Biased Surveys

Luca Gemmi[*] Rosen Valchev[†]

HEC Lausanne *Boston College*

Click here for the most recent version

August 30, 2022

Abstract

We provide evidence suggesting that survey of professional forecasters could be biased by strategic incentives. First, we find that individual professional forecasts over-react to private information but under-react to public information. Second, we show that this bias is absent in forecasts data not subject to strategic incentives, i.e. the central bank forecasts. We show that this is consistent with a theory of strategic diversification incentives in forecast reporting, where forecasters are rational but report a biased measure of their true expectations. This has two effects. First, it generates what looks like behavioral "over-reaction" in expectations, and second biases the information rigidity estimate further downward. Overall, our results caution against the use of survey of forecasts as a direct measure of expectations, and suggest that the true underlying beliefs are rational, but suffer from a much larger degree of imperfect information than previously thought. This has particularly profound implications for monetary policy, where inflation expectations play a key role.

Keywords: Expectations, Information, Forecasters, Reputation, Strategic diversification, Information rigidity, Surveys

JEL classification: C53, D82, D83, D84, E31

[*]Department of Finance, Faculty of Business and Economics (HEC Lausanne), University of Lausanne, 1015 Lausanne, CH; Email: luca.gemmi@unil.ch; website: sites.google.com/view/lucagemmi

[†]Department of Economics, Boston College, Chestnut Hill, MA 02467; Email: valchev@bc.edu; website: /sites.google.com/site/valchevr.

1 Introduction

Expectations play a crucial role in macroeconomic models, and hence the process through which agents form their expectation has been a fundamental, and often debated, topic. An important new development in the literature has emphasized the use of survey data, which holds the promise of providing direct, micro-level measurement of agent expectations. Using such data, Coibion and Gorodnichenko (2012, 2015) find significant evidence of incomplete and imperfect information, while another set of studies documents extensive predictability in individual forecast errors, which calls into question the classic paradigm of rational expectations itself (e.g. Bordalo et al. (2020)). Both strands of the literature, however, rely on the strong assumption that the information set of agents are contaminated with purely idiosyncratic errors, excluding any correlation in the noise of agent beliefs.

Our key findings are two-fold. First, we incorporate public information in the test for rational expectations and we find that, while individual professional forecasters tend to *over-react* to new information on average (in-line with previous findings of Bordalo et al. (2020)), the forecasts actually *under-react* to new public (i.e. common) information. Importantly, this finding is not consistent with some broadly used behavioral models as diagnostic expectations (Bordalo et al., 2018a).

Second, we show that this finding is in-line with models where strategic diversification incentives lead forecasters to provide a biased measure of their actual beliefs when responding to surveys (e.g. Ottaviani and Sørensen (2006)). First, we show that government expectations data do not exhibit the same over-reaction and under-reaction biases documented in professional forecasters surveys. Second, we document that professional forecasters disregard more signals about other forecasters beliefs than about fundamentals. Finally, to quantify this strategic diversification incentive and recover the true underlying expectations, we estimate a dynamic model of strategic incentives in reporting forecasts.

Our findings indicate that strategic incentives indeed play an important role, and hence caution against the use of survey forecasts as a direct measure of agent expectations. Specifically, the estimated model can fully account for the "over-reaction" puzzle in surveys that has received a lot of recent attention, suggesting that Rational Expectations is in fact a good model for the underlying true beliefs of agents. Moreover, the model estimates also show that the strategic incentives themselves bias the estimated information rigidity downward by a further 20% on average. Hence, our results indicate that expectations are rational after all, but the degree of imperfect information is significantly greater than previously thought.

In our empirical work we use data from the Survey of Professional Forecasters (SPF), which by now has become the common dataset for survey of macro forecasts. But while previous works considered a simple information structure with only private information, we start by documenting the importance of common noise in the forecast errors (i.e. public information) in survey expectations. We compare the seminal estimate of informational rigidity of Coibion and Gorodnichenko (2015), which is biased in presence of public information, with an unbiased estimate through a novel empirical strategy robust to public information (Goldstein, 2021). We document a systematic and large difference between the two measure for a large set of macroeconomic and financial variable, underlying the importance of public information. Such noise could be due to the incorporation of public signals in the forecasts, for example central bank's communications (e.g. Morse and Vissing-Jorgensen (2020)).

Then, we refine tests of rational expectations in survey data by also incorporating public signals and information in the benchmark regression specifications. We consider *lagged* consensus forecast as a proxy for public information, as it is both publicly available to all forecasters when they make their current forecasts and is also highly informative about the future realization of the variable being forecasted. We augment BGMS regression of individual forecast errors on individual forecast revisions by including our proxy for public signal. We find that individual forecasts in the SPF appear to *under-weight* the available public information while they over-weight the rest of their information set, interpreted as private information. This finding is not consistent with behavioral models that implies a homogeneous over- or under-reaction to new information, but consistent with strategic diversification, according to which pofessional forecasters over-weight private against public information to "stand-out from the crowd" (Ottaviani and Sørensen, 2006).

We present two additional evidence in support of the strategic diversification theory. First, we run the same regressions on the Tealbooks/Greenbooks projections from he Federal Reserve Board of Governors, which are about the same variables and same horizons, but not supposed to be subject to strategic incentives. In line with the theory, we don't find evidence of over-reaction to overall new information nor to differential over and underreaction to new private and public information. Second, we consider another type of public information – the past realization of the macroeconomic variable being forecasted (e.g. lagged inflation).This second type of public signal is under-weighted to a much smaller degree, which is again qualitatively consistent with the hypothesis of strategic incentives. Because the past consensus is not only a signal that eeveryone has access to, but is also a direct estimate of everyone else's recent beliefs, strategic diversification incentives imply that it will be doubly under-weighted. Thus, we provide an alternative, rational explanation of the over-reaction evidence, that is also consistent with additional, nuanced facts we uncover.

To rationalize our empirical findings, we build a global game model à la Morris and Shin (2002) with strategic substitutability, where the forecaster is balancing the desire to be right with the desire to stand-out. Intuitively, the forecaster would most like to both be right and also be the only person that gave a correct forecast, introducing strategic diversification incentives in forecast reporting as in Ottaviani and Sørensen (2006).[1] We assume agents have access to two types of noisy signals – a private signal with idiosyncratic noise, and a noisy public signal that is the same for everyone. We then show that, because of strategic substitutability incentives in responding to the survey, agents optimally decide to bias their response towards private information, leading to overreaction to private information and underreaction to public information, as we also find in the data. Moreover, we prove that in this setting it is always the case that individual forecasts appear as if they are *over-reacting* to new information on average, which can explain the recent findings in BGMS. Intuitively, because of agents' desire to stand out, when revising their expectations they put too high of a weight on their private signals which then results in forecastable errors that look like "over-reaction".

Finally, we estimate a dynamic, quantitative version of our model which allows us to back-out and measure the actual expectations of the forecasters, after removing the estimated bias due to strategic incentives. Our key results in this section are two-fold. First, we find that the reported consensus estimate is significantly more accurate than the true average belief – with the mean-squared error of the true average belief being roughly 30% to 100% higher, depending on the variable. This result is intuitive – the simultaneous

[1] This setting can be interpreted as a general version of a winner-take-all game, in which being accurate is rewarded but the prize is shared among correct forecasters (Ottaviani and Sørensen, 2006)

over-weighting of private information and under-weighting of public information acts as a positive externality in terms of the consensus estimate, as it limits the effects of common errors. Second, the true beliefs are also significantly less dispersed in the cross-section, with the cross-sectional standard deviation of beliefs being roughly 80% lower than the dispersion of the forecasts reported to SPF. This is also intuitive, and is a hallmark of the forecasters' attempts to "stand-out". It also speaks to the fact that the true disagreement and dispersion of beliefs is much lower than otherwise thought, and thus also consistent with an even higher degree of information stickiness.

Related literature This paper relates to three strands of the literature. First, papers using survey of professional forecasters to test the full information hypothesis. A common finding in this literature is consensus underreaction, meaning a positive relation between consensus forecast errors and consensus forecast revisions (Crowe, 2010; Coibion and Gorodnichenko, 2012, 2015). Similarly to Goldstein (2021), we highlight how public information bias the rigidity estimates in this literature downward and use a similar method to quantify this bias. In addition to this, we also highlight how the strategic incentives in forecast reporting biases the rigidity estimates even further, and use a structural model to estimate the actual information rigidity of honest beliefs.

Another strand of the literature uses surveys to test the rational expectation hypothesis. In particular, Bordalo et al. (2020) documents individual overreaction, meaning a negative relation between individual forecast errors and individual forecast revisions. As individual forecast errors should not be predictable using current information, the authors interpret this predictability as evidence of behavioral biases in belief formation. We show that this evidence can be explained by a departure from truthful revelation while preserving rational expectations. Moreover, we document underreaction to public information, which is consistent with a strategic incentive model but not with models of extrapolative beliefs. In a contemporaneous paper, Broer and Kohlhas (2018) also use public information to improve on the test of RE, and find mixed results in terms of under and over-reaction. The key difference is that in our empirical approach we isolate the surprise component of any given public signal, which leads to higher estimation precision, while they use the raw value of the public signal itself (which is correlated with other variables on the right-hand side of the main regressions). Moreover, we document how these biases appear in survey of professional forecasters, but not in non-strategic surveys as the Fed Greenbook.

A third group of papers investigate the role of strategic incentives in forecasters behavior (for a review, Marinovic et al. 2013). The most related is Ottaviani and Sørensen (2006), that propose two models of strategic substitutability and complementarity that leads forecasters to over or underweight private information in their reported forecast. While the spirit of the analysis is the same, we employ a more general Morris and Shin (2002) game and focus only on strategic substitutability. Moreover, we (i) empirically test this theory against alternative behavioral models, (ii) introduce public signals to distinguish between strategic incentive and behavioral theories and (iii) estimate a dynamic structural model to recover the underlying true expectations.

