Overconfidence in Private Information Explains Biases in Professional Forecasts[*]

Klaus Adam[†]
University College London & CEPR

Pei Kuang[‡]
University of Birmingham

Shihan Xie[§]
University of Illinois Urbana-Champaign

October 4, 2024

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 find that SPF forecasters overreact to residual information at almost all forecast horizons and for almost all forecast variables. In addition, forecasts are overly anchored to prior beliefs for all variables at all forecast horizons. We show analytically that overconfidence in private information qualitatively generates both of these features. It also implies that forecast errors correlate positively with past forecast revisions at the consensus level, but negatively at the individual level, as documented previously in the literature. Estimating Bayesian updating models on SPF data, we show that overconfidence in private information also replicates the observed patterns quantitatively. All estimated models display strong and statistically significant overconfidence in private information.

JEL classification: C53, D83, D84, E31
Keywords: expectations, information, belief updating, professional forecasts

[*]We would like to thank Marios Angeletos for helpful comments and suggestions. We are deeply grateful to the people who helped us determining the information provided to professional forecasters when surveys are collected, including Tom C. Stark (Philadelphia Fed, Survey of Professional Forecasters), Fatima Mboup (Philadelphia Fed, Livingston Survey), Steven Hubbard (Consensus Economics Inc.), and Aidan Meyler (European Central Bank).

[†]Department of Economics, University College London, United Kingdom. Email: k.adam@ucl.ac.uk
[‡]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 the assumption of full-information rational expectations (FIRE) has been the dominant workhorse assumption on expectation formation in macroeconomics. In a seminal paper, Lucas (1972) relaxed the FIRE assumption and studied expectation formation in a setup with incomplete information. Subsequently, macroeconomists continued studying models featuring learning, private information, and information frictions, e.g., Marcet and Sargent (1989); Woodford (2002); Mankiw et al. (2003); Sims (2003).

A key difficulty with testing the forecast implications of models featuring deviations from full information is that the information set available to forecasters can typically not be observed. This creates challenges for studying the efficiency properties of survey forecasts and for building empirically credible private information models that can be used in quantitative applications. To deal with this issue, Coibion and Gorodnichenko (2015) proposed using past forecasts as measures of the information available to agents. Using this approach, they showed that professional forecasts underreact to past forecast revisions 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 the existence of deviations from FIRE, they offer only indirect evidence on 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 *available to forecasters* at the time of forecasting and the present paper makes progress on this front.

Going back to the survey forms that get administered when collecting forecasts in the U.S. Survey of Professional Forecasters (SPF), we find that SPF forecasters are provided with the most recent data release of the variables they are requested to forecast in every forecasting round.[1] To the extent that this fact is common knowledge among forecasters, the latest data release represents public information that forecasters receive in between two forecasting rounds.

[1]We also show that this is a general feature of professional surveys: the Livingston Survey, the surveys run by Consensus Economics, and the European Central Bank's SPF all provide forecasters with the latest data release of the variables they are asked to forecast.

And since we observe forecasters' prior expectations about the newly released variables in the previous forecasting round, we can construct a high-dimensional measure of public *news* received by every forecaster.[2] Due to the heterogeneity in forecasters' prior expectations, the *news* contained in public information differs across forecasters.

In a first step, we use these forecaster-specific measures of public news to estimate how individual forecast revisions about macroeconomic variables over time depend on (i) public news, (ii) forecasters' prior beliefs, and (iii) a residual capturing information that is contained neither in the prior nor in the public news. In a second step, we regress individual ex-post forecast errors on forecast revisions explained by (i) public news, (ii) prior expectations, and (iii) the residual component.

With rational expectations, information used by forecasters to revise expectations does not predict forecast errors. Therefore, rational expectations implies that (i)-(iii) will not predict forecast errors. This holds independently of whether forecasters possess full information or not. We show, however, that this condition is strongly violated in the SPF data:

1. Forecasters' expectations are overly anchored to their prior expectations (ii). This holds true for all forecast variables and all forecast horizons in the survey.

2. Forecast revisions overreact to the residual component (iii). This holds true for the vast majority of forecast horizons and forecast variables.

3. Forecasters underreact to public news for the majority of variables and forecast horizons, but for a number of variables the opposite holds true.

While the first two findings are new to the literature, the last finding is broadly in line with evidence provided in Broer and Kohlhas (2022).

Matching this evidence requires both a deviation from full information and a deviation from rational expectations. Specifically, we show analytically that a simple Bayesian updating model featuring private and public information sources can qualitatively explain the three facts listed above, provided forecasters display overconfidence in the information content of their private

[2]As we explain in the main text, this is not possible for the other professional surveys mentioned in the preceding footnote.

information signal, in the sense that they *underestimate the noise contained in private information.* Importantly, the updating model with overconfidence in private information also generates underreaction to past belief revisions at the consensus level (Coibion and Gorodnichenko (2015)) and overreaction at the individual level (Bordalo et al. (2020)).

The Bayesian updating model replicates these facts because overconfidence in private information causes overreaction to private news. Since private news is reflected in the residual (iii), the model replicates overreaction to the residual (point 2. above). Overconfidence also implies that the prior expectations are viewed as more informative than they actually are, due to the accumulation of "informative" past private signals. This causes expectations to be overly anchored to prior expectations (point 1. above). And with the information content of the prior and of the private signal being overestimated by forecasters, public news tends to receive too little weight in updating (point 3. above).

Overconfidence in private information and the resulting overreaction to private information also causes overreaction of individual forecasts to forecast revisions (Bordalo et al. (2020)). Finally, the presence of private information causes underreaction of consensus beliefs to consensus revisions, as is the case with rational expectations (Coibion and Gorodnichenko (2015)).

Having shown that overconfidence in private information *qualitatively* generates the observed patterns in SPF forecasts, we turn consideration to the question whether the proposed model can also *quantitatively* match the evidence. To this end, we use the simulated method of moments to estimate Bayesian belief updating models that allow for overconfidence in private information. We show that a simple updating model quantitatively replicates a wide range of data moments surprisingly well, including the evidence listed in points 1-3 above. The estimated models robustly imply large and statistically significant amounts of overconfidence in private information: forecasters underestimate the standard deviation of the noise contained in private information by a factor of 2 to 5.

Taken together, our findings show that overconfidence in private information is a belief distortion that can single-handedly replicate a wide range of empirically documented deviations from FIRE in the SPF. While we do not rule out alternative explanations exist, we present additional evidence that further strengthens the case that private information is at the heart of the observed deviations from rational expectations.

In particular, our Bayesian updating model implies that the residual (iii) reflects forecasters' private information. This gives rise to additional testable implications: individual forecasts errors should fall, if forecasters based belief revisions on the average private signal rather than on their own private signal, because the average private signal purges some of the idiosyncratic noise contained in the individual signal. Conversely, replacing the private signal by the idiosyncratic component of the private signal should increase forecast errors. We test these predictions and find strong support for it in the SPF data, which further strengthens the case for overconfidence in private information.

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

In particular, Broer and Kohlhas (2022) document overreaction and underreaction patterns 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. The approach in these papers differ from ours because they assume that public information consists of past consensus forecasts. We treat the most recent data release as public information, in line with the information provided to forecasters on the SPF survey questionnaire.

Angeletos et al. (2021) provide 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. Yet, in line with their findings, our finding that forecasters' expectations are overly anchored to past beliefs implies (on average across shocks) underreaction to economic shocks in the impact period.

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 paper contributes by disciplining deviations from FIRE using information about a broad range of public signals available to forecasters.

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 qualitatively replicate all these facts. In Section 4, we present our estimated updating model and document that it performs surprisingly well in quantitatively replicating the empirical evidence and implies large and statistically significant degrees of overreaction to private information. Section 5 concludes.

2 New evidence on the source of forecast errors

This section explains how we identify the public information flow received by SPF forecasters in between survey rounds. Using the identified public information and forecasters' prior expectations, we compute the news contained in public information. We then decompose individual macroeconomic forecast revisions about the same variable in the same time period between two survey rounds into revisions that are due to (i) public news, (ii) prior expectations and (iii) residual information. In a final step, we show how individual ex-post forecast errors depend on these three 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

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

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).

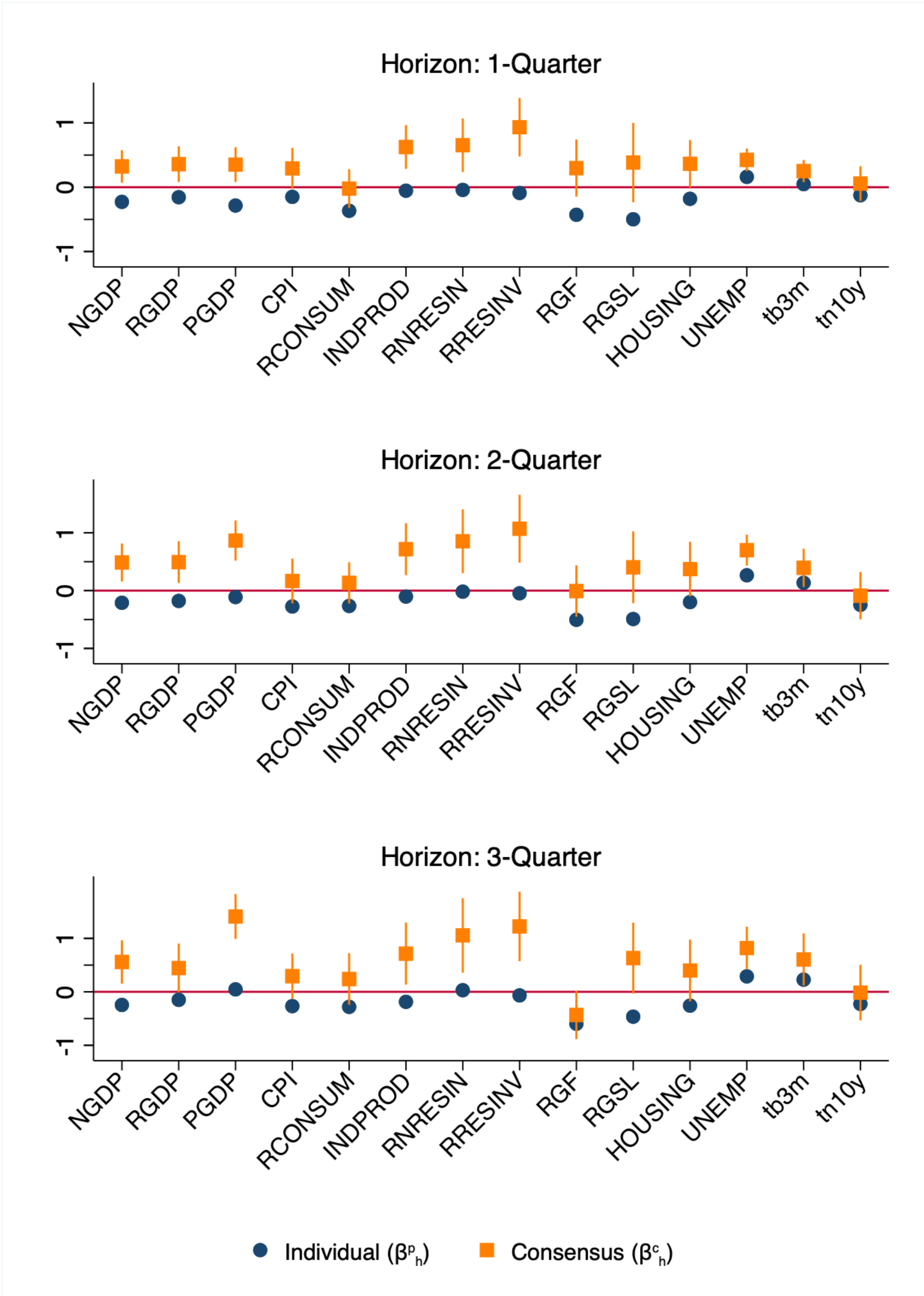


Figure 1: RESPONSES OF FORECAST ERRORS TO FORECAST REVISIONS AT THE CONSENSUS AND INDIVIDUAL LEVEL

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

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

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

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.

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 2 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. On these survey forms it is stated 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

SPF 2014:Q1

Forecaster:
Date:

Section 1. U.S. Business Indicators

		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.

Figure 2: SAMPLE SPF SURVEY FORM

survey management team. Figure A.1 in Appendix A.1 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 found out 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.2 provides a detailed discussion of the information available to forecasters participating in these surveys.

While supplying professional forecasters with the latest data release appears to be common practice in the administration of surveys, the decomposition exercise we implement below can only be performed with for the SPF forecast: the SPF is the only survey that includes in every round forecasts for four consecutive quarters, so that we observed how forecast for the same *variable* and same *time period* gets revised over time. Other surveys ask for forecasts only for a given longer horizon (usually one year or longer) or ask forecasters to forecast a fixed calendar year. As we explain below, the availability of successive rounds of forecasts over time for the same variable in the same quarter is key for our approach.

