prior $(\pi^i_{t+h|t-1})$. The derivation is as follows.

$$
\begin{aligned}
\begin{pmatrix}\alpha_1^h \\ \alpha_2^h\end{pmatrix} &= \begin{pmatrix} \mathbb{E}(s_t-\pi^i_{t-1|t-1})^2 & \mathbb{E}\left[(s_t-\pi^i_{t-1|t-1})\pi^i_{t+h|t-1}\right] \\ \mathbb{E}\left[(s_t-\pi^i_{t-1|t-1})\pi^i_{t+h|t-1}\right] & \mathbb{E}(\pi^i_{t+h|t-1})^2 \end{pmatrix}^{-1} \\
&\quad \mathbb{E}\left(\begin{pmatrix} s_t-\pi^i_{t-1|t-1} \\ \pi^i_{t+h|t-1}\end{pmatrix}(\pi_{t+h}-\pi^i_{t+h|t})\right) \\
&= \left(\mathbb{E}(s_t-\pi^i_{t-1|t-1})^2\mathbb{E}(\pi^i_{t+h|t-1})^2 - \left(\mathbb{E}\left[(s_t-\pi^i_{t-1|t-1})\pi^i_{t+h|t-1}\right]\right)^2\right)^{-1}
\end{aligned}
$$

$$
\begin{pmatrix}
\mathbb{E}(\pi^i_{t+h|t-1})^2 & -\mathbb{E}\left[(s_t-\pi^i_{t-1|t-1})\pi^i_{t+h|t-1}\right] \\
-\mathbb{E}\left[(s_t-\pi^i_{t-1|t-1})\pi^i_{t+h|t-1}\right] & \mathbb{E}(s_t-\pi^i_{t-1|t-1})^2
\end{pmatrix}
$$
$$
\mathbb{E}\left[\begin{pmatrix} s_t-\pi^i_{t-1|t-1} \\ \pi^i_{t+h|t-1}\end{pmatrix}(\pi_{t+h}-\pi^i_{t+h|t})\right]
$$

Note that the denominator is equivalent to Eqn. (C.20) and we omit the derivation here. Next, define the first and second elements of the numerator as $\mathcal{N}^\alpha_{1,h}$ and $\mathcal{N}^\alpha_{2,h}$,

$$
\begin{aligned}
\mathcal{N}^\alpha_{1,h} \equiv\ & \mathbb{E}(\pi^i_{t+h|t-1})^2 E\left[(s_t-\pi^i_{t-1|t-1})(\pi_{t+h}-\pi^i_{t+h|t})\right] \\
& -\mathbb{E}\left[(s_t-\pi^i_{t-1|t-1})\pi^i_{t+h|t-1}\right]\mathbb{E}\left[\pi^i_{t+h|t-1}(\pi_{t+h}-\pi^i_{t+h|t})\right] \\
=\ & \rho^{2h+2}\mathbb{E}(\pi^i_{t-1|t-1})^2\mathbb{E}\left[(s_t-\pi^i_{t-1|t-1})(\rho^{h+1}(1-\kappa_x-\kappa_y)(\pi_{t-1}-\pi^i_{t-1|t-1})-\rho^{h+1}\kappa_y\nu_{t-1})\right] \\
& -\rho^{2h+2}\mathbb{E}\left[(s_t-\pi^i_{t-1|t-1})\pi^i_{t-1|t-1}\right]\mathbb{E}\left[\pi^i_{t-1|t-1}(\rho^{h+1}(1-\kappa_x-\kappa_y)(\pi_{t-1}-\pi^i_{t-1|t-1}))\right] \\
=\ & \rho^{3(h+1)}\mathbb{E}(\pi^i_{t-1|t-1})^2\left[(1-\kappa_x-\kappa_y)\mathbb{E}\left[(\pi_{t-1}-\pi^i_{t-1|t-1})^2\right]-\kappa_y\sigma^2_\nu\right] \\
& -\rho^{3(h+1)}(1-\kappa_x-\kappa_y)\left(\mathbb{E}\left[(\pi_{t-1}-\pi^i_{t-1|t-1})\pi^i_{t-1|t-1}\right]\right)^2 \\
=\ & \rho^{3(h+1)}(1-\kappa_x-\kappa_y)\left[\mathbb{E}(\pi^i_{t-1|t-1})^2\left(\mathbb{V}\mathrm{ar}(FE^i)-\sigma^2_\tau\right)-\left(\mathbb{E}\left[(\pi_{t-1}-\pi^i_{t-1|t-1})\pi^i_{t-1|t-1}\right]\right)^2\right],
\end{aligned}
\tag{C.28}
$$

where the last equality follows the derivation in Eqn. (C.24).