Overall, our results also speak to the fact that imperfect and noisy information is the dominant paradigm in the data, supporting earlier results on the importance of information rigidities in the expectation formation process, such as Kiley (2007), Klenow and Willis (2007), Korenok (2008), Dupor et al. (2010), Knotek II (2010), Coibion and Gorodnichenko (2012), and Coibion and Gorodnichenko (2015). In contrast to this literature, however, we also specifically identify and quantify the contribution of common noise components in the (imperfect) information sets of agents, and of the biasing effects of strategic incentives survey responders face when reporting expectations.

Structure of the paper The remaining sections of the paper are organized as follows. In Section 2 we describe the data and replicate the existing empirical evidence of underreaction of consensus forecast and overreaction of individual forecasts, and we highlight the importance of public information in survey forecasts. Then we document a novel fact: forecast underreact to new public information and overreact to new private information. We provide additional evidence that support theories departing fro truthful revelation, and not rational expectations, to explain this finding. In section 3 we develop a static model of strategic substitutability in forecast reporting which can rationalize the empirical evidence and provide the additional empirical implication,. i.e. contemporaneous underreaction to new public information and overreaction to new private information. In section 4 we extend the model to a dynamic setting and estimate it, allowing us to recover honest forecast and correctly measure information rigidity.

2 Empirical Analysis

Theoretical framework We consider a general framework of belief updating with dispersed information. In particular, consider a random variable x_t with some arbitrary autoregressive process. Agents in t provide a forecast for some horizon $t+h$ after observing a private signal and a public signal

$$\begin{aligned} g_t &= x_{t+h} + e_t \\ s_t^i &= x_{t+h} + \eta_t^i \end{aligned} \tag{1}$$

where η_t^i represents a normally distributed mean-zero noise with variance τ^{-1} which is i.i.d. across time and across agents, while e_t represents a normally distributed mean-zero noise with variance ν^{-1} which is i.i.d. only across time, but common across agents. Each agent generates forecast $\tilde{E}_t^i[x_{t+h}]$ at time t about the variable at h periods ahead according to

$$\tilde{E}_t^i[x_{t+h}] = \tilde{E}_{t-1}^i[x_{t+h}] + G_1(g_t - \tilde{E}_{t-1}^i[x_{t+h}]) + G_2(s_t^i - \tilde{E}_{t-1}^i[x_{t+h}]) \tag{2}$$

where G_1 is the weight agent put on the public signal, G_2 the weight agent put on the private signal and $\tilde{E}$ is a potentially non-optimal expectation operator. This general format embeds the rational Bayesian case, denoted as $\tilde{E}_t[x_t] = E_t[x_t]$, with $G_1 = \frac{\nu}{\tau+\nu+\Sigma^{-1}}$, $G_2 = \frac{\tau}{\tau+\nu+\Sigma^{-1}}$, and $\Sigma \equiv var(x_{t+h} - E_t^i[x_{t+h}])$.

Data on forecasts We collect data on forecasts from the Survey of Professional Forecasters (SPF), currently run by the Federal Bank Reserve of Philadelphia. In each 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 the quarter, and the forecasters are anonymous but identified by forecasters IDs.

The SPF covers macroeconomic and financial outcomes, providing both consensus forecast and an unbalanced panel of individual forecasts. These variables include GDP, price indices, consumption, investment, unemployment, government consumption, yields on government bonds and corporate bonds.

While most macroeconomic variables are provided in SPF in level, we follow BGMS and CG in transforming them into implied growth rate. Because of the timing of the survey, the actual variable realization in $t-1$ is known to the forecasters at the time of the forecasts. Therefore, we compute the forecasted growth rate of the variable from $t-1$ to $t+3$. We apply this method for GDP, price indices, consumption, investment and government consumption, while we keep the forecast in level for unemployment and financial variables.

We winsorize outliers by removing forecasts that are more than 5 interquartile ranges away from the median of each horizon in each quarter. We keep forecasters with at least 10 observations in all analyses. Consensus forecast are computed as the average of the individual forecasts available in each quarter. Appendix A provides a description of variable construction.

Data on actual outcomes The values of macroeconomics variables are released quarterly but subsequently revised. At the time of the survey, the forecasters can observe the first release of the values of the variables in $t-1$. To match as closely as possible the information set of the forecasters, we follow BGMS and use the initial releases of macroeconomics variables from Philadelphia Fed's Real-Time Data Set for Macroeconomics. Financial variables are not revised, so we use historical data from the Federal Reserve Bank of St. Louis.

We use actual realization to compute forecast errors, defined as actual realization minus forecast, and forecast revisions, defined as forecast in t on some horizon $t+h$ minus forecast in $t-1$ about the same horizon $t+h$.

Summary Statistics Table 1 presents the summary statistics for each series. Columns 1-5 reports the statistics for the consensus errors and revisions, including mean, standard deviation and standard errors. Forecast errors are statistically indistinguishable from zero for most of the series except for the interest rates, for which the forecasts are systematically above realizations. As argued by BGMS, this is likely due to the downward trend of the interest rates during the sample period, to which the forecast adjust only partially.

Columns 6-8 reports the summary statistics of the individual forecasts, including forecasts dispersion, share forecast with no meaningful revisions[2] and the probability that less than 80 percent of forecasters revise in the same direction. The large dispersion of forecasts and revisions at each point in time suggest a role for dispersed information among

[2] We follow BGMS in categorizing a non-missing forecast as a non meaningful revision if the forecasts change by less than 0.01 percentage points.

Table 1: Summary Statistics

	Consensus					Individual		
	Errors			Revisions		Forecast dispersion	Nonrev share	Pr($<$ 80% revise same direction)
	Mean	SD	SE	Mean	SD			
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nominal GDP	-0.26	1.69	0.19	-0.14	0.68	1.00	0.02	0.80
GDP price index inflation	-0.28	0.58	0.08	-0.08	0.25	0.49	0.07	0.85
Real GDP	-0.26	1.64	0.19	-0.16	0.58	0.78	0.02	0.74
Consumer Price Index	-0.08	1.04	0.15	-0.11	0.68	0.54	0.06	0.66
Industrial production	-0.83	3.94	0.46	-0.49	1.19	1.57	0.01	0.72
Housing Start	-3.36	17.79	2.20	-2.31	5.93	8.34	0.00	0.68
Real Consumption	0.32	1.10	0.15	-0.06	0.41	0.61	0.03	0.78
Real residential investment	-0.46	8.32	1.19	-0.61	2.33	4.37	0.04	0.87
Real nonresidential investment	0.20	5.60	0.79	-0.22	1.71	2.31	0.03	0.74
Real state and local government consumption	0.04	2.96	0.38	0.14	1.10	2.09	0.07	0.91
Real federal government consumption	0.02	1.10	0.15	-0.05	0.33	0.98	0.11	0.93
Unemployment rate	0.01	0.68	0.08	0.05	0.32	0.30	0.18	0.66
Three-month Treasury rate	-0.51	1.14	0.16	-0.19	0.51	0.43	0.15	0.59
Ten-year Treasury rate	-0.48	0.73	0.11	-0.12	0.36	0.37	0.11	0.55
AAA Corporate Rate Bond	-0.46	0.82	0.11	-0.11	0.38	0.49	0.09	0.66

Notes: Columns 1 to 5 show statistics for consensus forecast errors and revisions. Errors are defined as actuals minus forecasts, where actuals are the realized outcome corresponding to the variable forecasted. Revisions are forecast provided in t minus forecasts provided in $t-1$ about the same horizon. Columns 6 to 8 show statistics for individual forecasts, with Newey West (1994) standard errors. Forecast dispersion is the average standard deviation of individual forecasts at each quarter. The share of nonrevisions is the average quarterly share of instances in which forecast revision is less than 0.01 percentage points. The final column shows the fraction of quarters where less than 80 percent of the forecasters revise in the same direction.

forecasts, which we embed in our model. The share of non revisions is often small, contrary to a sticky-information model a la Mankiw and Reis (2002), and revisions go in different direction, suggestion a noisy information setting instead.

2.1 Motivational evidence

This paper is motivated by two sets of evidence on professional forecasters surveys. First, the overreaction to new information documented at the individual level. Second, the importance of public signal in forecasters information set. We summarize and extend these results from the previous literature under our more general theoretical framework.

2.1.1 Individual over-reaction to new information

Bordalo et al. (2018b), hereafter BGMS, test the rational expectation hypothesis by regressing individual forecast errors on consensus forecast revisions. Intuitively, if forecasters were fully rational, it should not be possible to predict individual future errors using today individual revisions, which would be part of the forecasters' information sets.