2.4 Decomposing forecast revisions and their effects on forecast errors

This section decomposes individual forecast revisions into revisions associated with public news and residual 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.[3]

Consider the second month of quarter t, which is the month in which forecasts are collected. Let $s_t \in R^{14}$ denote the vector of public information presented to the forecasters, which consists of the latest data releases that came out between the second month in the last quarter and the second month in the current quarter. Letting $s^i_{t|t-1}$ denote forecaster i's forecast of 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 heterogeneous prior beliefs and 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 are interested in how this forecast gets *revised* from one quarter to the next, i.e., we are interested in $\pi^i_{t+h|t} - \pi^i_{t+h|t-1}$.

[3]The latest data release is public information, provided it is common knowledge that the latest release is on every forecaster's survey sheet, as is reasonable to assume.

Linear normal Bayesian updating implies that the forecast revision is a linear function of public news, $s_t - s^i_{t|t-1}$, prior beliefs $\pi^i_{t+h|t-1}$, and residual news that is not contained in public news. In particular, we can regress (for $h = 1,2,3$) the observed forecast revision on observed public news and the observed prior expectations:

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

where $\bar{\delta}^i_h$ 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 $\eta_h \in R^{14}$ captures the rate at which the weight on past information is reduced due to incoming news and the operator "$\circ$" indicates element-wise multiplication between vectors (Hadamard product). When agents follow Bayesian updating, we have $-1 \leq \eta_h \leq 0$, with the limiting cases $\eta_h = 0$ indicating the arrival of no new information and $\eta_h = -1$ indicating that new information is infinitely more informative than the information contained in the prior.[4] Figure 3 plots the η_h coefficients for all considered variables and forecast horizons. The vast majority of point estimates lie in the predicted range.

Note that equation (2.3) decomposes forecast revisions into those due to (i) a vector of public news, (ii) prior information becoming less relevant and (iii) a residual component $\epsilon^i_{h,t}$. If the public information signal s_t exhausts the set of public information, then the residual vector $\epsilon^i_{h,t}$ in equation (2.3) captures forecasts revisions that are due to forecasters' private news. Otherwise, the residual contains revisions that are due to a mix of unobserved public news and private news. Since the dynamics of macroeconomic variables can typically be described as being driven by less than a handful of common factors, see for instance Stock and Watson (2016), our 14 public signals represent - by macroeconomic standards - a high-dimensional public signal. This suggests that $\epsilon^i_{h,t}$ should predominantly reflects private information. We provide below empirical evidence supporting this view.

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 residual

[4]Inequalities involving vectors should be interpreted as applying to each element in the vector.

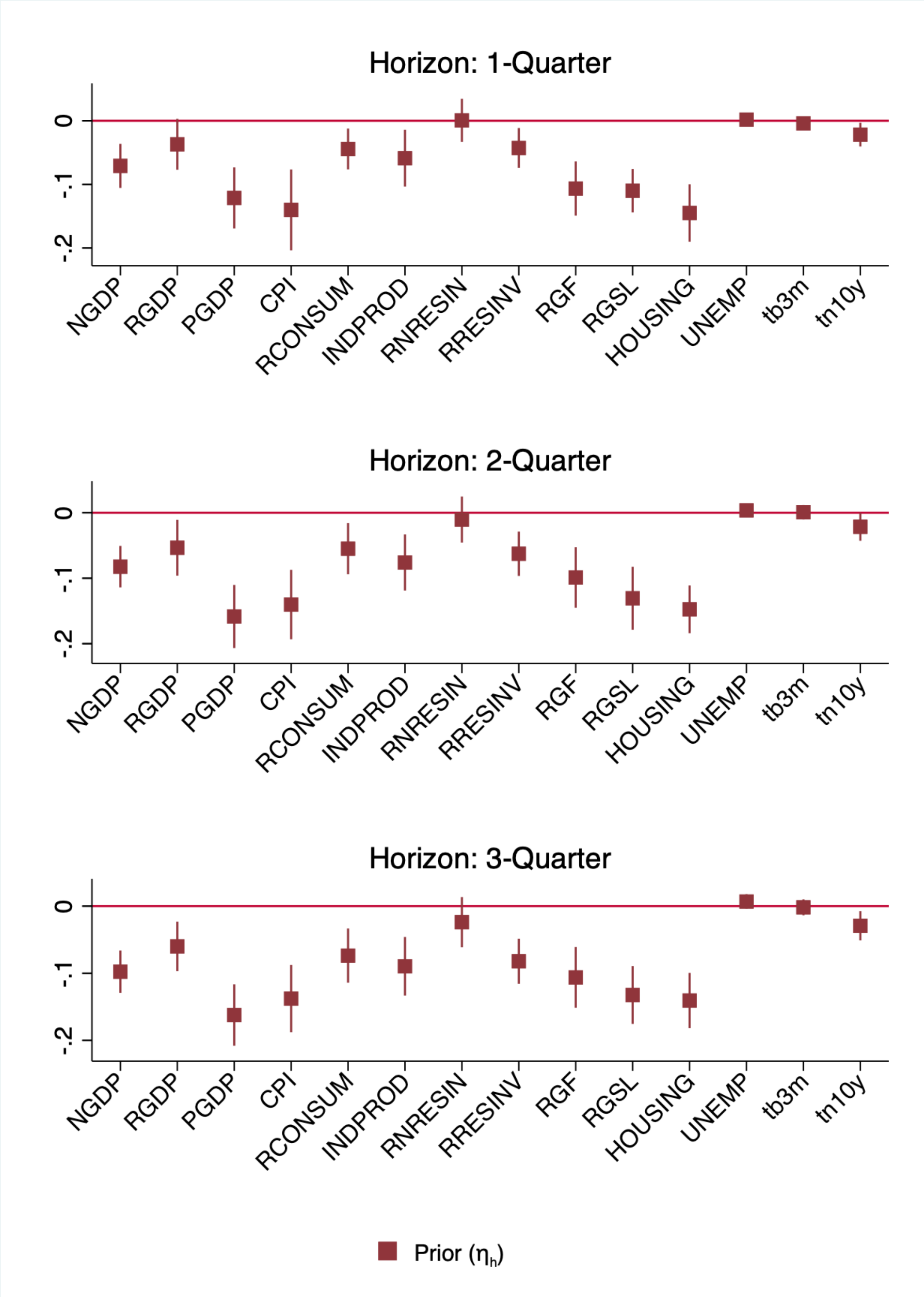


Figure 3: RESPONSES OF FORECAST REVISIONS TO PRIOR BELIEFS

Notes: This figure plots the coefficients of η_h on prior beliefs from Eqn. (2.3). 95% confidence intervals based on clustered standard errors are reported.

information, i.e., the regression residual:

$$\text{Predicted}^i_{h,t} \equiv \hat{\gamma}_h (s_t - s^i_{t|t-1}) + \hat{\eta}_h \circ \pi^i_{t+h|t-1}, \tag{2.4}$$

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

We then investigate whether these components predict individual forecast errors by considering regressions of the form

$$\pi_{t+h} - \pi^i_{t+h|t} = \bar{\bar{\delta}}^i_h + \beta_{1,h} \circ \text{Predicted}^i_{h,t} + \beta_{2,h} \circ \text{Residual}^i_{h,t} + \nu^i_{h,t}, \tag{2.6}$$

where the coefficient vectors $\beta_{i,h} \in R^{14}$ for $i = 1,2$ and the operator $\circ$ again indicates element-wise multiplication between vectors (Hadamard product). When forecasters hold rational expectations, we have $\beta_{1,h} = \beta_{2,h} = 0$ because the two regressors on the r.h.s. of the previous equation both reflect information that is available to forecasters at the time of forecasting.

Figure 4 reports the OLS estimates of $\beta_{1,h}$ (in green) and $\beta_{2,h}$ (in orange) for all considered variables and forecast horizons. It shows that these coefficients often significantly deviate from zero.[5] They also display a rather coherent pattern: for almost all variables and forecasting horizons, macroeconomic expectations underreact to forecast revisions induced by the prior and public news ($\beta_{1,h} > 0$). In addition, they overreact to the residual news component ($\beta_{2,h} < 0$).

We summarize these empirical findings as follows:

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 forces giving rise to Fact 1. To this end, we decompose the predicted component of forecast revisions constructed above into its two sub-components, i.e., 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 residual news from the regression (2.3). To do so, we

[5]Since our null hypotheses are $\beta_{1,h} = 0$ and $\beta_{2,h} = 0$, the standard errors do not have to be adjusted for the fact that our regressors are generated.

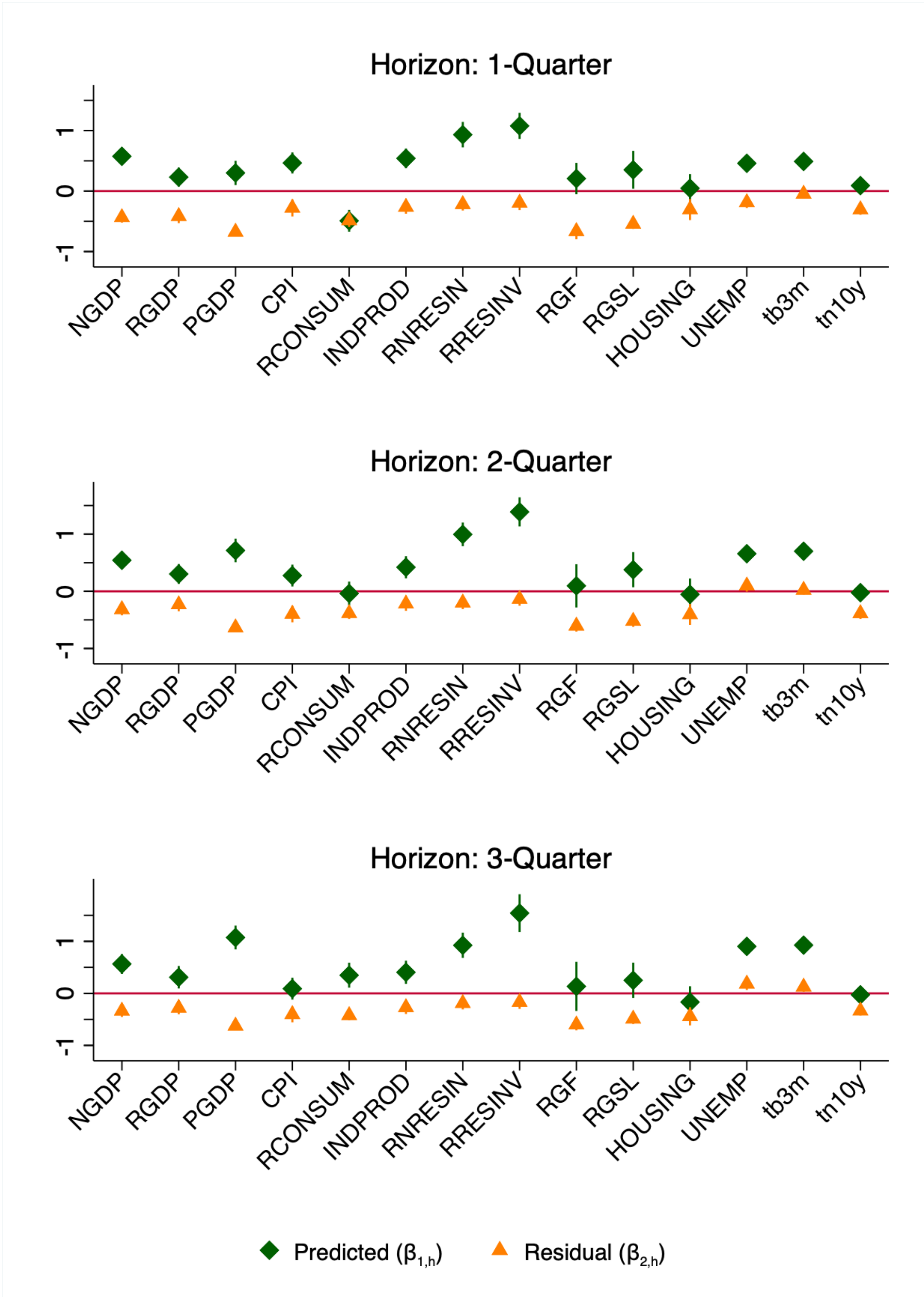


Figure 4: RESPONSES OF FORECAST ERRORS TO PREDICTED AND RESIDUAL COMPONENTS OF FORECAST REVISION

Notes: This figure plots the coefficients of $\beta_{1,h}$ on the predicted component of forecast revisions (in green) and $\beta_{2,h}$ on the residual component (in orange) from Eqn. (2.6). 95% confidence intervals based on clustered standard errors are reported.

define for each forecaster i and each forecast horizon h the forecast revision that is due to public information

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

where $\hat{\gamma}_h$ denotes the estimated coefficient matrix from equation (2.3). We then consider forecast-error regressions of the form

$$\pi_{t+h} - \pi^i_{t+h|t} = \tilde{\delta}^i_h + \alpha_{1,h} \circ \text{Public}^i_{h,t} + \alpha_{2,h} \circ \pi^i_{t+h|t-1} + \beta_{2,h} \circ \text{Residual}^i_{h,t} + \nu^i_{h,t}, \tag{2.7}$$

where $\circ$ again indicates element-wise multiplication. Again, rational expectations implies $\alpha_{1,h} = \alpha_{2,h} = 0$.