$$
\begin{aligned}
\mathcal{N}^\alpha_{2,h} \equiv\ & -\mathbb{E}\left[(s_t-\pi^i_{t-1|t-1})\pi^i_{t+h|t-1}\right]\mathbb{E}\left[(s_t-\pi^i_{t-1|t-1})(\pi_{t+h}-\pi^i_{t+h|t})\right] \\
& +\mathbb{E}\left[\pi^i_{t+h|t-1}(\pi_{t+h}-\pi^i_{t+h|t})\right]\mathbb{E}(s_t-\pi^i_{t-1|t-1})^2 \\
=\ & -\mathbb{E}\left[(s_t-\pi^i_{t-1|t-1})\pi^i_{t+h|t-1}\right]\mathbb{E}\left[(s_t-\pi^i_{t-1|t-1})(\rho^{h+1}(1-\kappa_x-\kappa_y)(\pi_{t-1}-\pi^i_{t-1|t-1})-\rho^{h+1}\kappa_y\nu_{t-1})\right] \\
& +\mathbb{E}\left[\pi^i_{t+h|t-1}(\rho^{h+1}(1-\kappa_x-\kappa_y)(\pi_{t-1}-\pi^i_{t-1|t-1}))\right]\mathbb{E}(s_t-\pi^i_{t-1|t-1})^2 \\
=\ & -\rho^{2h+2}(1-\kappa_x-\kappa_y)\mathbb{E}\left[(\pi_{t-1}-\pi^i_{t-1|t-1})\pi^i_{t-1|t-1}\right]\mathbb{E}\left[(\pi_{t-1}-\pi^i_{t-1|t-1})^2\right] \\
& +\rho^{2h+2}\mathbb{E}\left[(\pi_{t-1}-\pi^i_{t-1|t-1})\pi^i_{t-1|t-1}\right]\kappa_y\sigma^2_\nu \\
& +\rho^{2h+2}(1-\kappa_x-\kappa_y)\mathbb{E}\left[\pi^i_{t-1|t-1}(\pi_{t-1}-\pi^i_{t-1|t-1})\right]\mathbb{E}(\pi_{t-1}-\pi^i_{t-1|t-1})^2 \\
& +\rho^{2h+2}(1-\kappa_x-\kappa_y)\mathbb{E}\left[\pi^i_{t-1|t-1}(\pi_{t-1}-\pi^i_{t-1|t-1})\right]\sigma^2_\nu \\
=\ & \rho^{2h+2}\mathbb{E}\left[(\pi_{t-1}-\pi^i_{t-1|t-1})\pi^i_{t-1|t-1}\right](1-\kappa_x)\sigma^2_\nu.
\end{aligned}
\tag{C.29}
$$

Thus, $\alpha^h_1=\frac{\mathcal{N}^\alpha_{1,h}}{\mathcal{D}_h}$ and $\alpha^h_2=\frac{\mathcal{N}^\alpha_{2,h}}{\mathcal{D}_h}$.

Appendix D Proof of propositions

Appendix D.1 Proof of Proposition 1

Proof. Under RE, $\mathbb{V}\mathrm{ar}(FE^i) = \sigma_\tau^2$. Moreover, $\mathbb{E}\left[(\pi_{t-1} - \pi^i_{t-1|t-1})\pi^i_{t-1|t-1}\right] = 0$ since forecast errors $(\pi_{t-1} - \pi^i_{t-1|t-1})$ are not predictable by variables in forecaster i's information set at period $t-1$, and are therefore, orthogonal to the forecasts $(\pi^i_{t-1|t-1})$. We have the following:

1. The sign of β_1^h follows the sign of $\mathcal{N}^\beta_{1,h}$ (Eqn. C.22 and C.23). According to Eqn. (C.24) and Eqn. (C.25), $\beta_1^h = 0$ under RE.

2. The sign of β_2^h follows the sign of $\mathcal{N}^\beta_{2,h}$ (Eqn. C.26). Since $\mathcal{N}^\beta_{1,h} + \mathcal{N}^\beta_{2,h} \propto \mathbb{C}\mathrm{ov}(FE^i_t, FR^i_t) = 0$ under RE, given that $\beta_1^h = 0$, $\beta_2^h = 0$ under RE.

3. The sign of α_1^h follows the sign of $\mathcal{N}^\alpha_{1,h}$ (Eqn. C.28), which always equals 0 under RE.

4. The sign of α_2^h follows the sign of $\mathcal{N}^\alpha_{2,h}$ (Eqn. C.29). According to Eqn. (C.29), $\alpha_2^h = 0$ under RE.