Individual forecast error is defined as the actual realization minus the individual forecast: $fe^i_{t+h,t} = x_{t+h} - \tilde{E}^i_t[x_{t+h}]$. Similarly, individual forecast revision is defined as the individual forecast provided today minus the forecast provided in the previous period about the same horizon: $fr^i_{t+h,t} = \tilde{E}^i_t[x_{t+h}] - \tilde{E}^i_{t-1}[x_{t+h}]$. BGMS run the following regression.

$$fe^i_{t+h,t} = \alpha + \beta_{BMGS} fr^i_{t+h,t} + err^i_t \tag{3}$$

Proposition 1 *If agents forecasts follow 2, the coefficient of regression 3 is given by:*

$$\beta_{BGMS} = \frac{1-G}{G} - \frac{\frac{G_1^2}{K}\nu^{-1} + \frac{G_2^2}{G}\tau^{-1}}{G^2\tilde{\Sigma} + G_1^2\nu^{-1} + G_2^2\tau^{-1}} \tag{4}$$

with $\tilde{\Sigma} \equiv var(x_t - \tilde{E}^i_{t-1}[x_t])$ *(in the special case of AR(1) process,* $\tilde{\Sigma} = \frac{\rho^2[G_1^2\nu^{-1}+G_2^2\tau^{-1}]+\xi^{-1}}{1-\rho^2(1-G)^2}$*).*

Corollary 1 *If agents forecasts follow 2, under rational expectation the coefficient* β_{BGMS} *of regression 3 is equal to zero.*

According to RE, individual forecast errors should not be predictable using individual forecast revisions. A positive $\beta_{BMGS} > 0$ would imply that after an positive surprise today agents don't update their forecast enough and they consistently underestimate the future

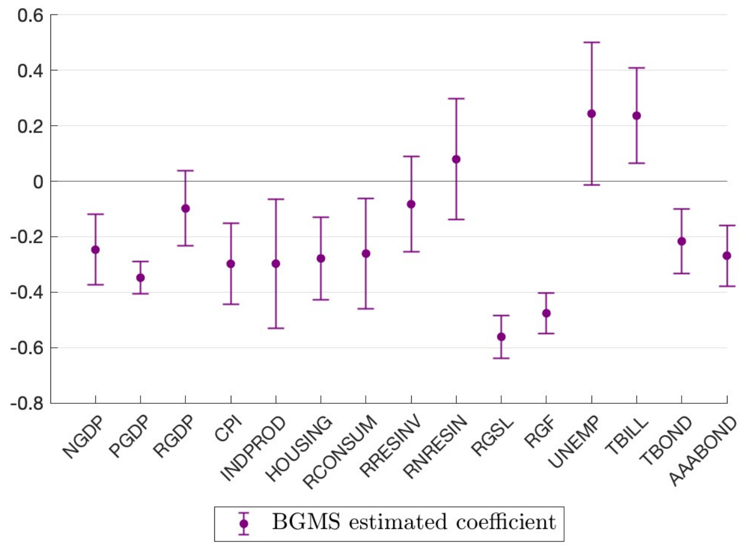

Figure 1: Forecast errors on forecast revisions and public information

Notes: this figure plots the coefficient from panel data regression 3 with individual fixed effect and horizon h=3. Bars reports the 90% confidence interval for the estimated coefficients. Standard errors are robust and clustered at both time and individual forecaster level.

value of x. On the opposite, a negative $\beta_{BMGS} < 0$ would imply that after an positive surprise today agents become too optimistic and they consistently overestimate the future value of x.

BMGS documents a robust $\beta_{BMGS} < 0$ for a wide range of macroeconomic and financial series. We replicate their panel data econometric specification with individual fixed effects in columns 1-3 of Table 2. However, the panel specification could introduce a bias in the coefficient.[3] Therefore we present also the median coefficients from the individual level regressions in the last column of Table 2, which confirms the panel results.[4]

2.1.2 Public information in forecasters information set

Coibion and Gorodnichenko (2015), hereafter CG, test the full information rational expectation hypothesis by regressing consensus forecast errors on consensus forecast revisions. While this regression is able to identify the degree of information dispersion (i.e. pri-

[3] RE implies that it is not possible for agents to predict their own forecast errors, $\beta_{BGMS} = 0$. However since the panel regression exploits the cross sectional variance in addition to the time series one, this specification effectively uses the average information set to pin down β_{BMGS} and not only the individual one.

[4] BGMS documents that overreaction holds also under the assumption that the fundamental process follows an AR(2). We replicate their finding in table 13 in appendix D.

Table 2: Individual errors on revisions

	3 quarters horizon				2 quarters horizon			
	β_{BGMS}	SE	p-value	Median	β_{BGMS}	SE	p-value	Median
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nominal GDP	-0.25	0.08	0.00	-0.19	-0.11	0.06	0.10	-0.08
GDP price index inflation	-0.35	0.04	0.00	-0.35	-0.25	0.04	0.00	-0.26
Real GDP	-0.10	0.08	0.24	0.07	-0.07	0.10	0.45	0.02
Consumer Price Index	-0.30	0.09	0.00	-0.29	-0.24	0.07	0.00	-0.24
Industrial production	-0.30	0.14	0.04	-0.31	-0.01	0.10	0.94	0.03
Housing Start	-0.28	0.09	0.00	-0.28	0.12	0.05	0.03	0.07
Real Consumption	-0.26	0.12	0.04	-0.24	-0.16	0.08	0.07	-0.16
Real residential investment	-0.08	0.10	0.44	-0.07	0.07	0.08	0.41	0.02
Real nonresidential investment	0.08	0.13	0.56	0.15	0.10	0.07	0.18	0.10
Real state and local government consumption	-0.56	0.05	0.00	-0.52	-0.30	0.05	0.00	-0.26
Real federal government consumption	-0.48	0.04	0.00	-0.40	-0.28	0.04	0.00	-0.27
Unemployment rate	0.24	0.16	0.13	0.19	0.20	0.11	0.09	0.20
Three-month Treasury rate	0.24	0.10	0.03	0.29	0.14	0.08	0.09	0.21
Ten-year Treasury rate	-0.22	0.07	0.01	-0.24	-0.24	0.09	0.01	-0.27
AAA Corporate Rate Bond	-0.27	0.07	0.00	-0.32	-0.22	0.06	0.00	-0.29

Notes: The table reports the coefficients from the BGMS regression (individual forecast errors on individual revisions). Columns (1) to (3) shows the panel data with fixed effect coefficient with standard errors and corresponding p-values. Standard errors are robust to heteroskedasticity and clustered at both forecaster and time level. Columns (7) shows the median coefficient of the BGMS regression at the individual level.

vate noise) among forecasters, it can't identify the degree of total information stickiness (i.e. private and public noise). However, a new empirical strategy employing individual forecasts developed in Goldstein (2021) allow to correctly identify the total information stickiness. We compare the two estimated coefficient to document the importance of public information in survey forecasts.

Consensus forecast error is defined as the average forecast error across forecasters, $\bar{fe}_{t+h,t} = \int_i fe^i_{t+h,t}di$, and similarly for consensus forecast revision, $\bar{fr}_{t+h,t} = \int_i fr^i_{t+h,t}di$. CG run the following regression

$$\bar{fe}_{t+h,t} = \alpha + \beta_{CG}\bar{fr}_{t+h,t} + err_t \tag{5}$$

In order to interpret the result, derive the structural equivalent from 2:

$$\bar{fe}_{t+h,t} = \frac{1-G}{G}\bar{fr}_{t+h,t} - \frac{G_1}{G}\rho^h e_t \tag{6}$$

First, consider a setting without public information, i.e. $G_1 = 0$. In this case, $\beta_{CG} = \frac{1-G}{G}$.

A $\hat{\beta}_{CG} = 0$ would imply $G = 1$, meaning forecast adjust completely to new information, as implied by the FIRE hypothesis. On the other hand, $\hat{\beta}_{CG} > 0$ would imply $G < 1$, meaning stickiness in forecast updating, as in the noisy information setting. Therefore, in absence of public information, $\hat{\beta}_{CG} > 0$ rejects full information models (but not necessarily rational expectation).[5] While this intuition is accurate in absence of public information, it is not in the presence of it, i.e. $G_1 > 0$. Because of the bias introduced by the public noise in the regression error, $\hat{\beta}_{CG}$ does not identify the information gain G.

Proposition 2 *If agents forecasts follow 2, the coefficient from regression 5 is given by:*

$$\beta_{CG} = \frac{\tilde{\bar{\Sigma}} - [G\tilde{\bar{\Sigma}} + \frac{G_1^2}{G}\nu^{-1}]}{G\tilde{\bar{\Sigma}} + \frac{G_1^2}{G}\nu^{-1}} \tag{7}$$

with $\tilde{\bar{\Sigma}} \equiv var(x_t - \tilde{\bar{E}}_{t-1}[x_t])$ *(in the special case of AR(1) process,* $\tilde{\bar{\Sigma}} = \frac{\rho^2[G_1^2\nu^{-1}]+\xi^{-1}}{1-\rho^2(1-G)^2}$*). If* $G_1 = 0$*,* $\beta_{CG} = \frac{1-G}{G}$.

Corollary 2 *If agents forecasts follow 2, under rational expectation the coefficient* β_{CG} *of regression 3 is equal to zero when the forecasts are equal among forecasters. This happens when private info is not informative,* $\tau = 0$*, or when agents are fully informed,* $\tau \to \infty$ *or* $\nu \to \infty$*.*

Intuitively, the CG regression coefficients identifies the degree of information *dispersion* among forecasters (i.e. private information), not overall information *stickiness* (i.e. gain G).[6]

Corollary 3 *If agents forecasts follow 2, the difference between* $\hat{G}_{CG} = \frac{1}{1+\hat{\beta}_{CG}}$*, where* $\hat{\beta}_{CG}$ *is the coefficient estimated from regression 5, and the actual gain* $G = G_1 + G_2$ *from equation 2 is given by:*

$$\hat{G}_{CG} - G = G\left(\frac{G_1^2\nu^{-1}}{G^2\tilde{\bar{\Sigma}}}\right) > 0 \tag{8}$$

[5] Intuitively, in a dispersed information setting individual forecasters do not observe the information of the others, and therefore the average forecast revisions can predict average forecast errors. Because of private noise the individual signal is more noisy and less accurate that the average signal, even if each individual update their forecast optimally given their signal the average forecast is suboptimally sticky with respect to the average signal.