Figure 5 plots the OLS estimates of $\alpha_{1,h}$ (blue colour) and $\alpha_{2,h}$ (brown colour) for all considered variables and forecast horizons. It shows that rational expectations are rejected in most cases.[6] Specifically, the results indicate a negative coefficient on the prior expectation ($\alpha_{2,h} < 0$) for all forecast variables and all forecast horizons. Since $\eta_h < 0$ in equation (2.3), this implies that forecasters do not reduce the weight on prior expectations sufficiently strongly. As a result, their expectations remain too strongly anchored to prior beliefs. In addition, Figure 5 shows that forecast errors covary mostly positively with public news ($\alpha_{1,h} > 0$). However, this feature is less consistent across variables and forecast horizons. It nevertheless indicates that forecasters predominantly underreact to public news. Both sub-components thus tend to contribute to the positive coefficient on $\text{Predicted}^i_{h,t}$ documented in Figure 4.

We summarize these empirical findings as follows:

Fact 3: At the individual level, forecasters' expectations mostly underreact to public news ($\alpha_{1,h} > 0$), 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 investigate whether the estimated residual $\hat{\epsilon}^i_{h,t}$ in equation (2.3) displays patterns that are

[6] By construction, the regressor $\text{Residual}^i_{h,t}$ is orthogonal to the news component $(s_t - s^i_{t|t-1})$ and the prior $(\pi^i_{t+h|t-1})$, so that the estimate of $\beta_{2,h}$ in equation (2.7) will be identical to the one in equation (2.6) and is thus not shown here.

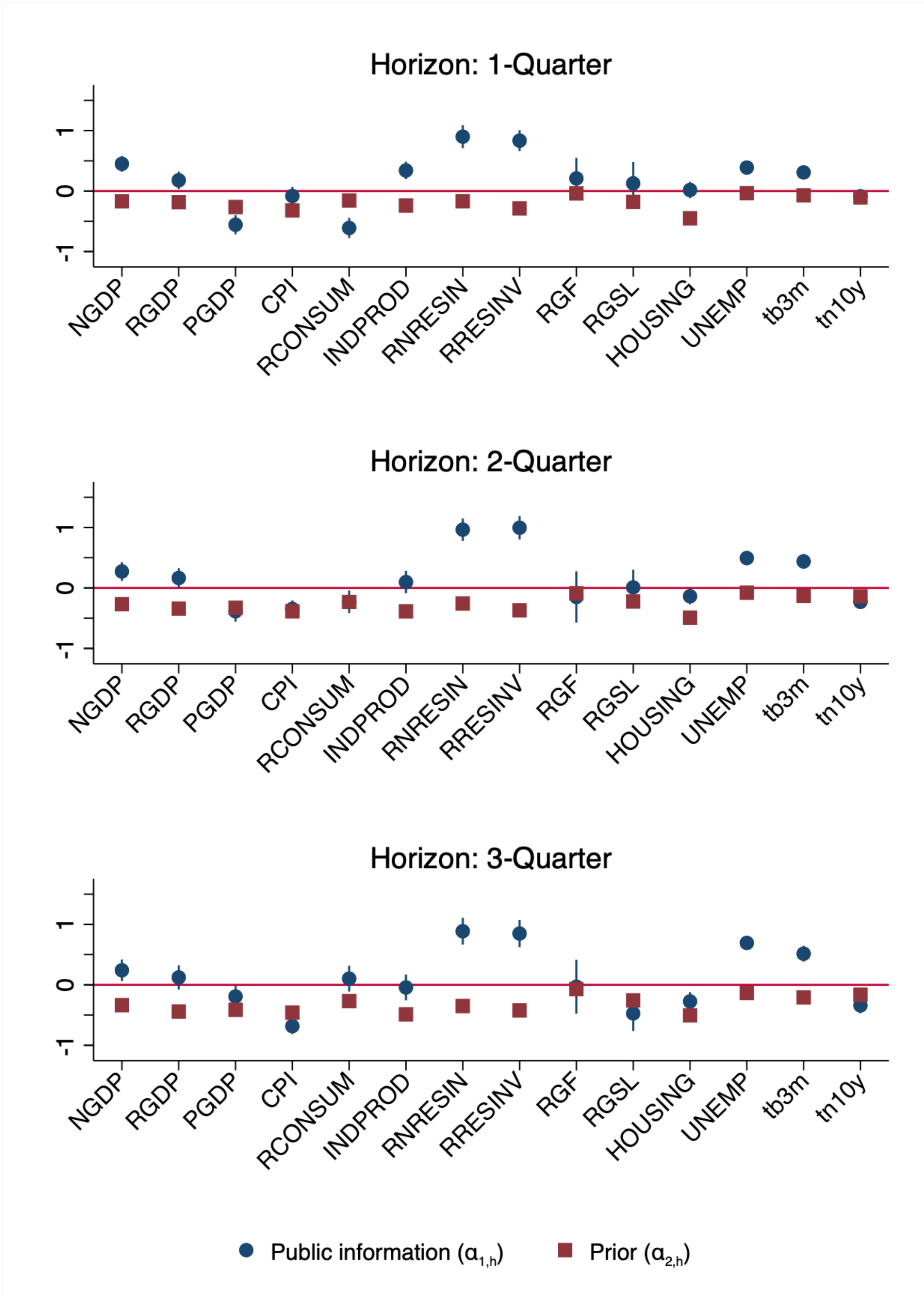


Figure 5: RESPONSES OF FORECAST ERRORS TO PUBLIC NEWS AND PRIOR EXPECTATIONS

Notes: This figure plots the estimated coefficients of $\alpha_{1,h}$ (in blue) and $\alpha_{2,h}$ (in brown) from Eqn. (2.7). 95% confidence intervals based on clustered standard errors are reported.

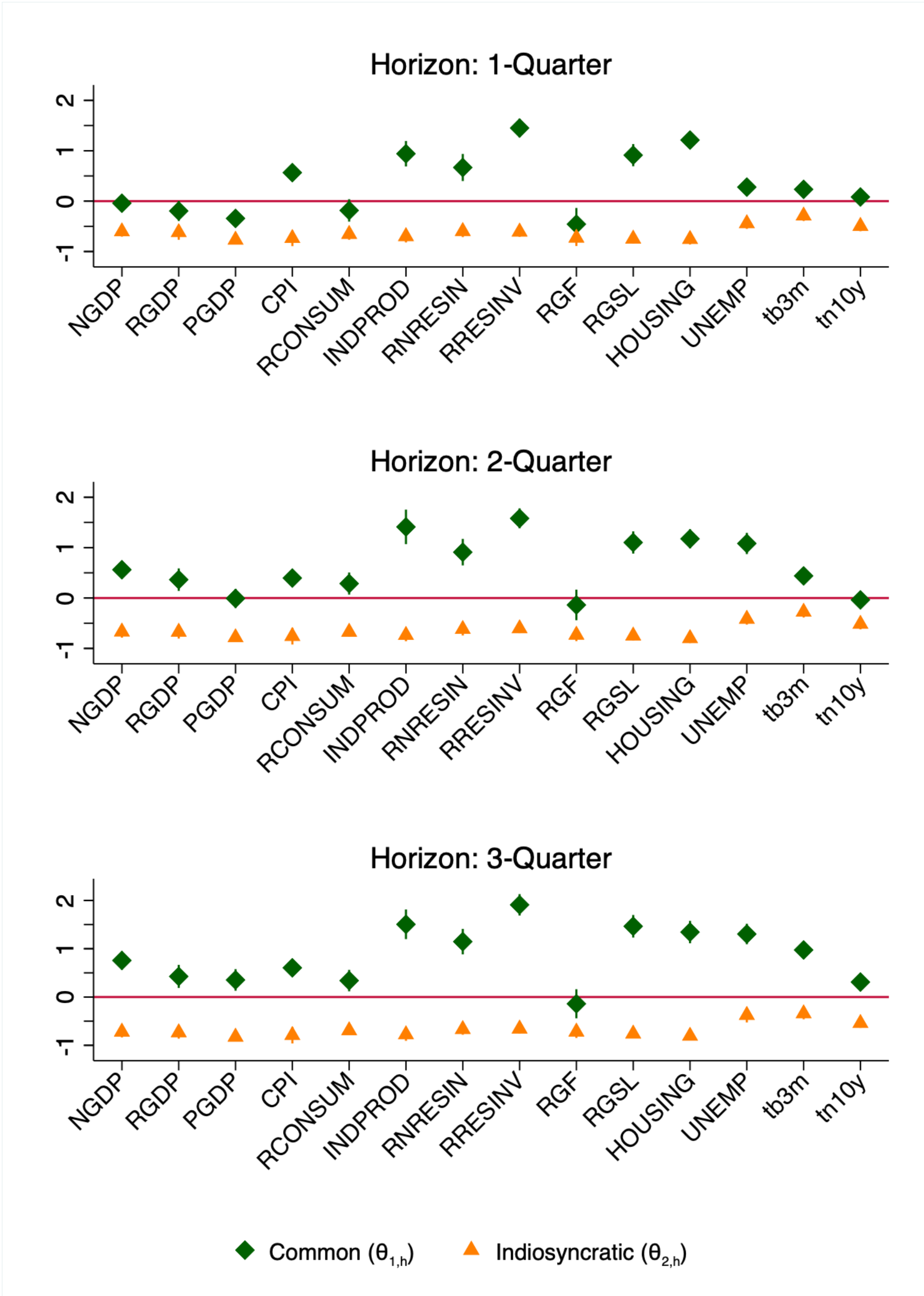


Figure 6: RESPONSES OF FORECAST ERRORS TO COMMON AND IDIOSYNCRATIC COMPONENTS OF PRIVATE INFORMATION

Notes: This figure plots the estimated coefficients of $\theta_{1,h}$ (in green) and $\theta_{2,h}$ (in orange) from Eqn. (2.10). 95% confidence intervals based on clustered standard errors are reported.

consistent with these belief revisions being due to the presence of (noisy) private information. To this end, we decompose residual forecast revisions (at a given point in time) into a common and an idiosyncratic component

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

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

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

$$\begin{aligned} \pi_{t+h} - \pi^i_{t+h|t} =& \tilde{\tilde{\delta}}^h_i + \alpha_{1,h} \circ \text{Public}^i_{h,t} + \alpha_{2,h} \circ \pi^i_{t+h|t-1} \\ &+ \theta_{1,h} \circ \text{Common}^i_{h,t} + \theta_{2,h} \circ \text{Idiosync}^i_{h,t} + \nu^i_{h,t}. \end{aligned} \tag{2.10}$$

Figure 6 plots the OLS estimates of $\theta_{1,h}$ (in green) and $\theta_{2,h}$ (in orange). It shows that the idiosyncratic component of the residual has a negative coefficient ($\theta_{2,h} < 0$) for all variables and all horizons, while the coefficient on the common component is generally positive ($\theta_{1,h} > 0$). This pattern is fully consistent with residual forecast revisions being due to private information. Specifically, it shows that if forecasters had access to the private information of other forecasters, they could improve forecast errors by reducing the updating weight on their own idiosyncratic noise component ($\theta_{2,h} < 0$) and by reacting more strongly ($\theta_{1,h} > 0$) to the (less noisy) average private signal than they react to their own (more noisy) private signal. We show in the next section that this is consistent with a situation in which residual information only reflects 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 very similar findings as the ones presented above.

In Appendix B.2 we conduct a robustness analysis by adding news about consensus forecasts to the public news available to forecasters.[7] The inclusion of news about consensus forecasts

[7] Strictly speaking, we do not observe news about consensus forecasts, as agents do not forecast consensus forecast in the SPF. However, we use revisions of consensus forecasts from one quarter to the next, i.e., $\pi^c_{t+h|t-1} - \pi^c_{t+h|t-2}$, as a proxy for public news.

again leads to very similar findings.

3 Explaining the evidence

This section presents a simple Bayesian belief updating 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.

Section 3.1 introduces the updating model, which allows for departures from full information and from rational expectations. Departures from full information take the form of noisy public and private information, while departures from rational expectations take the form of subjective beliefs about the noise variance contained in public and private information. Section 3.2 shows analytically that the model misses nearly all empirical facts mentioned above when forecasters hold rational expectations. Section 3.3 then considers a setting where forecasters are overconfident about their private information, i.e., underestimate the noise contained in private information. It shows analytically that overconfidence in private information allows qualitatively replicating all empirical facts mentioned in the previous section. The quantitative performance of the model with overconfidence will be explored in detail in Section 4.

3.1 The setup

We consider a setting with $i = 1, 2, ..., I$ forecasters that receive private and public signals about an underlying state that drives the realization of observable variables. In line with the empirical analysis in the previous section, public information consists of the most recent data release, while private information provides noisy information about the current value of the underlying state. To be able to derive analytic results, we consider a univariate setting.

In period t, forecasters seek to forecast future releases of the variable $s_{t+h} \in R$ for $h \geq 1$, which evolves according to

$$s_t = \pi_{t-1} + \nu_t, \tag{3.1}$$

where $\pi_{t-1} \in R$ is the unobserved state and $\nu_t \sim_{iid} N(0, \sigma_\nu^2)$ a variable-specific noise component.

The underlying state evolves according to

$$\pi_t = \rho\pi_{t-1} + u_t, \tag{3.2}$$

where $\rho \in (0,1)$ and $u_t \sim_{iid} N(0,\sigma_u^2)$.