□

Appendix D.2 Proof of Proposition 2

Proof. First, Eqn. (3.6) yields

$$\sigma_\nu^2 = \frac{1-\omega}{\omega}\sigma_\tau^2 \tag{D.1}$$

Eliminating σ_ν^2 from Eqn. (3.7) and using $\widehat{\sigma}_\tau^2 = \tau\sigma_\tau^2$, we get

$$\sigma_\epsilon^2 = \frac{(1-\kappa_x)\left(\rho^2(1-\omega)\sigma_\tau^2 + \sigma_u^2\right)}{\kappa_x \tau}. \tag{D.2}$$

Second, substituting Eqn. (D.1) and (D.2) into (3.10) and solve for σ_τ, we obtain

$$\sigma_\tau^2 = \frac{(1-\kappa_x)}{1-\rho^2(1-\kappa_x)(1-\omega)}\sigma_u^2. \tag{D.3}$$

Under overconfidence of private information, $\mathbb{V}\mathrm{ar}(FE^i) > \widehat{\sigma}_\tau^2$. Therefore, from Eqn. (C.7) and (D.3), we get

$$\mathbb{E}\left[(\pi_{t-1} - \pi^i_{t-1|t-1})\pi^i_{t-1|t-1}\right] = \widehat{\sigma}_\tau^2 - \mathbb{V}\mathrm{ar}(FE^i) < 0. \tag{D.4}$$

Eqn. (C.6) yields $E(\pi_t^2) - E(\pi_t \pi_{t|t}^i) = \widehat{\sigma}_\tau^2$, which in turn leads to

$$E(\pi_t \pi_{t|t}^i) = \frac{\sigma_u^2}{1-\rho^2} - \widehat{\sigma}_\tau^2 \tag{D.5}$$

Eqn. (C.7) gives

$$E(\pi_t \pi_{t|t}^i) - E\left((\pi_{t|t}^i)^2\right) = \widehat{\sigma}_\tau^2 - \mathbb{V}\mathrm{ar}(FE^i) \tag{D.6}$$

Combining Eqn. (D.5) and (D.6), we get

$$E\left((\pi_{t|t}^i)^2\right) = \mathbb{V}\mathrm{ar}(FE^i) - 2\sigma_\tau^2 + \frac{\sigma_u^2}{1-\rho^2} \tag{D.7}$$

Before continuing the proof of this proposition, we note that the individual-level CG coefficient is negative under overconfidence ($\beta_h^p < 0$) due to the inequality (D.4) and Eqn. (C.11). We now have the following:

1. The sign of β_1^h follows the sign of $\mathcal{N}_{1,h}^\beta$ (Eqn. C.23). The sum of the first two components of $\mathcal{N}_{1,h}^\beta$ (Eqn. C.22) is (C.24), which is positive because $\mathbb{V}\mathrm{ar}(FE^i) > \widehat{\sigma}_\tau^2$. The third component of $\mathcal{N}_{1,h}^\beta$ in Eqn. (C.23) is given by Eqn. (C.25), which is positive too. Thus, $\beta_1^h > 0$.

2. The sign of β_2^h follows the sign of $\mathcal{N}_{2,h}^\beta$ (Eqn. C.26). Since $\mathcal{N}_{1,h}^\beta + \mathcal{N}_{2,h}^\beta \propto \mathbb{C}\mathrm{ov}(FE_t^i, FR_t^i) < 0$ under overconfidence, given that $\beta_1^h > 0$, it follows that $\beta_2^h < 0$.

3. The sign of α_1^h follows the sign of $\mathcal{N}_{1,h}^\alpha$ (Eqn. C.28).

$$\begin{aligned}
\mathcal{N}_{1,h}^\alpha \propto\ & (1-\kappa_x-\kappa_y)\left[\mathbb{E}(\pi_{t-1|t-1}^i)^2\left(\mathbb{V}\mathrm{ar}(FE^i)-\sigma_\tau^2\right) - \left(\mathbb{E}\left[(\pi_{t-1}-\pi_{t-1|t-1}^i)\pi_{t-1|t-1}^i\right]\right)^2\right] \\
&= (1-\kappa_x-\kappa_y)\left[\left(\mathbb{V}\mathrm{ar}(FE^i)-2\sigma_\tau^2+\frac{\sigma_u^2}{1-\rho^2}\right)\left(\mathbb{V}\mathrm{ar}(FE^i)-\sigma_\tau^2\right) - \left(\mathbb{V}\mathrm{ar}(FE^i)-\sigma_\tau^2\right)^2\right] \\
&= (1-\kappa_x-\kappa_y)\left(\mathbb{V}\mathrm{ar}(FE^i)-\sigma_\tau^2\right)\left(\frac{\sigma_u^2}{1-\rho^2}-\sigma_\tau^2\right)
\end{aligned}$$

   The second equation above uses Eqn. (D.7) and (D.6). Note that $1-\kappa_x-\kappa_y \geq 0$ with equality when $\tau = 0$; $\mathbb{V}\mathrm{ar}(FE^i) - \widehat{\sigma}_\tau^2 > 0$ under overconfidence; since $\mathbb{V}\mathrm{ar}(\pi_t) = \frac{\sigma_u^2}{1-\rho^2}$ is the unconditional variance of π_t, $\sigma_\tau^2 < \frac{\sigma_u^2}{1-\rho^2}$ always holds when σ_v and σ_ϵ are finite. Therefore, $\alpha_1^h > 0$ when $\tau \in (0,1)$.

4. The sign of α_2^h follows the sign of $\mathcal{N}_{2,h}^\alpha$ (Eqn. C.29). Because of the inequality (D.4), Eqn. (C.29) implies $\alpha_2^h < 0$ under overconfidence.