[6] If there were no private but only public information, (no *dispersion*, $G_2 = 0, G_1 > 0$), under rational expectations the coefficient would still be $\hat{\beta}_{CG} = 0$ even if $G < 1$.

with $\bar{\tilde{\Sigma}} \equiv var(x_t - \tilde{E}_{t-1}[x_t])$ *(in the special case of AR(1) process,* $\bar{\tilde{\Sigma}} = \frac{\rho^2[G_1^2\nu^{-1}]+\xi^{-1}}{1-\rho^2(1-G)^2}$*).*

We compare the CG implied stickiness with an unbiased measure of the actual stickiness to document the importance of public information, G_1. We employ and estimation procedure which relies on the difference between individual forecasts and the average forecast to get rid of the common noise component. Rewrite 2 in terms of forecast revision on forecast errors at some horizon h and demean it using consensus revisions.[7]

$$(fr^i_{t+h,t}) - (\bar{f}r_{t+h,t}) = G(\bar{\tilde{E}}_{t-1}[x_{t+h}] - \tilde{E}^i_{t-1}[x_{t+h}]) - G_2\rho^h\eta^i_t \tag{10}$$

Equation 10 provide an unbiased strategy to measure information stickiness. The coefficient estimated by regressing the difference between individual and consensus forecast revision on the difference between individual and consensus prior converge to the posted weight on new information G. In particular, consider the regression

$$fr^i_{t+h,t} - \bar{f}r_{t+h,t} = \beta(\bar{\hat{x}}_{t+h,t-1} - \hat{x}^i_{t+h,t-1}) + err^i_t \tag{11}$$

the OLS coefficient $\hat{\beta}$ is an efficient estimator of gain G. This approach is more general than the CG regression as it doesn't rely on the assumption of no public information.

We run regression 11 in a panel data with time fixed effect to demean forecast revisions and priors at each quarter. Table 3 reports the estimated coefficient, standard errors and p-value of the panel data regression, and the median coefficient from demeaned individual regressions. We estimate the gains for both 3 quarters and 2 quarters horizons. There are two important observations. First, our estimated gains are relatively stable across variables at the same horizon. Second, the gains are systematically larger at the shorter horizon, consistently with the idea of more accurate information about shorter horizons.

Table 4 reports the estimated gain G from regression 11 in columns (1) -(2) and the

[7] In the particular setting laid down in 1 and 2, the demeaning is not strictly needed, as you can rewrite 2 as $fr^i_{t+h,t} = G(fs^i_{t+h,t}) + G_1 e_t + G_2\eta^i_t$. The coefficient from regressing individual forecast revisions on individual surprise would converge to the gain G. However, this would not be true if the variable forecasted has some unlearnable error component, e.g. $y_t = x_t + \epsilon_t$, where x_t follows **??**. In this case, 2 would become

$$fr^i_{t+h,t} = G(fs^i_{t+h,t}) + G_1 e_t + G_2\eta^i_t - \epsilon_{t+h} \tag{9}$$

The estimated coefficient from regressing individual forecast revisions on individual surprise would be biased by the correlation between fs^i_{t+h} and ϵ_{t+h} to $t+h$. The deameaning takes care of this.

Table 3: Stickiness estimation

	3 quarters horizon				2 quarters horizon			
	β	SE	p-value	Median	β	SE	p-value	Median
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nominal GDP	0.53	0.02	0.00	0.49	0.61	0.01	0.00	0.62
GDP price index inflation	0.49	0.03	0.00	0.52	0.63	0.02	0.00	0.68
Real GDP	0.56	0.03	0.00	0.52	0.63	0.02	0.00	0.62
Consumer Price Index	0.49	0.02	0.00	0.53	0.70	0.02	0.00	0.71
Industrial production	0.50	0.03	0.00	0.52	0.59	0.02	0.00	0.63
Housing Start	0.49	0.03	0.00	0.55	0.53	0.02	0.00	0.56
Real Consumption	0.49	0.03	0.00	0.48	0.63	0.03	0.00	0.62
Real residential investment	0.41	0.03	0.00	0.44	0.56	0.02	0.00	0.64
Real nonresidential investment	0.48	0.02	0.00	0.49	0.61	0.03	0.00	0.61
Real state and local government consumption	0.43	0.04	0.00	0.40	0.60	0.05	0.00	0.56
Real federal government consumption	0.47	0.04	0.00	0.48	0.62	0.03	0.00	0.62
Unemployment rate	0.49	0.02	0.00	0.54	0.56	0.02	0.00	0.62
Three-month Treasury rate	0.55	0.02	0.00	0.59	0.63	0.03	0.00	0.67
Ten-year Treasury rate	0.51	0.02	0.00	0.54	0.60	0.02	0.00	0.63
AAA Corporate Rate Bond	0.54	0.02	0.00	0.56	0.61	0.02	0.00	0.62

Notes: The table shows the result from regression 11. Columns (1)-(3) report coefficients, standard errors and p-values from the panel data regression with time and individual fixed effect. Column (4) reports the median coefficient from individual regressions. Columns (5)-(8) reports the same statistics for the 2 quarters horizon. Standard errors are robust to heteroskedasticity and clustered at both forecaster and time level.

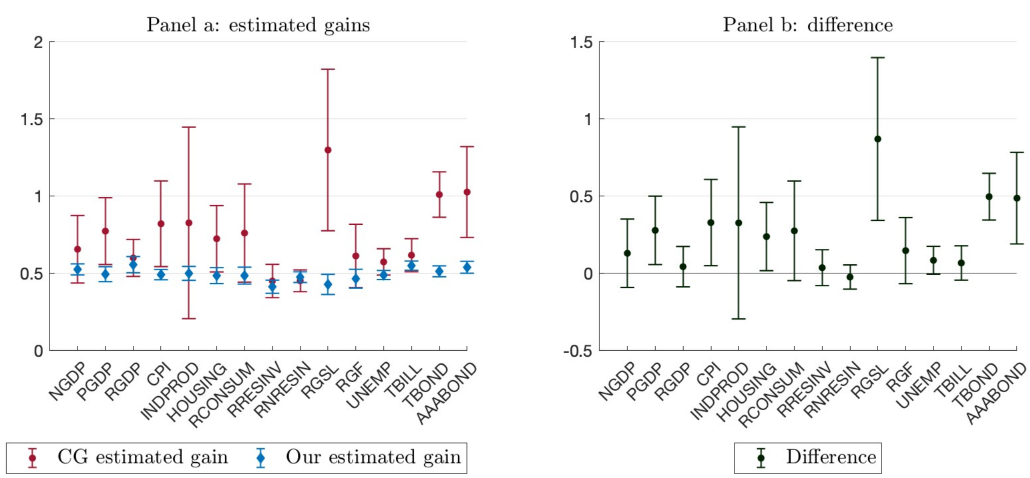

Figure 2: Estimated gains: our measure vs CG

Notes: Panel a: the red circles represent the implied gain from estimated coefficients in regression 5 at horizon h=3. Standard errors are robust to heteroskedasticity and Newey-West with the automatic bandwidth selection procedure of Newey and West (1994). The blue diamonds represent the gain estimated from 11 with individual fixed effect at horizon h=3. Standard errors are robust to heteroskedasticity and clustered at both forecaster and time level. Bars reports the 90% confidence interval for the estimated coefficients. Panel b: difference between the gains reported in panel a.

Table 4: Difference between estimated gains

Variable	G_{CG} (1)	SE (2)	G (3)	SE (4)	Difference (5)	SE (6)	p-value (7)
Nominal GDP	0.66	0.13	0.53	0.02	0.13	0.13	0.17
GDP price index inflation	0.77	0.13	0.49	0.03	0.28	0.13	0.02
Real GDP	0.60	0.07	0.56	0.03	0.04	0.08	0.29
Consumer Price Index	0.82	0.17	0.49	0.02	0.33	0.17	0.03
Industrial production	0.83	0.38	0.50	0.03	0.33	0.38	0.19
Housing Start	0.72	0.13	0.49	0.03	0.24	0.13	0.04
Real Consumption	0.76	0.19	0.49	0.03	0.28	0.20	0.08
Real residential investment	0.45	0.07	0.41	0.03	0.04	0.07	0.30
Real nonresidential investment	0.45	0.04	0.48	0.02	-0.02	0.05	0.69
Real state and local government consumption	1.30	0.32	0.43	0.04	0.87	0.32	0.00
Real federal government consumption	0.61	0.12	0.47	0.04	0.15	0.13	0.13
Unemployment rate	0.57	0.05	0.49	0.02	0.08	0.05	0.06
Three-month Treasury rate	0.62	0.07	0.55	0.02	0.07	0.07	0.16
Ten-year Treasury rate	1.01	0.09	0.51	0.02	0.50	0.09	0.00
AAA Corporate Rate Bond	1.03	0.18	0.54	0.02	0.49	0.18	0.00

Notes: Columns (1)-(2) reports the implied gain from CG regressions of table **??**. Columns (3)-(4) replicate the gain estimate in table 3. Columns (5)-(8) reports the difference between column (1) and (3), its standard error and the probability of rejecting the null of column (5) lower or equal to zero.

gain G_{CG} implied by CG estimate 5 in absence of public information in columns (3)-(4). Figure 2 panel a shows the comparison graphically. Our estimate gain is less volatile across variables and consistently lower than the one implied by CG. We report the difference $G_{CG} - G$ in Table 4, columns (5)-(6), and plot it graphically in figure 2 panel b. The difference is consistently positive as implied by proposition 2. We test whether the difference is statistically larger than zero and report the p-value in column (7). The null hypothesis is rejected at the 10% confidence level for 7 variables out of 15.

The evidence indicate that public information is in fact an important part of the information set of forecasters. While the CG estimate implies very different gains across variables, with some series with no apparent stickiness (Ten-year Treasury rate, AAA Corporate Bond and Real federal government consumption), our novel approach suggests that the overall gain on new information is instead similar across variables but with differences in the role of public information. In particular, public exceed private information in importance for financial variable, consistently with the idea that most of private information is priced in an efficient market.