In period t, before forecasting s_{t+h} for $h \geq 1$, forecasters observe the realization of the variable of interest s_t from the previous quarter, which is a function of the lagged state. This captures the fact that forecasters observe the lagged outcomes of the variable they seek to forecast. Each forecaster i also receives an idiosyncratic private signal x_t^i about the value of the current state

$$x_t^i = \pi_t + \epsilon_{xt}^i, \tag{3.3}$$

where $\epsilon_{xt}^i \sim_{iid} N(0,\sigma_\epsilon^2)$ is idiosyncratic observation noise.

The information Ω_t^i available to forecaster i in period t consists of all current and lagged values of the outcome variable and of the private signal, i.e., $\Omega_t^i = \left\{s_\tau, x_\tau^i\right\}_{\tau=0}^t$. Given this information, forecaster i formulates expectations about future outcomes $\mathbb{E}^{\mathscr{P}}[s_{t+h}|\Omega_t^i]$ for $h \geq 1$, where $\mathscr{P}$ denotes a potentially subjective probability measure, as described further below. We assume that professional forecasters truthfully report their expectations when filling out the survey. Since

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

forecasting future realizations for s amounts to forecasting the underlying state (one period lagged).

Importantly, we allow for the possibility that forecasters' probability measure $\mathscr{P}$ is subjective. Specifically, we consider subjective point beliefs about the value of the variances $(\sigma_u^2, \sigma_\nu^2, \sigma_\epsilon^2)$, which we denote by $(\hat{\sigma}_u^2, \hat{\sigma}_\nu^2, \hat{\sigma}_\epsilon^2)$. In the special case where $(\hat{\sigma}_u^2, \hat{\sigma}_\nu^2, \hat{\sigma}_\epsilon^2) = (\sigma_u^2, \sigma_\nu^2, \sigma_\epsilon^2)$ we are in a situation in which forecasters hold rational expectations.

When forecasters' prior beliefs $\pi_{t|t-1}^i \equiv \mathbb{E}^{\mathscr{P}}\left[\pi_t \mid \Omega_{t-1}^i\right]$ are normally distributed and if prior uncertainty is equal to the steady-state value of uncertainty implied by the subjective Kalman filter, then forecaster i finds it optimal to use a prediction rule of the form

$$\mathbb{E}^{\mathscr{P}}[s_{t+1}|\Omega_t^i] = \underbrace{\mathbb{E}^{\mathscr{P}}\left[\pi_t \mid \Omega_t^i\right]}_{\equiv \pi_{t|t}^i} = (1-\kappa_x-\kappa_y)\pi_{t|t-1}^i + \kappa_x x_t^i + \kappa_y \rho s_t, \tag{3.5}$$

where κ_x and κ_y denote the weights implied by the (subjective) Kalman filter.[8] The previous equation can equivalently be written as

$$\pi_{t|t}^i = \kappa_x x_t^i + (1-\kappa_x)\rho\left[\omega s_t + (1-\omega)\pi_{t-1|t-1}^i\right],$$

where the Kalman filter parameters are now given by (κ_x, ω) with $\omega \equiv \kappa_y/(1-\kappa_x)$ and

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

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

where $\hat{\sigma}_\tau^2$ is the (stationary subjective) uncertainty about π_t given information Ω_t^i, which is given by

$$\widehat{\sigma}_\tau^2 = \frac{\widehat{\kappa}_x^2\widehat{\sigma}_\epsilon^2 + (1-\widehat{\kappa}_x)^2\widehat{\sigma}_u^2 + \rho^2(1-\widehat{\kappa}_x)^2\widehat{\omega}^2\widehat{\sigma}_v^2}{1-\rho^2(1-\widehat{\kappa}_x)^2(1-\widehat{\omega})^2}. \tag{3.8}$$

For the case with rational beliefs ($(\hat{\sigma}_u^2, \hat{\sigma}_v^2, \hat{\sigma}_\epsilon^2) = (\sigma_u^2, \sigma_v^2, \sigma_\epsilon^2)$), the previous equations deliver the rational Kalman filter weights that we denote by ω^* and κ_x^*.

3.2 Model performance with rational expectations

We first explore the predictions of the updating model under rational expectations. In this setup, deviations from full information rational expectations (FIRE) are exclusively due to deviations from full information, i.e., due to the presence of (i) an unobserved state and (ii) private information. The following proposition shows that the model then fails to replicate almost all empirical facts:

[8] If prior uncertainty is not equal to the steady-state value, then the Kalman filter weights depend on time but deterministically converge to their steady-state values κ_x and κ_y.

Proposition 1. *Under rational expectations:*

1. *Forecasters' expectations neither over- nor under-react to public news-related forecast revisions ($\beta_{1,h} = 0$), contrary to Fact 1.*
2. *Forecasters' expectations neither over- nor under-react to the residual component of forecast revisions ($\beta_{2,h} = 0$), contrary to Fact 2.*
3. *Forecasters' expectations neither over- nor under-react to public news ($\alpha_{1,h} = 0$), contrary to Fact 3.*
4. *Forecasters' expectations are correctly anchored to prior expectations ($\alpha_{2,h} = 0$), contrary to Fact 4.*
5. *Forecasters' expectations neither over- nor under-react to past forecast revisions at the individual level ($\beta_h^p = 0$), contrary to the Fact in Figure 1.*
6. *Consensus forecasts underreact to past consensus forecast revisions ($\beta_h^c > 0$), consistent with the Fact in Figure 1.*

The proof of proposition 1 is in Appendix D.1. Perhaps not surprisingly, with rational expectations, forecast errors cannot be explained by information available to agents at the time of forecasting, in contrast to Facts 1 to 4 and in contrast to the evidence provided in Bordalo et al. (2020). With rational expectations, the model only matches the evidence in Coibion and Gorodnichenko (2015): since forecasters know that private information is contaminated by noise, they adjust beliefs only gradually to private information. Since this is true for all forecasters, this causes the forecast errors associated with the average forecasts (across forecasters) to be predicted by past revisions in average forecasts ($\beta_h^c > 0$).

3.3 Overconfidence in private information

We now introduce a single belief distortion and show that the Bayesian updating model then qualitatively replicates all documented deviations from FIRE. In particular, individuals perceive the standard error of the observation noise in their private signal to be given by

$$\widehat{\sigma}_\epsilon = \tau \sigma_\epsilon, \tag{3.9}$$

for some $\tau \geq 0$. When $\tau < 1$ forecasters are overconfident in the information content of their private signal because they underestimate the noise contained in the signal.[9] Forecasters hold rational beliefs about all other parameters, i.e., $(\hat{\sigma}_u^2, \hat{\sigma}_v^2) = (\sigma_u^2, \sigma_v^2)$.

Given these beliefs, agents will find it optimal to update their expectations using the following (subjective) Kalman weights:

$$\widehat{\omega} = \frac{(\sigma_v^2)^{-1}}{(\widehat{\sigma}_\tau^2)^{-1} + (\sigma_v^2)^{-1}}, \tag{3.10}$$

$$\widehat{\kappa}_x = \frac{(\widehat{\sigma}_\epsilon^2)^{-1}}{(\widehat{\sigma}_\epsilon^2)^{-1} + \left[\rho^2 \left(\widehat{\omega}^2 \sigma_v^2 + (1-\widehat{\omega})^2 \widehat{\sigma}_\tau^2\right) + \sigma_u^2\right]^{-1}}, \tag{3.11}$$

where agents' prior uncertainty is

$$\widehat{\sigma}_\tau^2 = \frac{\widehat{\kappa}_x^2 \widehat{\sigma}_\epsilon^2 + (1-\widehat{\kappa}_x)^2 \sigma_u^2 + \rho^2 (1-\widehat{\kappa}_x)^2 \widehat{\omega}^2 \sigma_v^2}{1 - \rho^2 (1-\widehat{\kappa}_x)^2 (1-\widehat{\omega})^2}. \tag{3.12}$$

The following proposition presents our main analytic result:

Proposition 2. *When agents are overconfident about the information content of their private signal* ($0 \leq \tau < 1$)*, then:*

1. *Forecasters' expectations underreact to public news-related forecast revisions* ($\beta_{1,h} > 0$)*, consistent with Fact 1.*
2. *Forecasters' expectations overreact to the residual component of forecast revisions* ($\beta_{2,h} < 0$)*, 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 past forecast revisions at the individual level* ($\beta_h^p < 0$)*, consistent with the Fact in Figure 1.*
6. *If* $\tau > 1/I$*, then consensus forecasts underreact to past consensus forecast revisions* ($\beta_h^c > 0$)*, consistent with the Fact in Figure 1.*

[9] Conversely, for $\tau > 1$ forecasters are underconfident because they overestimate the standard deviation of the noise.

The proof of the proposition can be found in Appendix D.2. Intuitively, when forecasters are overly optimistic about the noise contained in private information ($\tau < 1$), they overreact to private signals ($\hat{\kappa}_x > \kappa_x^*$) and underreact to the forecast revision related to public news. Overreaction to private information also explains why belief revisions to be "too strong", so that expectations overreact to past forecast revisions at the individual level. The high perceived information content of private information also causes prior uncertainty to be lower than with rational expectations ($\hat{\sigma}_\tau^2 < \sigma_\tau^{2*}$). As a result, agents overly anchor beliefs to prior information ($\hat{\omega} < \omega^*$). However, the response of period-by-period belief revisions to prior beliefs, $(1-\widehat{\kappa_x})\hat{\omega}$, can be larger or smaller than with rational expectations, $(1-\kappa_x^*)\omega^*$. Yet, as the proposition shows, one always gets that beliefs that are overly anchored to prior beliefs ($\alpha_{2,h} < 0$), when considering the full dynamic outcome. Finally, overconfidence in private information is consistent with underreaction of consensus forecasts to past consensus forecast revisions, as in the case with rational expectations, provided $\tau > 1/I$. Since we observe approximately 40 forecasters in the SPF, the latter condition is lax and very close to zero.

3.4 Further tests of the overconfidence model

The overconfidence model in the previous section implies that the residuals in the empirical forecast revision equation (2.3) are due to private information. This interpretation of residual information gives rise to further testable predictions. In this section, we derive these predictions and show 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_t^i. We can decompose this reaction into a component that is common across forecasters, $\kappa_x \frac{1}{N} \sum_i x_t^i$, where N denotes the number of forecasters, and an idiosyncratic component.[10] When N is large, then the common component represents very precise information about the variable that gets forecasted, see equation (3.3). In contrast, the idiosyncratic component of private information reflects observation noise that is detrimental to forecasting performance. This implication can be tested in the data.

Specifically, consider the common and idiosyncratic components (2.8)-(2.9) of the residual

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

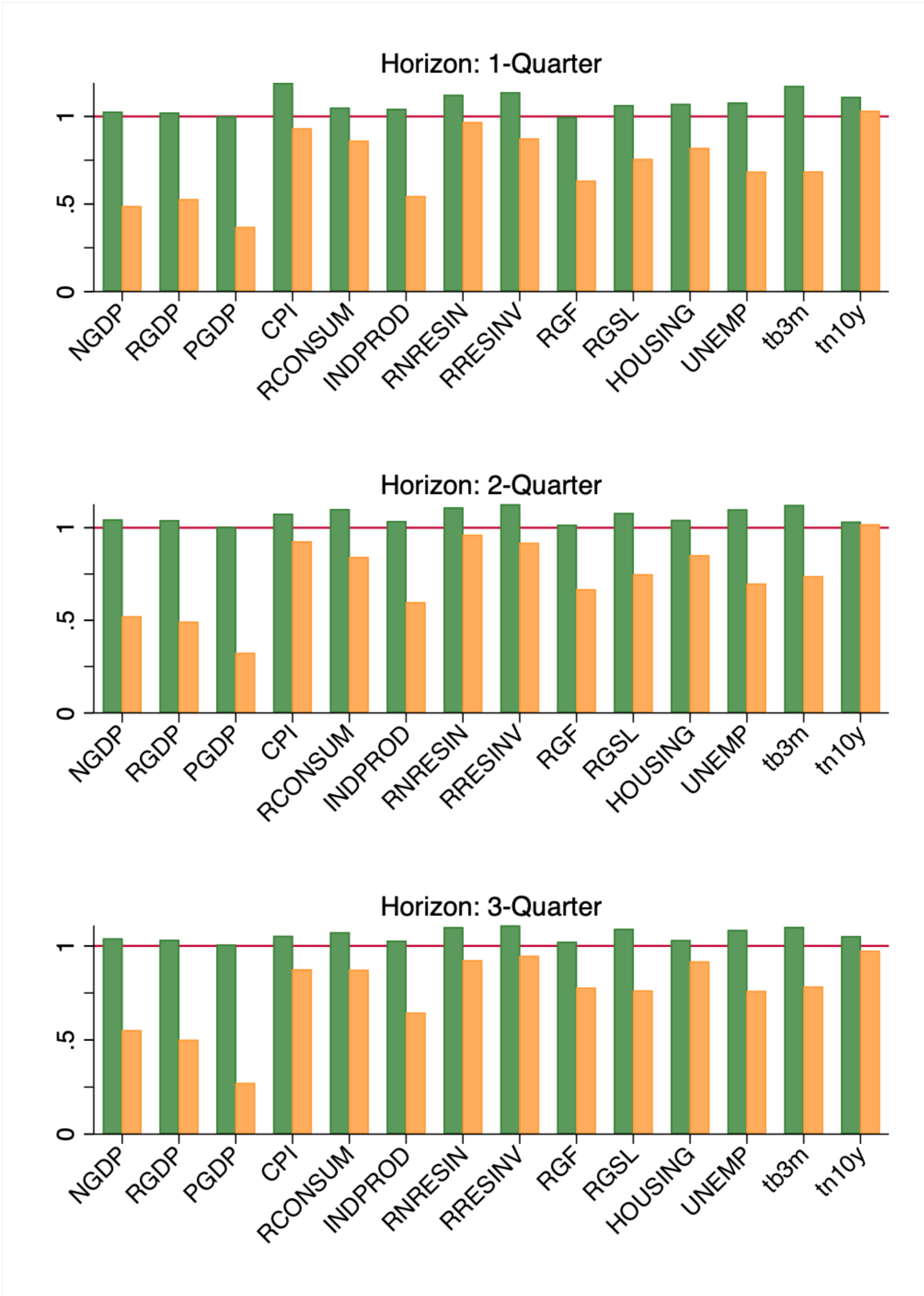


Figure 7: INDIVIDUAL FORECAST ERRORS: COMMON VS. IDIOSYNCRATIC COMPONENTS OF RESIDUAL INFORMATION

Notes: This figure compares the individual forecast errors implied by the updating equation (2.3) to those implied when replacing the residuals $\epsilon^i_{h,t}$ 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 (2.3), which uses both the common and the idiosyncratic components.

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

Figure 7 computes the resulting mean squared forecast errors (averaged across all forecasters) for each variable and forecast horizon, relative to the forecast errors implied by agents' actual forecasts, which is tantamount to using both the idiosyncratic *and* the common component in the updating equation (3.5). The figure shows that using the common component instead of $\epsilon^i_{h,t}$ 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 which implies that residual information is due to noisy private information.

4 Quantitative performance of the overconfidence model

This section provides a quantitative assessment of the ability of our Bayesian updating model to capture the documented empirical patterns in professional forecasts. We estimate the model using the simulated method of moments, evaluate its quantitative fit, and present estimates of the overconfidence parameter τ.

4.1 Estimation approach

We use the simulated method of moments to estimate the parameter vector

$$x \equiv (\tau, \sigma_\epsilon/\sigma_u, \sigma_\nu/\sigma_u, \rho, \sigma_u) \in R^5, \tag{4.1}$$

targeting the eight data moments

$$\widehat{\Gamma} \equiv (\widehat{\alpha}_{1,h}, \widehat{\alpha}_{2,h}, \widehat{\beta}_{1,h}, \widehat{\beta}_{2,h}, \widehat{\beta}^p_h, \widehat{\beta}^c_h, \sigma(FE), \sigma(FR)) \in R^8 \tag{4.2}$$

for $h = 1$, where the first six moments are the regression coefficients discussed at length in the previous sections, $\sigma(FE)$ the standard deviation of individual one-step-ahead forecast errors

$(\pi_{t+1} - \pi^i_{t+1|t})$, and $\sigma(FR)$ the standard error of individual forecast revisions $(\pi^i_{t+1|t} - \pi^i_{t+1|t-1})$. We add the last two moments as estimation targets to ensure that the forecast errors and forecast revisions behave in line with the data, following Bordalo et al. (2020).

Given the overconfidence parameter τ, the noise-to-signal ratios $(\sigma_\epsilon/\sigma_u, \sigma_\nu/\sigma_u)$, and the persistence parameter ρ, we can compute the Kalman filter weights (ω, κ_x) by solving equations (3.9)-(3.12) using a fixed-point search algorithm. Given these solutions, we can compute $\sigma(FR)$ and $\sigma(FE)$ using equations (C.2) and (C.5) from the appendix, the individual CG coefficient β^p_1 using analytic results from appendix Appendix C.2, and the regression coefficients $(\alpha_{1,1}, \alpha_{2,1}, \beta_{1,1}, \beta_{2,1})$ using the analytic formulas in appendices Appendix C.4 - Appendix C.6. We do not have closed-form expressions for the consensus CG coefficient β^c_1, thus compute it using a simulation approach.[11]

For each forecast series k, we let $\hat{\Gamma}_k$ denote the empirical moments and $\Gamma(x_k)$ the model moments implied by parameter vector x_k. We then estimate $\widehat{x}_k$ as

$$\widehat{x}_k = \arg\min_{x_k} \quad (\hat{\Gamma}_k - \Gamma(x_k))' I (\hat{\Gamma}_k - \Gamma(x_k)),$$

where I is the identity matrix. We impose the estimation bounds $\rho \in [0,1]$ and $\sigma_\varepsilon/\sigma_u, \sigma_\nu/\sigma_u \in [0,10]$ to ensure that parameters remain within a-priori reasonable ranges. Without these bounds, the fit of the model with the data would improve further.[12] Importantly, however, we leave the overconfidence parameter $\tau \geq 0$ in equation (3.9) unrestricted in the estimation.[13]

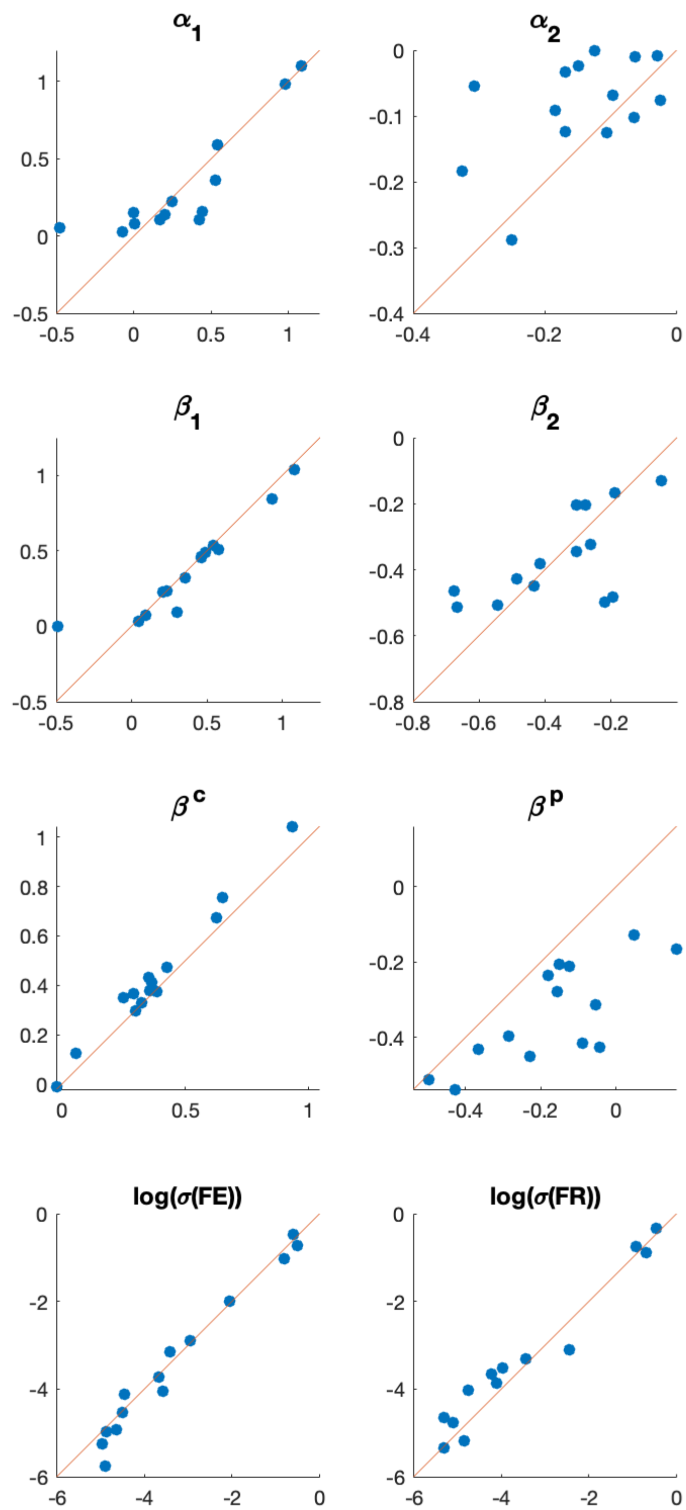


Figure 8: TARGETED MOMENTS: DATA VS. MODEL

Notes: This figure plots the data moments on the horizontal axis, the moments of the estimated models on the vertical axis, and 45^o lines in red.

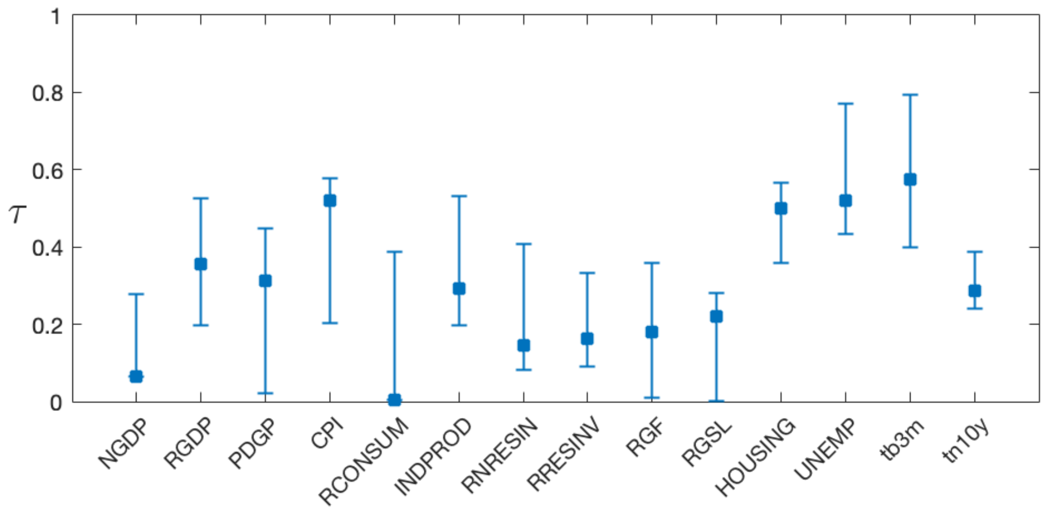


Figure 9: ESTIMATED DEGREE OF OVERCONFIDENCE τ IN EQUATION (3.9)

Notes: This figure plots estimated values of the overconfidence parameter τ and bootstrapped 90% confidence intervals.

4.2 Estimation outcome

Figure 8 reports - for each of the 8 targeted data moments - a scatter plot displaying for the 14 forecast variables the empirical moment (on the horizontal axis) and the model moment (on the vertical axis). The figure also depicts 45^o lines (in red), which indicate a perfect model fit. Our simple estimated models manage to replicate the data surprisingly well, with most estimates aligning well around the 45^o lines. The only systematic deviation occurs for the individual CG coefficient β^p, which the model predicts to be consistently more negative than in the data. Appendix E provides further evaluations of the model fit for longer forecast horizons ($h = 2, 3$) and shows that the model also performs well at longer forecast horizons, even though these

[11]We proceed as follows: (i) we simulate the AR(1) process for $\tilde{\pi}_t$ for $t = 1, \ldots, 100$; (ii) we simulate a time series of private and public signals, $\tilde{x}^i_t = \tilde{\pi}_t + \epsilon^i_t$ and $\tilde{s}_t = \tilde{\pi}_t + \nu_t$, where ϵ^i_t is drawn from $N(0, \sigma^2_\epsilon)$, i.i.d. across time and forecasters, for $i = 1, \ldots, 50$, and ν_t is drawn from $N(0, \sigma^2_\nu)$ i.i.d. across time; (iii) we simulate the forecasts using equation (3.5), setting initial forecasts equal to zero (the unconditional mean of the forecasted variables); (iv) we use these forecasts to compute the consensus forecasts and then use consensus forecasts to compute consensus forecast revisions and consensus forecast errors; (v) we estimate the consensus CG coefficient β^c_1; (vi) we repeat the process described in (i)-(v) 500 times and then use the average coefficient estimate as the expected value of the consensus CG coefficient implied by the considered parameter vector.

[12]The quantitative findings about model fit and estimates for τ are robust to lifting the bounds on the signal-to-noise ratios.

[13]The zero bound for τ is required to insure that standard deviations in the model remain positive, see equation (3.9).

moments have not been used in the estimation. Overall, Figure 8 shows that our simple updating model performs surprisingly well in *quantitatively* replicating the empirical evidence.

Of primary interest are the implied estimates of the parameter τ for the 14 considered forecast variables. The model displays overconfidence in private information whenever $\tau < 1$, see equation (3.9). Figure 9 reports the point estimates together with 90% confidence intervals obtained from bootstrapping for each of the considered forecast variables.[14] It shows that all estimates $\widehat{\tau}_k$ are statistically significantly below 1, with most of them ranging between 0.2 and 0.5. This shows that professional forecasters significantly underestimate the standard deviation of noise contained in their private information by a factor ranging between 2 and 5.

5 Conclusion

Observing public information available to professional forecasters, we document a number of new facts about the behavior of forecasts in the Survey of Professional Forecasters. A simple model in which forecasters overreact to (noisy) private information explains these new facts, but also explains 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 professional forecasters' over-reliance on private information.