2.2 Under-reaction to public information in professional forecast

Motivated by the importance of public information in forecasters information set, we distinguish individual forecast reaction to private and public information. We document that while individual forecasts over-react to overall new information, they under-react to new *public* information.

In order to measure public information, we use the lagged consensus forecast, namely the average of the individual forecasts provided in the previous quarter about the same horizon. The consensus forecast is available to the forecasters at the time of the survey. To capture the surprise component in the public information, we compute the difference between the public signal and individual prior about the signal: $pi_{t,t+h} = g_t - \tilde{E}^i_{t-1}[x_{t+h}]$. We run the following regression:

$$fe^i_{t+h,t} = \alpha + \beta_1 fr^i_{t+h,t} + \beta_2 pi_{t+h,t} + err^i_t \tag{12}$$

where g_t is a public signal providing information about the variable at horizon $t+h$.

Proposition 3 *If agents forecasts follow 2, the coefficients of regression 12 are given by:*

$$\begin{aligned}
\hat{\beta}_1 &= \frac{1-G_2}{G_2} - \frac{(\tilde{\Sigma}+\nu^{-1})G_2\frac{1}{\tau}}{(\tilde{\Sigma}+\nu^{-1})(G^2\tilde{\Sigma}^{-1}+G_1^2\nu^{-1}+G_2^2\tau^{-1})-(G\tilde{\Sigma}+G_1\nu^{-1})^2} \\
\hat{\beta}_2 &= -\frac{G_1}{G_2} + \frac{(G\tilde{\Sigma}+G_1\nu^{-1})G_2\frac{1}{\tau}}{(\tilde{\Sigma}+\nu^{-1})(G^2\tilde{\Sigma}+G_1^2\nu^{-1}+G_2^2\tau^{-1})-(G\tilde{\Sigma}+G_1\nu^{-1})^2}
\end{aligned} \tag{13}$$

with $\tilde{\Sigma} \equiv var(x_t - \tilde{E}^i_{t-1}[x_t])$ *(in the special case of AR(1) process,* $\tilde{\Sigma} = \frac{\rho^2[G_1^2\nu^{-1}+G_2^2\tau^{-1}]+\xi^{-1}}{1-\rho^2(1-G)^2}$*).*

Corollary 4 *If agents forecasts follow 2, under rational expectation the coefficient* β_1 *and* β_2 *of regression 12 are equal to zero.*

Under RE and truthful revelation, it would not be possible to predict individual errors using individual information sets. Panel A of Table 5 reports the panel data regressions with individual fixed effects and the median from individual regressions. Both specifications display a consistent $\beta_1 < 0$ and $\beta_2 > 0$ across variables, with few exceptions, meaning individual overreaction to private information and underreaction to public information. Figure 3 provide a graphical representation of the estimated coefficient from the

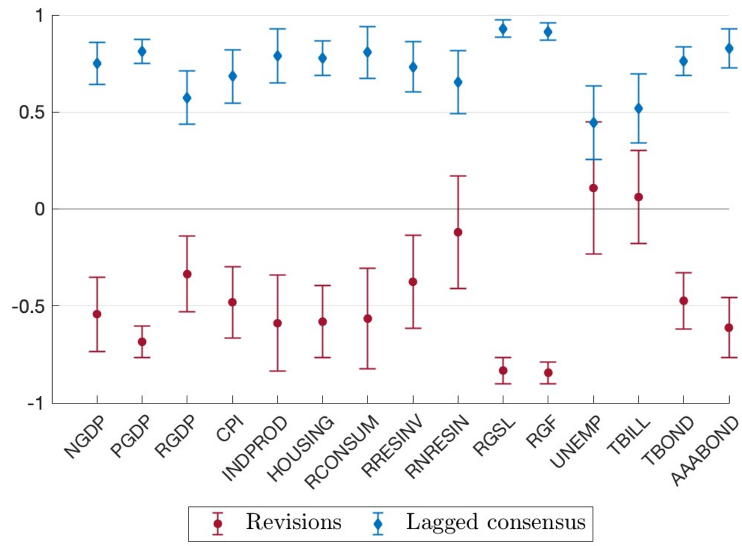

Figure 3: Forecast errors on forecast revisions and public information

Notes: this figure plots the coefficient from panel data regression 12 with individual fixed effect and horizon h=3. The red circles represent the coefficient β_1 while the blue diamonds represent the coefficient β_2. Bars reports the 90% confidence interval for the estimated coefficients. Standard errors are robust and clustered at both time and individual forecaster level.

panel data regression. In table 14 in appendix D we show that these results holds also under the assumption that the fundamental follows an AR(2) process.

Discussion While the overreaction to new information documented in the literature seems to indicates a departure for the rational expectation hypothesis, we document that forecasters overreact to the first but underreact to the second. This is not consistent with model of overreaction to all new information as diagnostic expectations Bordalo et al. (2020) or natural extrapolation (...), but it is consistent with other two distinct frameworks.

First, strategic diversification in forecast reporting, according to which forecasters form their beliefs rationally but do not truthfully reveal their beliefs in the survey due to strategic incentives. Their objective is not only want to provide accurate forecasts, but also to stand out with respect to the average forecast (Ottaviani and Sørensen, 2006). The likelihood of strategic interactions in reporting is known in the forecasting literature. For example, Croushore (1997) suggests that “some [survey] participants might shade their forecasts more toward the consensus (to avoid unfavorable publicity when wrong), while others

Table 5: Private and public information

Panel A: 3 quarters horizon

	Revision				Public signal			
	β_1	SE	p-value	Median	β_2	SE	p-value	Median
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nominal GDP	-0.54	0.12	0.00	-0.44	0.75	0.07	0.00	0.76
GDP price index inflation	-0.68	0.05	0.00	-0.64	0.81	0.04	0.00	0.83
Real GDP	-0.34	0.12	0.01	-0.18	0.57	0.08	0.00	0.61
Consumer Price Index	-0.48	0.11	0.00	-0.46	0.68	0.08	0.00	0.69
Industrial production	-0.59	0.15	0.00	-0.60	0.79	0.08	0.00	0.78
Housing Start	-0.58	0.11	0.00	-0.53	0.78	0.05	0.00	0.71
Real Consumption	-0.57	0.16	0.00	-0.58	0.81	0.08	0.00	0.81
Real residential investment	-0.38	0.15	0.01	-0.39	0.73	0.08	0.00	0.66
Real nonresidential investment	-0.12	0.18	0.50	-0.10	0.65	0.10	0.00	0.51
Real state and local government consumption	-0.83	0.04	0.00	-0.81	0.93	0.03	0.00	0.89
Real federal government consumption	-0.84	0.03	0.00	-0.77	0.91	0.03	0.00	0.87
Unemployment rate	0.11	0.21	0.61	-0.02	0.44	0.11	0.00	0.42
Three-month Treasury rate	0.06	0.15	0.68	0.11	0.52	0.11	0.00	0.38
Ten-year Treasury rate	-0.47	0.09	0.00	-0.40	0.76	0.04	0.00	0.83
AAA Corporate Rate Bond	-0.61	0.09	0.00	-0.67	0.83	0.06	0.00	0.87

Panel B: 2 quarters horizon

	Revision				Public signal			
	β_1	SE	p-value	Median	β_2	SE	p-value	Median
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nominal GDP	-0.35	0.09	0.00	-0.27	0.62	0.06	0.00	0.63
GDP price index inflation	-0.55	0.06	0.00	-0.50	0.70	0.04	0.00	0.66
Real GDP	-0.26	0.13	0.05	-0.14	0.54	0.08	0.00	0.54
Consumer Price Index	-0.38	0.09	0.00	-0.36	0.52	0.08	0.00	0.52
Industrial production	-0.16	0.12	0.19	-0.14	0.49	0.08	0.00	0.50
Housing Start	-0.15	0.08	0.08	-0.15	0.54	0.05	0.00	0.56
Real Consumption	-0.37	0.11	0.00	-0.29	0.64	0.07	0.00	0.69
Real residential investment	-0.13	0.11	0.23	-0.16	0.49	0.07	0.00	0.43
Real nonresidential investment	-0.02	0.10	0.81	-0.04	0.41	0.07	0.00	0.44
Real state and local government consumption	-0.63	0.08	0.00	-0.51	0.79	0.04	0.00	0.72
Real federal government consumption	-0.71	0.06	0.00	-0.64	0.80	0.04	0.00	0.74
Unemployment rate	0.09	0.15	0.57	0.03	0.39	0.10	0.00	0.37
Three-month Treasury rate	0.02	0.11	0.89	0.10	0.48	0.10	0.00	0.39
Ten-year Treasury rate	-0.46	0.11	0.00	-0.45	0.71	0.07	0.00	0.67
AAA Corporate Rate Bond	-0.49	0.08	0.00	-0.52	0.70	0.06	0.00	0.71

Notes: this table reports the coefficients of regression 12 (individual forecast errors on individual revisions and public information). Columns 1 to 3 show coefficient β_1 (forecast revision) from the panel regression with individual fixed effect, with standard errors and corresponding p-values. Standard errors are robust and clustered by forecaster. Column 4 shows the median coefficient of the same regression at the individual level. Columns 5 to 7 show coefficient β_2 (public information) from the panel regression with individual fixed effect, with standard errors and corresponding p-values. Standard errors are robust and clustered by forecaster. Column 8 shows the median coefficient of the same regression at the individual level. Panel A uses forecast at 3 quarters horizon and panel B uses forecast at 2 quarters horizon.

might make unusually bold forecasts, hoping to stand out from the crowd".[8] Strategic diversification considerations are therefore more important for professional forecasters and imply a larger underreaction to signals on average forecasters consensus.