[14] The estimates of the remaining parameters are reported in Appendix E.

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.

Marcet, Albert and Thomas J. Sargent, "Convergence of Least Squares Learning Mechanisms in Self Referential Linear Stochastic Models," *Journal of Economic Theory*, 1989, *48*, 337–368.

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

Figure A.1 presents a sample of the historical SPF data sheet mentioned in the main text.

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.1: 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.2 - A.4 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.5 and A.6 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.7, 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.[15]

[15] 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.

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.2: 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	SAME AS	
14. 10-YR U.S. TREASURY BOND	Secondary Market, %	2.98	4.10				1.52	DEC 30,	DEC 29	
15. 90-DAY U.S. TREASURY BILL	Secondary Market, %	1.66	4.06				0.06	2022	2023	
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.3: SAMPLE LIVINGSTON SURVEY FORM AND HISTORICAL DATA SHEET (PAGE 2)

Table B

HISTORICAL DATA for DECEMBER SURVEY

QUARTERLY INDICATORS		Quarterly Data							Annual Data		
		2021				2022					
		Q1	Q2	Q3	Q4	Q1	Q2	Q3	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							Annual Data		
		2022									
		APRIL	MAY	JUNE	JULY	AUG	SEPT	OCT	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							Annual, End-of-Period		
		2022									
		29-Apr	31-May	30-Jun	29-Jul	31-Aug	30-Sep	31-Oct	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.4: 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 of 2

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.5: 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	INCREASE		NO CHANGE		DECREASE		Total
June 14, 2023 ? *NOT THE MEETING ON MAY 3		+		+		=	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*

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.6: 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

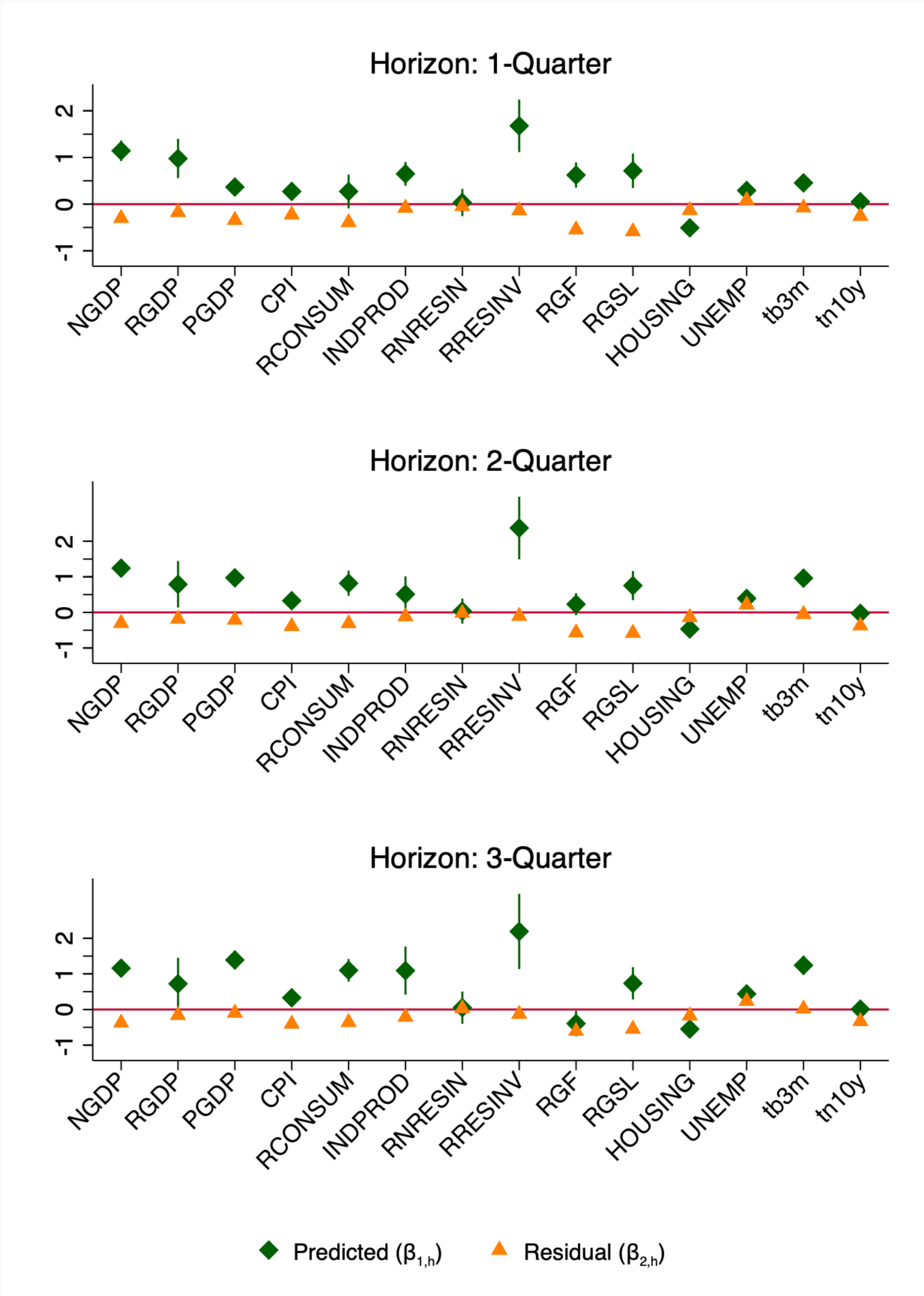
HICP inflation (March 2003)	2.4%
Annual GDP growth (2002 Q4)	1.3% (according to the ESA 95 definition)
Unemployment rate (February 2003)	8.7%

Appendix Figure A.7: ECB SPF SURVEY INFORMATION

Appendix B Additional results on empirical analyses

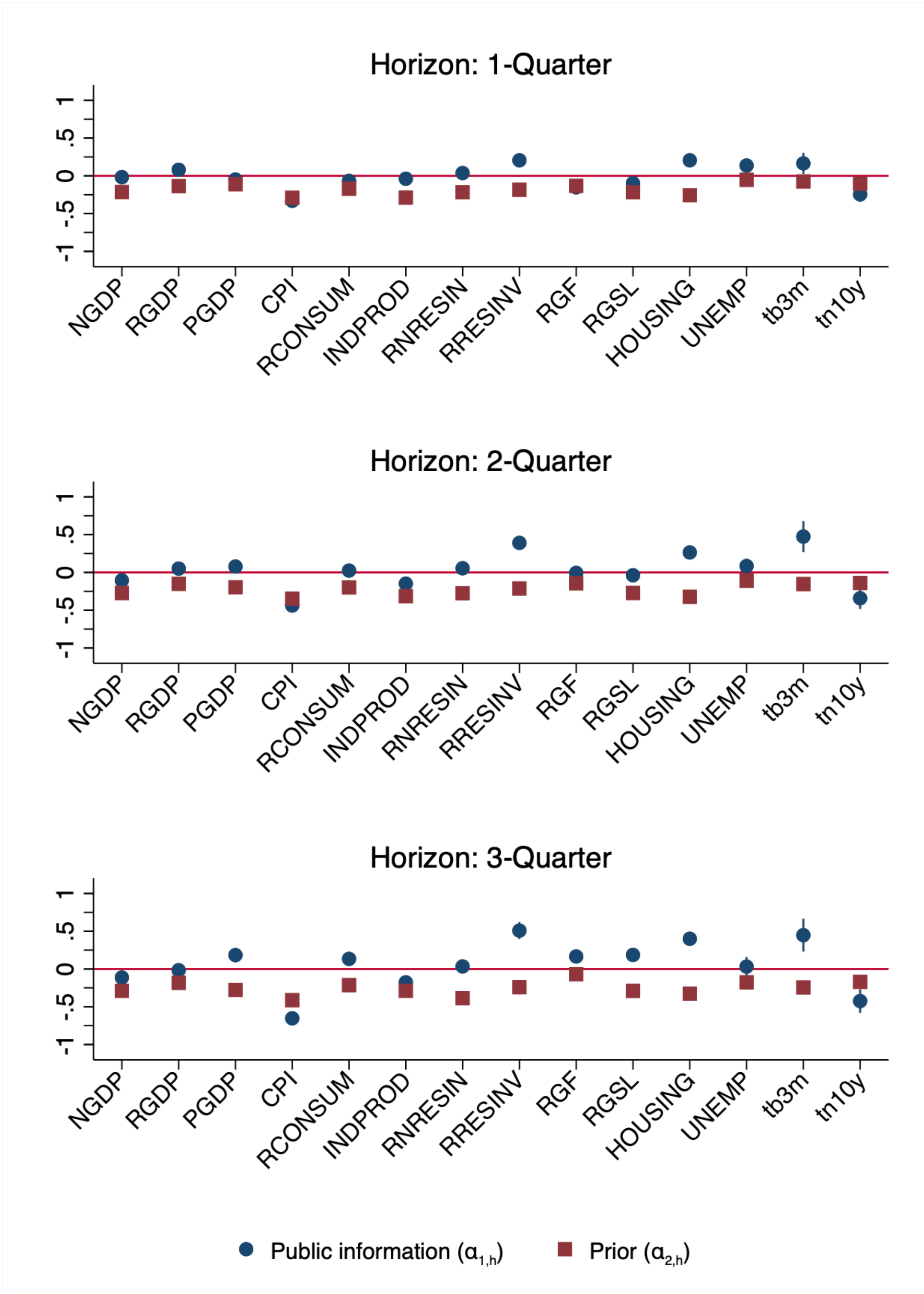
Appendix B.1 One-dimensional public information

In this appendix, we consider a special case where s_t in Eqn. (2.3) is one-dimensional. Specifically, s_t is the most recent release on the dependent variable π, the realized value of π in the previous period. We repeat the analysis in Section 2.4 and report the results in Figure B.8 - B.10.



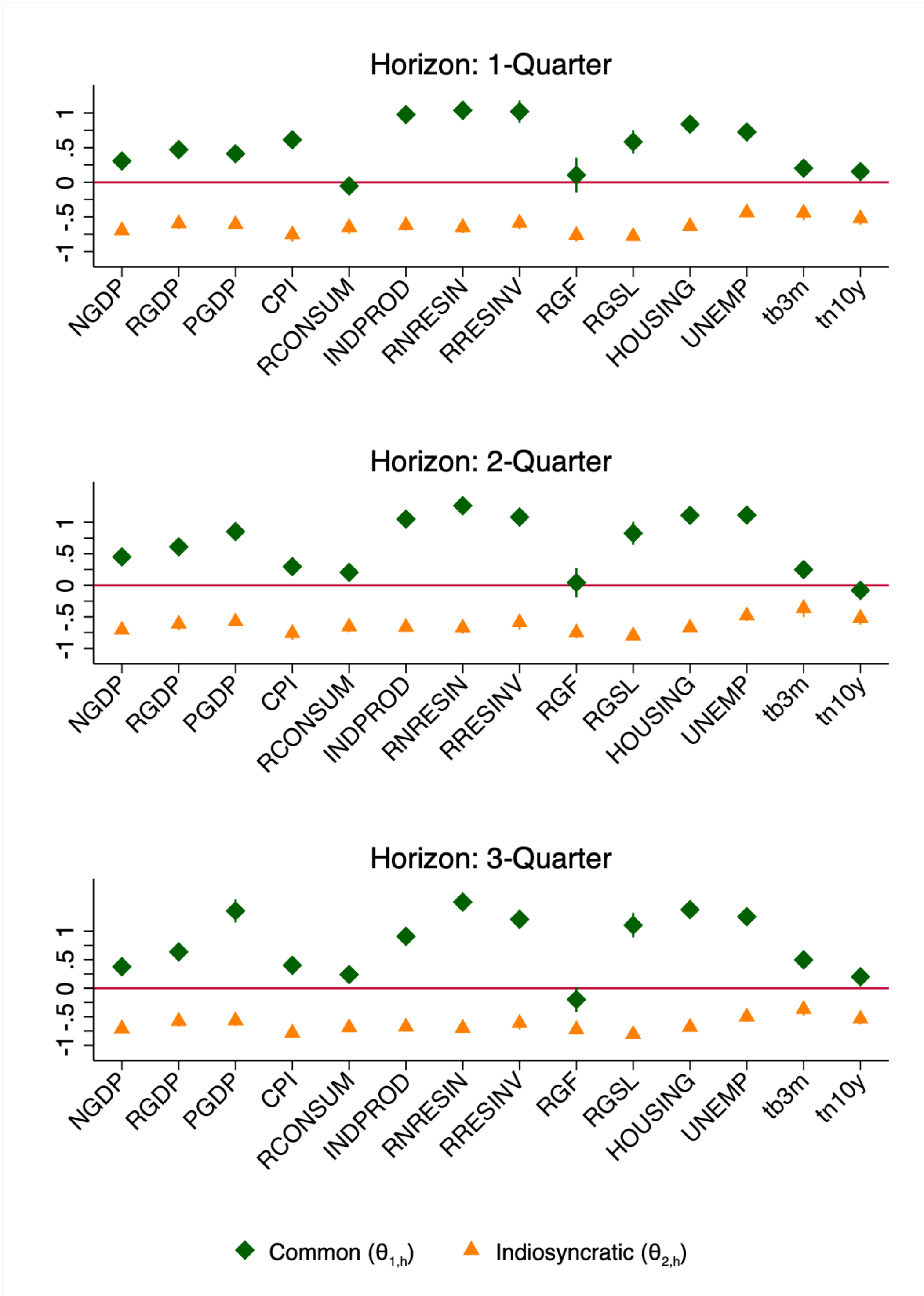
Appendix Figure B.8: RESPONSES OF FORECAST ERRORS TO FORECAST REVISION DECOMPOSITION: 1-DIMENSIONAL SIGNAL

Notes: This figure plots the coefficients of $\beta_{1,h}$ (in green) and $\beta_{2,h}$ (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.9: RESPONSES OF FORECAST ERRORS TO PRIOR AND REAL-TIME DATA RELEASE: 1-DIMENSIONAL SIGNAL

Notes: This figure plots the estimated coefficients of $\alpha_{1,h}$ (in blue) and $\alpha_{2,h}$ (in maroon) from Eqn. (2.7). 95% confidence intervals based on clustered standard errors are reported.



Appendix Figure B.10: RESPONSES OF FORECAST ERRORS TO PRIVATE INFORMATION DECOMPOSITION: 1-DIMENSIONAL SIGNAL

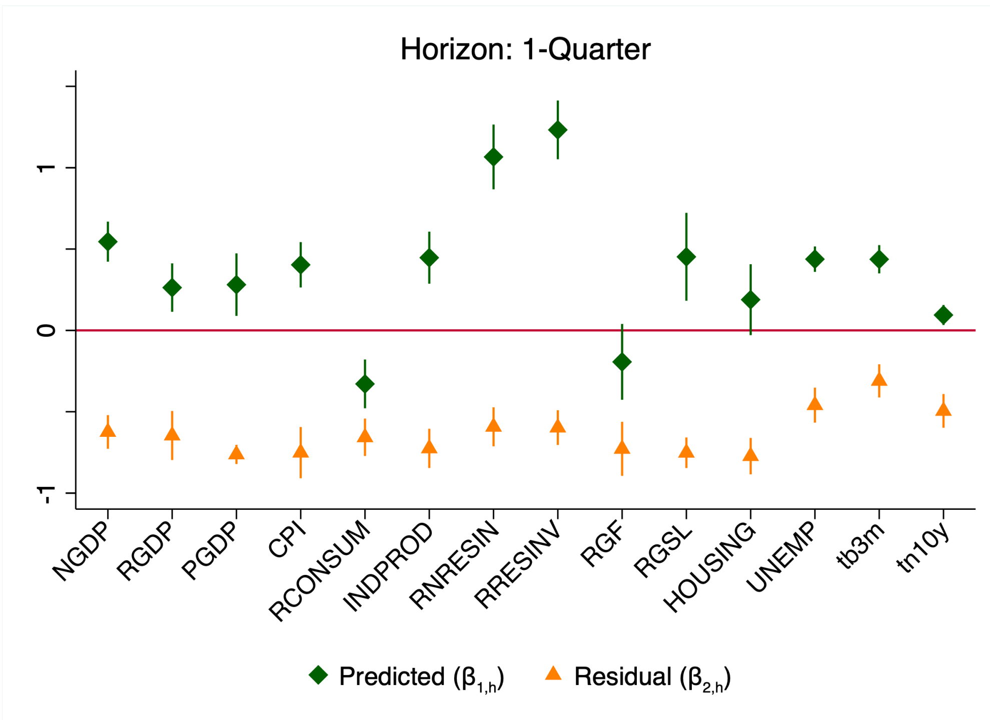
Notes: This figure plots the estimated coefficients of $\theta_{1,h}$ (in green) and $\theta_{2,h}$ (in orange) from Eqn. (2.10). 95% confidence intervals based on clustered standard errors are reported.

Appendix B.2 Consensus forecast as public information

In this appendix, we conduct a robustness analysis by including consensus forecasts as a subset of the public news in Eqn. (2.3). Specifically, consider the following regression

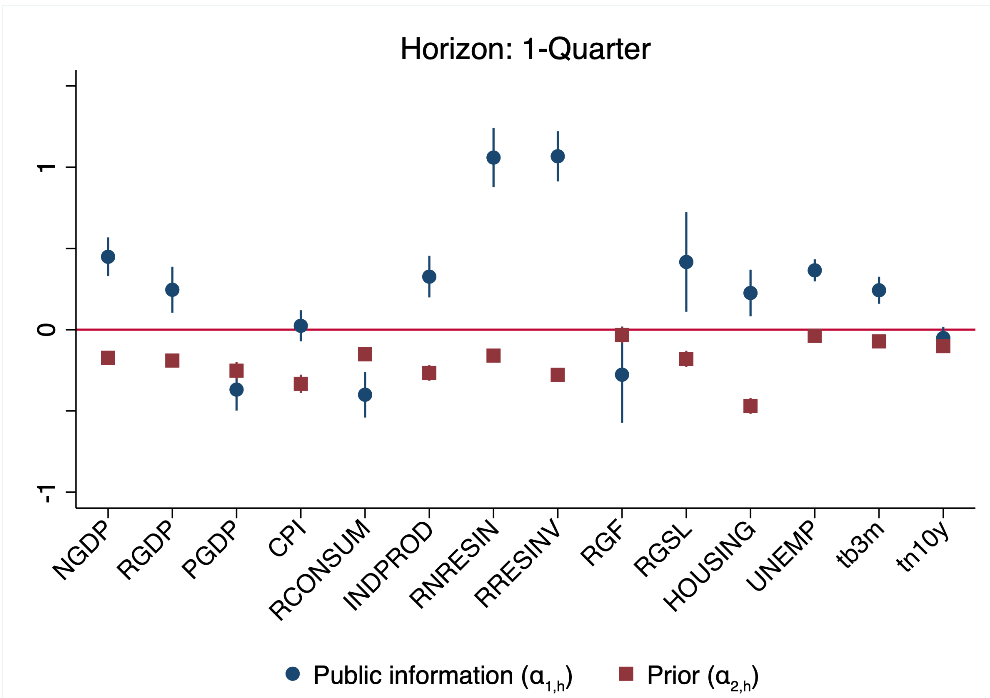
$$\pi^i_{t+h|t} - \pi^i_{t+h|t-1} = \tilde{\tilde{\delta}}^i_h + \tilde{\gamma}_h x_t + \tilde{\eta}_h \circ \pi^i_{t+h|t-1} + \epsilon^i_{h,t}, \tag{B.1}$$

where x_t is a vector containing the public news $(s_t - s^i_{t|t-1})$ and the revisions of consensus forecasts $(\pi^c_{t+h|t-1} - \pi^c_{t+h|t-2})$. The coefficient matrix $\tilde{\gamma}_h \in R^{28\times 28}$ captures how forecasters respond to public news as well as news in consensus forecasts. Since this analysis requires knowledge of $\pi^c_{t+h|t-2}$, we repeat the analysis in Section 2.4 for $h = 1$ only and report the results in Figure B.11 - B.13.



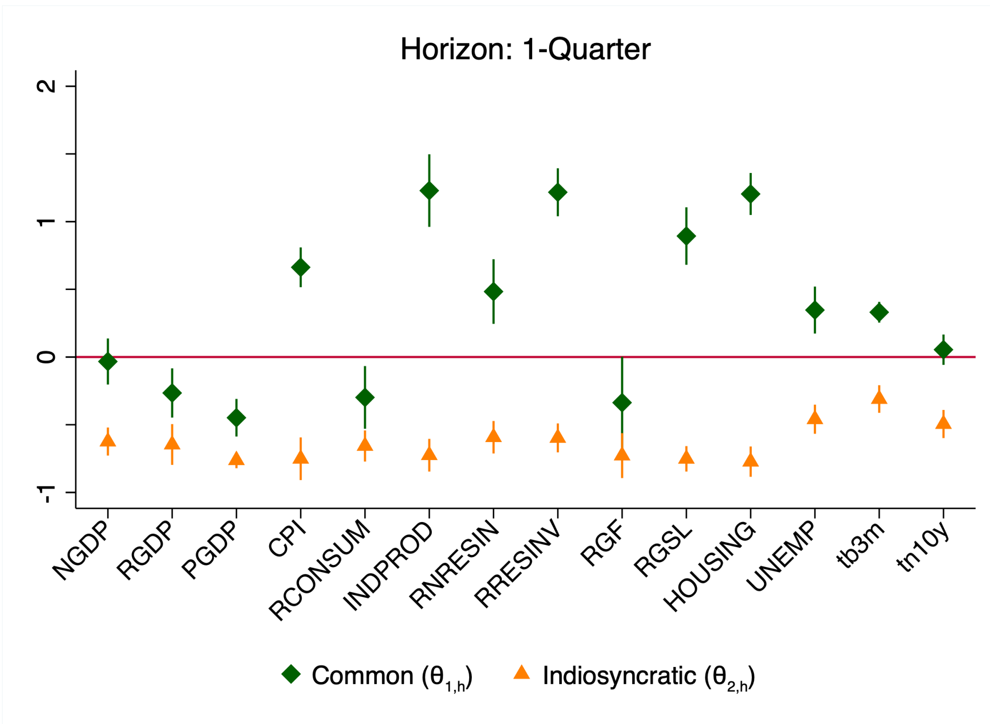
Appendix Figure B.11: RESPONSES OF FORECAST ERRORS TO FORECAST REVISION DECOMPOSITION: CONSENSUS FORECASTS AS ADDITIONAL PUBLIC INFORMATION

Notes: This figure plots the coefficients of $\beta_{1,h}$ (in green) and $\beta_{2,h}$ (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.12: RESPONSES OF FORECAST ERRORS TO PRIOR AND REAL-TIME DATA RELEASE: CONSENSUS FORECASTS AS ADDITIONAL PUBLIC INFORMATION

Notes: This figure plots the estimated coefficients of $\alpha_{1,h}$ (in blue) and $\alpha_{2,h}$ (in maroon) from Eqn. (2.7). 95% confidence intervals based on clustered standard errors are reported.



Appendix Figure B.13: RESPONSES OF FORECAST ERRORS TO PRIVATE INFORMATION DECOMPOSITION: CONSENSUS FORECASTS AS ADDITIONAL PUBLIC INFORMATION

Notes: This figure plots the estimated coefficients of $\theta_{1,h}$ (in green) and $\theta_{2,h}$ (in orange) from Eqn. (2.10). 95% confidence intervals based on clustered standard errors are reported.

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:

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

The forecast revision at time t is:

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

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

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

Therefore,

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

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

Appendix C.2 Compute individual CG coefficients

The individual-level CG coefficient is

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

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

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

In particular,

$$\begin{aligned}\mathbb{C}\text{ov}\left(FE^i_t, FR^i_t\right) &= (1-\kappa_x-\kappa_y)(\kappa_x+\kappa_y)\rho^2\mathbb{V}\text{ar}(FE^i) + (1-\kappa_x)\kappa_x\sigma^2_u - \kappa^2_x\sigma^2_\epsilon - \rho^2\kappa^2_y\sigma^2_v \\ &= \left[1-(1-\kappa_x-\kappa_y)\rho^2\right]\mathbb{E}\left[(\pi_t - \pi^i_{t|t})\pi^i_{t|t}\right]\end{aligned} \tag{C.11}$$

When forecasts are optimal, $\mathbb{E}\left[(\pi_t - \pi^i_{t|t})\pi^i_{t|t}\right] = 0$ since forecast errors $(\pi_t - \pi^i_{t|t})$ are not predictable, and are therefore, orthogonal to the forecasts $(\pi^i_{t|t})$. As a result, $\mathbb{C}\text{ov}\left(FE^i_t, FR^i_t\right) = 0$, 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 $x_t = \frac{\sum_i x^i_t}{N_t}$ with $\mathbb{V}\text{ar}(x_t) = \frac{\sigma^2_\epsilon}{N_t}$, and N_t is the number of forecasters in period t.