Second, behavioral overconfidence, according to which agents do not update beliefs rationally, but overestimate the actual precision of their private signal against public signals (among the others, Daniel et al. (1998); Eyster et al. (2019); Broer and Kohlhas (2018)). While these theories differ in terms of how and which signal agents under and overreact to, they apply to both professional and non-professional forecasters, differently to strategic considerations which apply only to professional forecasters

In the following section we provide empirical evidence supporting the strategic diversification theory against alternative behavioral theories. First, we show that non-professional forecasters survey forecasts do not exhibit the same biases documented in professional forecast survey forecasts. Second, we show that the level underreaction to public information of professional forecasters is larger for signal directly related to forecasters consensus beliefs.

2.3 Comparison between professional and government forecast

In this section we distinguish between behavioral and strategic diversification theories by comparing professional and non-professional forecasters survey. In particular, we use the Tealbooks/Greenbooks forecasts of the Federal Reserve Board of Governors to measure non-professional forecasters survey and show that these forecasters do not display the biases documented in professional forecasters surveys (overreaction to new information and underreaction to public information). This finding is consistent with the strategic diversification theory, which applies to professional but not non-professional forecasters, but not with common behavioral theories, which make no distinction.

We collect data on forecasts from the Tealbooks/Greenbooks (GB) projection of the Federal Reserve Board of Governors, provided by the Federal Reserve Bank of Philadelphia. Using an assumption about monetary policy, the Research staff at the Board of Governors prepares projections about how the economy will fare in the future. These projections are made available to the public after a lag of five years.

[8] While strategic considerations apply more intuitively to non-anonymous survey, in Appendix F we argue that they apply to anonymous survey as well, as forecasters are likely to provide the same forecast to both surveys.

Table 6: Summary Statistics: non-professional forecasters

	Errors			Revisions	
	Mean	SD	SE	Mean	SD
Variable	(1)	(2)	(3)	(4)	(5)
Nominal GDP	-0.06	1.46	0.16	-0.10	0.87
GDP price index inflation	-0.07	0.58	0.09	0.06	0.47
Real GDP	-0.13	1.55	0.18	-0.15	0.82
Consumer Price Index	0.15	1.08	0.16	0.00	0.69
Industrial production	-0.94	3.58	0.38	-0.46	1.69
Housing Start	-2.14	15.64	1.83	-2.97	9.51
Real residential investment	0.60	7.13	0.87	-0.92	4.85
Real nonresidential investment	0.68	4.93	0.62	-0.36	2.84
Real state and local government consumption	0.87	3.06	0.34	0.22	1.70
Real federal government consumption	-0.19	1.30	0.17	-0.08	0.89
Unemployment rate	-0.09	0.59	0.06	0.01	0.41

Notes: Columns 1 to 5 show statistics for consensus forecast errors and revisions using forecasts from the Federal Reserve Green Book dataset. Errors are defined as actuals minus forecasts, where actuals are the realized outcome corresponding to the variable forecasted. Revisions are forecast provided in t minus forecasts provided in $t-1$ about the same horizon.

While the projections are produced before each meeting of the Federal Open Market Committee, we group them at quarterly level by keeping only the last forecasts of the quarter. The forecasts are produced for up to 9 quarters ahead in the future, but in order to compare them with the SPF we keep up to 4 quarters ahead. The variables forecasted are 15, but we keep only the 11 appearing also in the SPF (see tables 6).

While most macroeconomic variables are provided in quarter-on-quarter growth, we follow the same procedure as with the SPF and transform them into forecast about annual growth rate from $t-1$ to $t+3$ (with the exception of unemployment) . We compare forecasts withe same actual data we used for the SPF, described in the section above. Appendix A provides a description of variable construction.

Summary Statistics Table 6 presents the summary statistics for each series. Columns 1-5 reports the statistics for the consensus errors and revisions, including mean, standard deviation and standard errors. Forecast errors are statistically indistinguishable from zero for most of the series except for industrial production and real real non-residential investment.

No over-reaction in individual forecasts Figure 4 reports the coefficients estimated by running regression 3 on the SPF forecast data (from figure 1) and on the GB forecast data. It shows that in large part the non-professional GB forecasts do not exhibit the same overreaction bias to new information exhibited by professional SPF data (with the

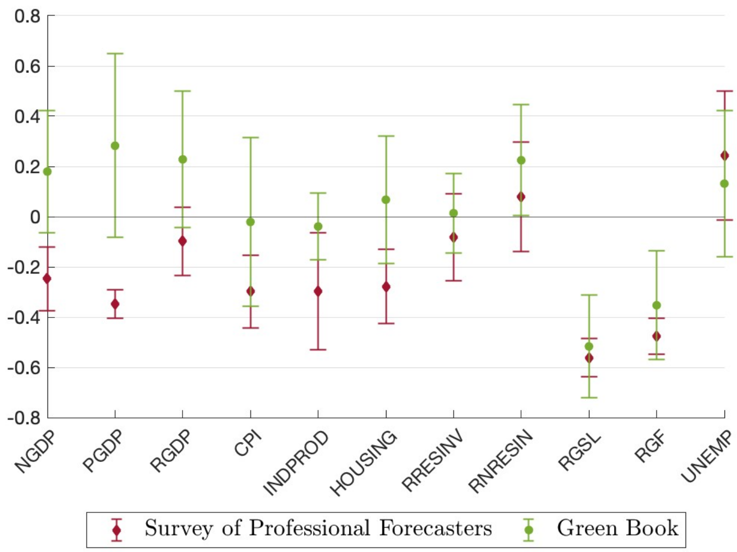

Figure 4: Forecast errors on forecast revisions and public information

Notes: This figure compare reaction to new information in the Fed Survey of Professional Forecasters (*professional*) and the Fed Green Book (*non-professional*) at horizon h=3. (i) In red the coefficient from regression 3 with individual fixed effect using the Survey of Professional Forecasters. Standard errors are robust and clustered at both time and individual forecaster level. (ii) In green the coefficient from regression 3 using the Fed Green Book. Bars reports the 90% confidence interval for the estimated coefficients. Standard errors are robust to heteroskedasticity and Newey-West with the automatic bandwidth selection procedure of Newey and West (1994).

exception of the government purchase forecasts). Table 7 reports the same coefficients for the non-professional survey at both 3 and 2 quarters horizons.

No under-reaction to public information in individual forecast Similarly as in our previous section, we use the lagged consensus forecast from the SPF as a public signal, which is available to the Federal reserve Board at the time of the projection.[9] To capture the surprise component in the public information, we compute the difference between the public signal and individual prior about the signal: $pi_{t,t+h} = g_t - \tilde{E}^i_{t-1}[x_{t+h}]$, where g_t is the public signal.

Figure 5 reports the coefficients estimated by running regression 12 on the GB forecast data. It shows that in large part the non-professional GB forecasts do not exhibit neither overreaction to new private information (red bars), nor underreaction to public information (blue bar) (with the exception of the government purchase forecasts and CPI inflation). Table 8 reports the same coefficients at both 3 and 2 quarters horizons.

[9] The opposite would not be true, as the Green Book projection are published with a lag of 5 years.

Table 7: Individual errors on revisions: non-professional forecasters

	3 quarters horizon			2 quarters horizon		
	β_{BGMS}	SE	p-value	β_{BGMS}	SE	p-value
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Nominal GDP	0.18	0.15	0.23	0.09	0.10	0.36
GDP price index inflation	0.28	0.22	0.21	0.31	0.18	0.09
Real GDP	0.23	0.16	0.17	0.17	0.16	0.28
Consumer Price Index	-0.02	0.20	0.92	0.16	0.08	0.03
Industrial production	-0.04	0.08	0.63	0.12	0.12	0.33
Housing Start	0.07	0.15	0.66	0.02	0.10	0.83
Real residential investment	0.01	0.10	0.88	0.02	0.09	0.80
Real nonresidential investment	0.22	0.13	0.10	0.18	0.12	0.15
Real state and local government consumption	-0.52	0.12	0.00	-0.19	0.10	0.06
Real federal government consumption	-0.35	0.13	0.01	-0.41	0.11	0.00
Unemployment rate	0.13	0.18	0.46	-0.03	0.14	0.83

Notes: The table reports the coefficients from the BGMS regression (individual forecast errors on individual revisions) using forecasts from the Federal Reserve Green Book. Columns (1) to (3) shows the regression coefficients with standard errors and corresponding p-values. Standard errors are robust to heteroskedasticity and Newey-West with the automatic bandwidth selection procedure of Newey and West (1994).

2.4 Comparison between public signals

According to the strategic diversification hypothesis, professional forecasters do not only want to provide an accurate forecasts, but also to differentiate themselves from the average forecast in the survey. As a result, they disregard more (and therefore underreact to) public signal that directly relate to the average information of the forecasters in the survey.

We provide empirical evidence supporting this implication by comparing two different measure of public signals. First, we use our baseline measure for public signal, the consensus forecasts about the same variable at the same horizon in the survey in the previous quarter. We demean using forecaster's prior to isolate the new information content.

$$pi^i_{1,t+h,t} \equiv \bar{\tilde{E}}_{t-1}[x_{t+h}] - \tilde{E}^i_{t-1}[x_{t+h}] \tag{14}$$

Second, we include the current realization of the variable forecasted. In the case of financial variable, this is the realization in the second month of the quarter (when the survey is collected), in case of macroeconomic variable this is the realization in the previous quarter. We demean using forecaster's prior to isolate the new information content.

$$pi^i_{2,t+h,t} \equiv x_{t-1} - \tilde{E}^i_{t-1}[x_{t-1}] \tag{15}$$

Table 8: Private and public information: non-professional forecasters

Panel A: 3 quarters horizon

	Revision			Public signal		
	β_1	SE	p-value	β_2	SE	p-value
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Nominal GDP	0.18	0.13	0.17	-0.03	0.24	0.91
GDP price index inflation	0.20	0.21	0.35	0.39	0.37	0.28
Real GDP	0.24	0.21	0.25	0.22	0.18	0.23
Consumer Price Index	-0.03	0.19	0.86	1.11	0.39	0.01
Industrial production	-0.05	0.09	0.56	0.01	0.14	0.96
Housing Start	0.11	0.16	0.52	0.01	0.18	0.96
Real residential investment	0.10	0.13	0.45	0.01	0.15	0.94
Real nonresidential investment	0.24	0.15	0.10	0.17	0.19	0.38
Real state and local government consumption	-0.60	0.13	0.00	0.38	0.10	0.00
Real federal government consumption	-0.60	0.23	0.01	0.46	0.21	0.03
Unemployment rate	0.12	0.18	0.48	-0.10	0.09	0.24

Panel B: 2 quarters horizon

	Revision			Public signal		
	β_1	SE	p-value	β_2	SE	p-value
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Nominal GDP	0.10	0.10	0.32	0.07	0.11	0.50
GDP price index inflation	0.25	0.16	0.11	0.24	0.20	0.22
Real GDP	0.18	0.18	0.30	0.23	0.09	0.01
Consumer Price Index	0.18	0.09	0.04	0.59	0.26	0.02
Industrial production	0.12	0.14	0.38	0.18	0.10	0.08
Housing Start	0.03	0.10	0.79	-0.12	0.11	0.26
Real residential investment	0.09	0.12	0.48	0.01	0.10	0.96
Real nonresidential investment	0.24	0.15	0.12	0.11	0.11	0.31
Real state and local government consumption	-0.26	0.08	0.00	0.43	0.09	0.00
Real federal government consumption	-0.65	0.21	0.00	0.49	0.20	0.01
Unemployment rate	-0.03	0.14	0.82	-0.02	0.12	0.84

Notes: this table reports the coefficients of regression 12 (individual forecast errors on individual revisions and public information) using forecasts from the Federal Reserve Green Book dataset. Columns 1 to 3 show coefficient β_1 (forecast revision), with standard errors and corresponding p-values. Columns 4 to 6 show coefficient β_2 (public information), with standard errors and corresponding p-values. Standard errors are robust to heteroskedasticity and Newey-West with the automatic bandwidth selection procedure of Newey and West (1994). Panel A uses forecast at 3 quarters horizon and panel B uses forecast at 2 quarters horizon.

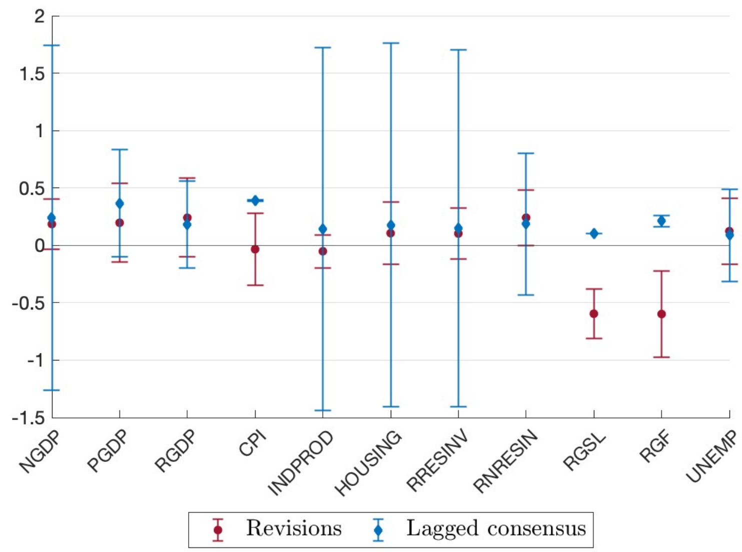

Figure 5: Forecast errors on forecast revisions and public information

Notes: this figure plots the coefficient from panel data regression 12 using forecasts from the Federal Reserve Green Book and horizon h=3. The blue diamonds represent the coefficient β_1 while the red circles represent the coefficient β_2. Bars reports the 90% confidence interval for the estimated coefficients. Standard errors are robust to heteroskedasticity and Newey-West with the automatic bandwidth selection procedure of Newey and West (1994).

We include both variables in regression 12

$$fe^i_{t+h,t} = \alpha + \beta_1 fr^i_{t+h,t} + \beta_2 pi_{1,t+h,t} + \beta_3 pi_{2,t+h,t} + err^i_t \tag{16}$$

As our first public signal related directly to the average beliefs of forecasters, we expect forecasters to disregard it more than the second in the forecast they provide to the survey: $\beta_2 > \beta_3$. The empirical estimates of equation 16 reported in figure confirm it: $\hat{\beta}_2$ is generally larger than $\hat{\beta}_3$, which is still weakly larger than zero with the exception of CPI inflation and unemployment. The tables with the estimated coefficients at both three and two quarters horizon are reported in Appendix E.

Discussion The evidence indicate that the biases documented in surveys of professional forecasters are not present in a survey of non-professional forecasters, as the Federal reserve Board Green Book projections. This evidence is constent with a strategic diversification theory, that affects only professional forecasters, but not with a typical behavioral theory, that makes no such difference.

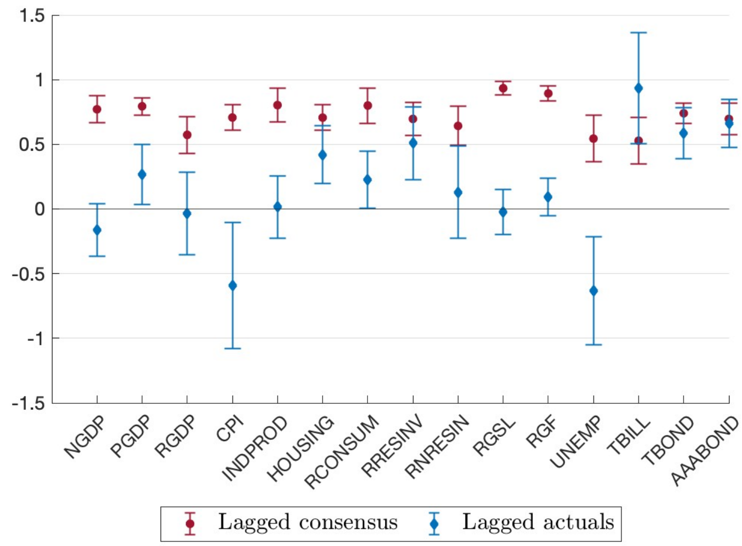

Figure 6: Comparison between public signals

Notes: this figure plots the coefficient from regression 16 using forecasts from the Survey of Professional Forecasters and horizon h=3. The red circles represent the coefficient β_2 while the blue diamonds represent the coefficient β_3. Bars reports the 90% confidence interval for the estimated coefficients. Standard errors are robust to heteroskedasticity and Newey-West with the automatic bandwidth selection procedure of Newey and West (1994).

An important implication of the strategic diversification theory is that beliefs reported in the surveys are only a biased measure of the underlying honest agents' beliefs. As a result, any estimate of the stickiness or forecast dispersion of agents beliefs based on these raw forecast data would turn out ot be biased as well. In the remaining part of the paper, we provide a general theoretical framework of strategic incentives consistent with our four empirical facts and estimate it structurally, in order to recover the actual stickiness of belief updating and forecast dispersion.

3 Strategic incentives in Forecast Reporting

3.1 Static setting

In this section we present a model of strategic interactions in forecast reporting, in which forecasters don't only want to provide accurate forecasts, but also to stand out with respect to the average forecast. We therefore depart from the assumption of truthful reporting by introducing strategic substitutability between forecasters. While the intuition of the model follow from Ottaviani and Sørensen (2006)'s model of winner-take-all game, we employ a more general Morris and Shin (2002) global game setting and introduce public information.

In particular, the forecasters' problem is not only to minimize the expected squared forecast error, but also the expected squared distance with the average forecast:

$$min_{\hat{x}^i} \quad u^i = E^i\left[(\hat{x}^i - x)^2 - \lambda(\hat{x}^i - \bar{\hat{x}})^2\right] \tag{17}$$

where x is the true state and $\bar{\hat{x}} = \int \hat{x}^i di$ is the average of the reported forecast $\hat{x}^i$; $0 < \lambda < 1$ measures the degree of strategic substitutability in agent's reported forecasts.

The first order condition is:

$$\hat{x}^i = \frac{1}{1-\lambda}E^i[x] - \frac{\lambda}{1-\lambda}E^i[\bar{\hat{x}}] \tag{18}$$

If $\lambda = 0$, agents report their honest beliefs. If $\lambda > 0$, agents not only want to be accurate, but also to stand out with respect to the average forecast.

Information Suppose the actual x is unobserved. Forecasters have a common prior $x \sim N(\mu, \chi^{-1})$. Moreover, they received a private and a public signal, both unbiased and centered around the true x with some noise.

$$\begin{aligned} s^i &= x + \eta^i \\ g &= x + e \end{aligned} \tag{19}$$

with $\eta^i \sim N(0, \tau^{-1})$ and $e \sim N(0, \nu^{-1})$.

The resulting honest posterior is

$$E^i[x] = \mu + \gamma_1(g - \mu) + \gamma_2(s^i - \mu) \tag{20}$$

with $\gamma_1 = \frac{\nu}{\tau+\nu+\chi}$, $\gamma_2 = \frac{\tau}{\tau+\nu+\chi}$.