The consensus CG coefficient is

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

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

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):

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

We derive the OLS coefficient estimates as follows:

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

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

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

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

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

$$
\begin{aligned}
\mathscr{N}^\gamma_h \equiv\; &\mathbb{E}(\pi^i_{t+h|t-1})^2 \mathbb{E}\left[(s_t - \pi^i_{t-1|t-1})(\pi^i_{t+h|t} - \pi^i_{t+h|t-1})\right] \\
&- \mathbb{E}\left[(s_t - \pi^i_{t-1|t-1})\pi^i_{t+h|t-1}\right] \mathbb{E}\left[\pi^i_{t+h|t-1}(\pi^i_{t+h|t} - \pi^i_{t+h|t-1})\right] \\
=\; &\rho^{3h+2}\mathbb{E}(\pi^i_{t-1|t-1})^2 \mathbb{E}\left[(s_t - \pi^i_{t-1|t-1})(\pi^i_{t|t} - \pi^i_{t|t-1})\right] \\
&- \rho^{3h+2}\mathbb{E}\left[(s_t - \pi^i_{t-1|t-1})\pi^i_{t-1|t-1}\right] \mathbb{E}\left[\pi^i_{t-1|t-1}(\pi^i_{t|t} - \pi^i_{t|t-1})\right] \\
=\; &\rho^{3(h+1)}(\kappa_x + \kappa_y)\left(\mathbb{E}(\pi^i_{t-1|t-1})^2 \mathbb{E}\left[(\pi_{t-1} - \pi^i_{t-1|t-1})^2\right] - \left(\mathbb{E}\left[(\pi_{t-1} - \pi^i_{t-1|t-1})\pi^i_{t-1|t-1}\right]\right)^2\right) \\
&+ \rho^{3(h+1)}\mathbb{E}(\pi^i_{t-1|t-1})^2\kappa_y\sigma^2_\nu \\
\mathscr{N}^\eta_h \equiv\; &-\mathbb{E}\left[(s_t - \pi^i_{t-1|t-1})\pi^i_{t+h|t-1}\right] \mathbb{E}\left[(s_t - \pi^i_{t-1|t-1})(\pi^i_{t+h|t} - \pi^i_{t+h|t-1})\right] \\
&+ \mathbb{E}\left[\pi^i_{t+h|t-1}(\pi^i_{t+h|t} - \pi^i_{t+h|t-1})\right] \mathbb{E}(s_t - \pi^i_{t-1|t-1})^2 \\
=\; &-\rho^{2h+1}\mathbb{E}\left[(\pi_{t-1} - \pi^i_{t-1|t-1})\pi^i_{t-1|t-1}\right] \mathbb{E}\left[(\pi_{t-1} - \pi^i_{t-1|t-1})(\pi^i_{t|t} - \pi^i_{t|t-1})\right] \\
&- \rho^{2h+1}\mathbb{E}\left[(\pi_{t-1} - \pi^i_{t-1|t-1})\pi^i_{t-1|t-1}\right]\kappa_y\rho\sigma^2_\nu \\
&+ \rho^{2h+1}\mathbb{E}\left[\pi^i_{t-1|t-1}(\pi^i_{t|t} - \pi^i_{t|t-1})\right] \mathbb{E}(\pi_{t-1} - \pi^i_{t-1|t-1})^2 \\
&+ \rho^{2h+1}\mathbb{E}\left[\pi^i_{t-1|t-1}(\pi^i_{t|t} - \pi^i_{t|t-1})\right]\sigma^2_\nu
\end{aligned}
$$

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

Thus, $\gamma_h = \frac{\mathcal{N}_h^\gamma}{\mathcal{D}_h}$ and $\eta_h = \frac{\mathcal{N}_h^\eta}{\mathcal{D}_h}$. In particular, $0 < \gamma_h \le \rho^{h+1}(\kappa_x+\kappa_y)$ 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):

$$
\pi_{t+h}-\pi^i_{t+h|t} = \beta_{1,h}\times \text{Predicted}^i_{h,t} + \beta_{2,h}\times \text{Residual}^i_{h,t} + \nu^i_{h,t}. \tag{C.21}
$$

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

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

where the numerator of β_1 is

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

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

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

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

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

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

The numerator of β_2 is

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

Note that the numerator of $\beta_{1,h}$ and the numerator of $\beta_{2,h}$ 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):

$$
\pi_{t+h}-\pi^i_{t+h|t} = \alpha_{1,h}(s_t-\pi^i_{t-1|t-1}) + \alpha_{2,h}\pi^i_{t+h|t-1} + \beta_{2,h}\times \mathrm{Residual}^i_{h,t} + \nu^i_{h,t}. \qquad \text{(C.27)}
$$

Note that by construction, $\mathrm{Residual}^i_{h,t}$ 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, $\alpha_{1,h}=\frac{\mathcal{N}^\alpha_{1,h}}{\mathcal{D}_h}$ and $\alpha_{2,h}=\frac{\mathcal{N}^\alpha_{2,h}}{\mathcal{D}_h}$.

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 $\beta_{1,h}$ 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 $\beta_{2,h}$ 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 $\alpha_{1,h}$ follows the sign of $\mathcal{N}^\alpha_{1,h}$ (Eqn. C.28), which always equals 0 under RE.

4. The sign of $\alpha_{2,h}$ 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 $\hat{\sigma}_\tau^2 = \tau\sigma_\tau^2$, 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 (3.12) 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, $\mathbb{V}\mathrm{ar}(FE^i) > \hat{\sigma}_\tau^2$. Therefore, from Eqn. (C.7) and (D.3), we get

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

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

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

Eqn. (C.7) gives

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

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

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

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 $\beta_{1,h}$ 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 $\mathbb{V}\text{ar}(FE^i) > \hat{\sigma}_\tau^2$. 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 $\beta_{2,h}$ follows the sign of $\mathcal{N}_{2,h}^\beta$ (Eqn. C.26). Since $\mathcal{N}_{1,h}^\beta + \mathcal{N}_{2,h}^\beta \propto \mathbb{C}\text{ov}(FE_t^i, FR_t^i) < 0$ under overconfidence, given that $\beta_{1,h} > 0$, it follows that $\beta_{2,h} < 0$.

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

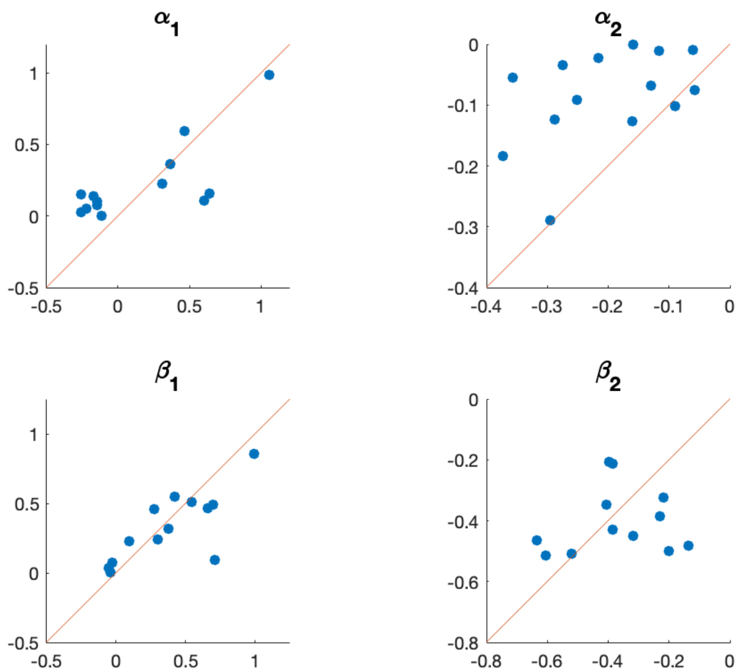
$$\begin{aligned}
\mathcal{N}_{1,h}^\alpha \propto\ & (1-\kappa_x-\kappa_y)\left[\mathbb{E}(\pi_{t-1|t-1}^i)^2\left(\mathbb{V}\text{ar}(FE^i)-\hat{\sigma}_\tau^2\right) - \left(\mathbb{E}\left[(\pi_{t-1}-\pi_{t-1|t-1}^i)\pi_{t-1|t-1}^i\right]\right)^2\right] \\
=\ & (1-\kappa_x-\kappa_y)\left[\left(\mathbb{V}\text{ar}(FE^i)-2\hat{\sigma}_\tau^2+\frac{\sigma_u^2}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

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

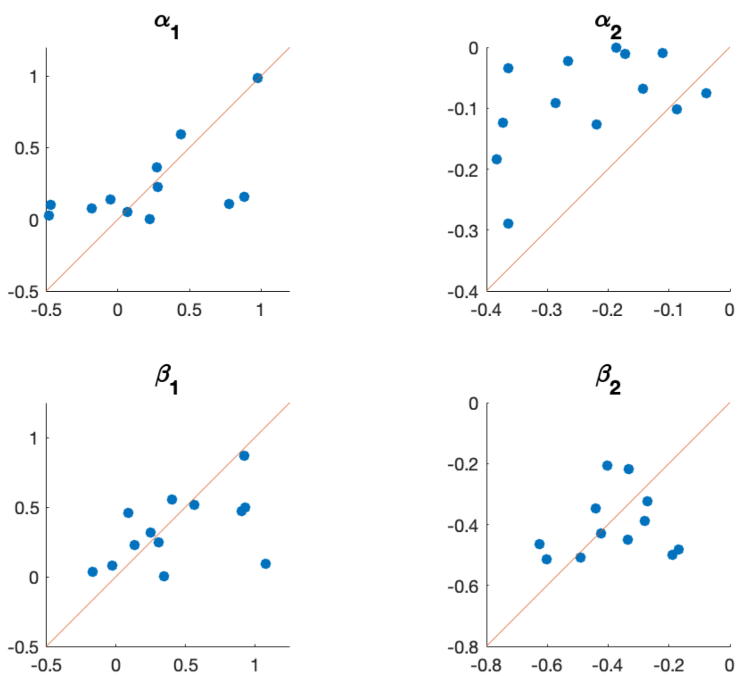
4. The sign of $\alpha_{2,h}$ follows the sign of $\mathcal{N}_{2,h}^\alpha$ (Eqn. C.29). Because of the inequality (D.4), Eqn. (C.29) implies $\alpha_{2,h} < 0$ under overconfidence.

□

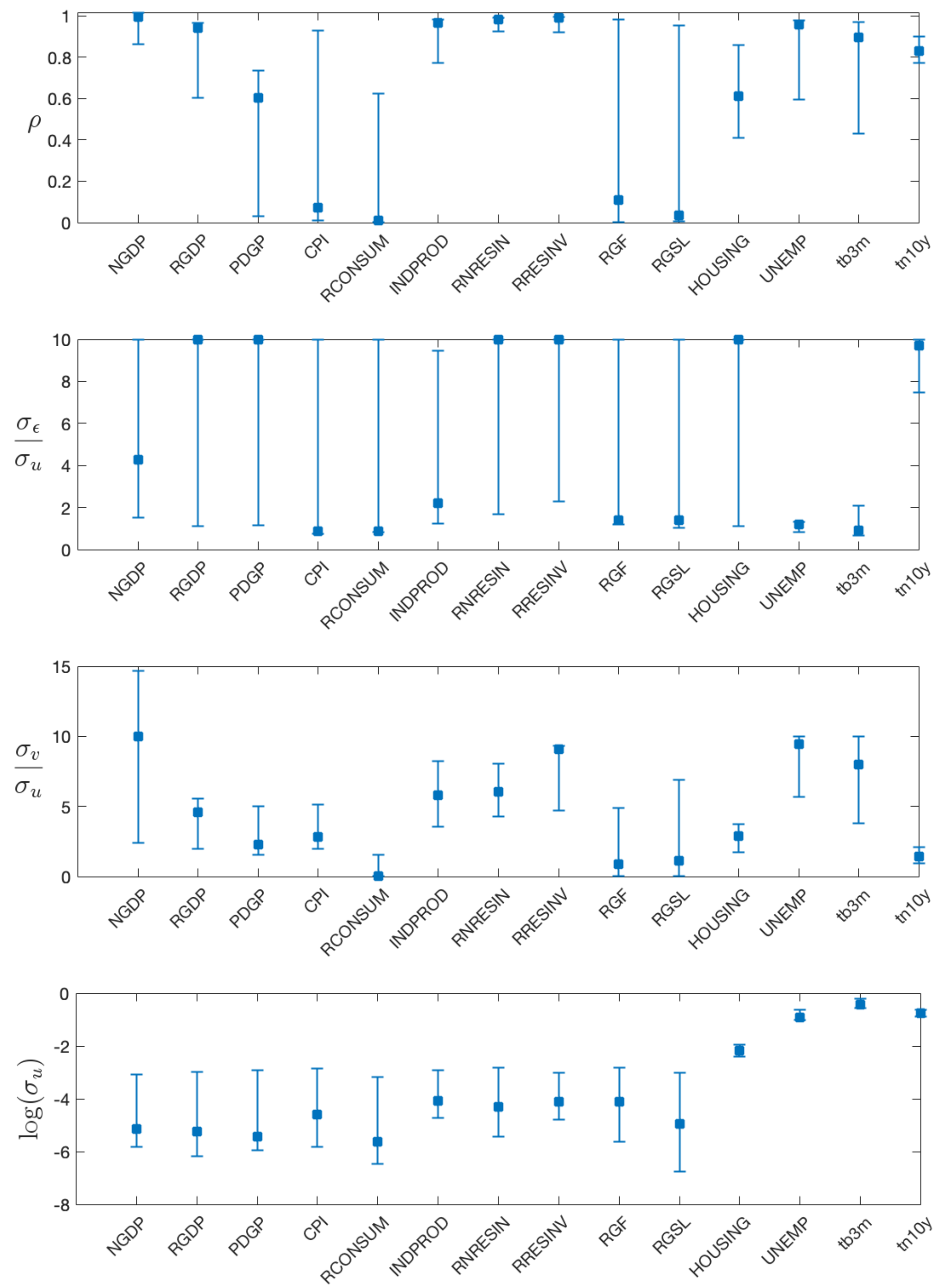
Appendix E Further parameter estimates and moment predictions of the estimated model from section 4



Appendix Figure E.14: DATA V.S. MODEL COEFFICIENTS FOR $h = 2$



Appendix Figure E.15: DATA V.S. MODEL COEFFICIENTS FOR $h = 3$



Appendix Figure E.16: ESTIMATED MODEL PARAMETERS OTHER THAN τ, WHICH IS SHOWN IN FIGURE 9 IN THE MAIN TEXT