Introduce now strategic substitutability in expectation reporting as in equation 18. We guess a linear solution, solve for the fixed point problem and we get

$$\hat{x}^i = \mu + \delta_1(g - \mu) + \delta_2(s^i - \mu) \tag{21}$$

where $\delta_2 = \frac{\gamma_2}{(1-\lambda)+\lambda\gamma_2}$, $\delta_1 = \frac{(1-\lambda)\gamma_1}{(1-\lambda)+\lambda\gamma_2}$, $1 - \delta_1 - \delta_2 = \frac{(1-\lambda)(1-\gamma_1-\gamma_2)}{(1-\lambda)+\lambda\gamma_2}$. In order to stand out from the crowd, the forecasters overweight new private information in his posted forecast with respect to his actual beliefs ($\delta_2 > \gamma_2$) and underweight new public information ($\delta_1 < \gamma_1$). At the same time, since the prior is common and new information partly private, the agent overweight new information as a whole ($1 - \delta_1 - \delta_2 < 1 - \gamma_1 - \gamma_2$).

Proposition 4 *In a strategic substitutability game as in 18, with $0 < \lambda < 1$, the coefficient of the individual regression 3 is given by:*

$$\beta_{BGMS} = \frac{-\lambda\tau\chi}{([(1-\lambda)\nu + \tau]^2 + (1-\lambda)^2\nu\chi)} \tag{22}$$

Thus $\beta_{BGMS} < 0$ if $\lambda > 0$.

If $\lambda = 0$, there is no strategic interaction between forecasters and they simply report their honest beliefs. In that case, $\beta_{BGMS} = 0$, as forecast errors are not correlated with any information available in time t, and in particular her forecast revisions. This result follows directly from rational expectation. On the other hand, if $\lambda > 0$, agents overweight private information to stand out from the crowd, which results in $\beta_{BGMS} < 0$, meaning overreaction to new information. Ottaviani and Sørensen (2006) derive a similar result in a specific winner-take-all game only considering private information. The model reconciles the empirical result in section 2.

Proposition 5 *In a strategic substitutability game as in 18, with $0 < \lambda < 1$, the coefficient of the consensus regression 5 is given by:*

$$\beta_{CG} = \frac{(1-\lambda)\tau\chi}{([(1-\lambda)\nu + \tau]^2 + [(1-\lambda)^2\nu + \tau]\chi)} \tag{23}$$

Thus $\beta_{CG} > 0$ if $\lambda < 1$.

If $\lambda = 0$, there is no strategic interaction between forecasters and they simply report their honest beliefs. In that case, $\beta_{CG} > 0$: the average forecast is sub-optimally sticky with respect to the average signal, which is less noisy than the individual one as shown by CG. On the other hand, if $\lambda > 0$, agents overweight new information and the average forecast is less sticky. The higher is the strategic incentive λ, the lower is the rigidity of posted forecast. In the limit case of $\lambda \to 1$, individual forecasters adjust one-to-one their posteriors to new information, making the average forecast not sticky. It is not possible to have $\beta_{CG} < 0$,consistently with the data.

Proposition 6 *In a strategic substitutability game as in 18, with $0 < \lambda < 1$, the coefficient of the individual regression 12 is given by:*

$$\begin{aligned}\beta_1 &= \frac{-\lambda(\nu+\chi)}{(\tau+\nu+\chi)} \\ \beta_2 &= \frac{\lambda\nu}{(\tau+\nu+\chi)}\end{aligned} \tag{24}$$

Thus $\beta_1 > 0$ and $\beta_2 < 0$ if $\lambda > 0$.

Proposition 7 represents our main theoretical result. If $\lambda = 0$, forecasters report their honest beliefs and both $\beta_1 = 0$ and $\beta_2 = 0$ as implied by rational expectation. However, with strategic incentives $\lambda > 0$, forecasters overweight private information and underweight public information, in order to stand from the crowd. This leads to an underreaction to public information, as measured by $\beta_2 > 0$, and overreaction to private information, as measured by $\beta_1 < 0$. This result reconciles our new empirical fact four documented in section 2.

3.2 Dynamic setting

We now extend the previous strategic incentives model to a dynamic setting. Assume the series follows a AR(1) process

$$x_t = \rho x_{t-1} + u_t \tag{25}$$

with $u \sim N(0, \xi^{-1})$.

Each agent receive a private signal s_t^i and a public signal g_t

$$\begin{aligned}s_t^i &= x_t + \eta_t^i \\ g_t &= x_t + e_t\end{aligned} \tag{26}$$

with $\eta_t^i \sim N(0, \tau^{-1})$, $e_t \sim N(0, \nu^{-1})$.

Honest beliefs Agents form beliefs about x at horizon h: $E_t^i[x_{t+h}] \equiv x_{t+h,t}^i$. The honest posterior belief about x is given by the Kalman filter

$$x_{t,t}^i = x_{t,t-1}^i + K_{1,1}(g_t - x_{t,t-1}^i) + K_{1,2}(s_t^i - x_{t,t-1}^i)$$

where the Kalman gains are

$$\begin{aligned} K_{1,1} &= \frac{\nu}{\Sigma^{-1} + \nu + \tau} \\ K_{1,2} &= \frac{\tau}{\Sigma^{-1} + \nu + \tau} \end{aligned} \tag{27}$$

and the posterior forecast error variance

$$\begin{aligned} \Sigma &\equiv E[(x_t - x_{t,t-1}^i)(x_t - x_{t,t-1}^i)'] \\ &= \frac{-[(\rho^2 - 1)\xi + (\tau + \nu)] + \sqrt{[(\rho^2 - 1)\xi + (\tau + \nu)]^2 + 4(\tau + \nu)\xi}}{2} \end{aligned} \tag{28}$$

Strategic interactions As in the previous section, the strategic substitutability in agents objective function leads them to report

$$\begin{aligned} \hat{x}_{t,t}^i &= \frac{1}{1-\lambda} x_{t,t}^i - \frac{\lambda}{1-\lambda} E^i[\bar{\hat{x}}_{t,t}] \\ \hat{x}_{t+h,t}^i &= \rho^h \hat{x}_{t,t}^i \end{aligned} \tag{29}$$

where $\hat{x}_{t+h,t}^i$ is the forecast provided by individual i in t about realization in $t+h$, and $\bar{\hat{x}}_{t+h,t} = \int^i \hat{x}_{t+h,t}^i di$ is the average of forecasts provided in t about realization in $t+h$. If $\lambda = 0$, agents report their true beliefs. With $0 < \lambda < 1$, agents not only want to be accurate, but also to stand out with respect to the average forecast.

The model builds on Woodford (2001) and Coibion and Gorodnichenko (2012).[10] Following them, we average $\hat{x}_{t,t}^i$ across agents and use repeated substitution in 29 to express the reported average forecast as

$$F_t = -\frac{1}{1-\lambda} \sum_{k=0}^{\infty} \left(\frac{\lambda}{1-\lambda}\right)^k \bar{E}^{(k)}[x_t] = \frac{1}{1-\lambda} \bar{x}_{t+h,t} - \frac{\lambda}{1-\lambda} \bar{E}_t[\bar{\hat{x}}_{t+h,t}] \tag{30}$$

[10] Our model depart from the latter in two dimensions. First, while they consider only strategic complementarity, we focus on strategic substitutability. Second, while they only consider consensus forecasts, we are interested in individual forecasts

We guess and verify the law of motion for F_t and the other unobserved state variables. In particular, we conjecture that the state vector evolves according to[11]

$$Z \equiv \begin{bmatrix} x_t \\ F_t \\ w_t \end{bmatrix} = MZ_{t-1} + m \begin{bmatrix} u_t \\ e_t \end{bmatrix} \tag{31}$$

Where

$$M = \begin{bmatrix} \rho & 0 & 0 \\ G & L & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad m = \begin{bmatrix} 1 & 0 \\ m_{2,1} & m_{2,2} \\ 0 & 1 \end{bmatrix} \tag{32}$$

the observable variables are the two signals about x_t

$$V_t^i \equiv \begin{bmatrix} g_t \\ s_t^i \end{bmatrix} = HZ_t + \begin{bmatrix} 0 \\ \eta_t^i \end{bmatrix} \tag{33}$$

where

$$H = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \tag{34}$$

Agents use their conjecture law of motion 31 and the observables 33 to infer the state using the individual Kalman filter. The posterior estimate of the state vector by agent i is

$$\begin{aligned} E_t^i[Z_t] &= ME_{t-1}^i[Z_{t-1}] + K(V_t^i - E_{t-1}^i[V_t]) \\ &= (I - KH)ME_{t-1}^i[Z_{t-1}] + KHMZ_{t-1} + KHm \begin{bmatrix} u_t \\ e_t \end{bmatrix} + K \begin{bmatrix} 0 \\ \eta_t^i \end{bmatrix} \end{aligned} \tag{35}$$

Where K is the Kalman gain. Average 35 to find the consensus believe on the state vector.

$$\bar{E}_t[Z_t] = (I - KH)M\bar{E}_{t-1}[Z_{t-1}] + KHMZ_{t-1} + KHm \begin{bmatrix} u_t \\ e_t \end{bmatrix} \tag{36}$$

[11] w_t takes care of the correlation between public signal and higher order beliefs F_t

From the definition on F_t in 30 it follows that

$$\begin{aligned} F_t &= \begin{bmatrix} \frac{1}{1-\lambda} & -\frac{\lambda}{1-\lambda} & 0 \end{bmatrix} \bar{E}_t[Z_t] \equiv \xi \bar{E}_t[Z_t] \\ &= \xi(I - KH)M\bar{E}_{t-1}[Z_{t-1}] + \xi KHMZ_{t-1} + \xi KHm \begin{bmatrix} u_t \\ e_t \end{bmatrix} \end{aligned} \tag{37}$$

$$\begin{aligned} F_t =& ((1-\alpha)\rho + \alpha G)\bar{E}_{t-1}[x_{t-1}] + \alpha L\bar{E}_{t-1}[F_{t-1}] - C\rho\bar{E}_{t-1}[x_{t-1}] \\ & + C\rho x_{t-1} + C_1 e_t + Cu_t \\ =& [\rho(1-\alpha) + \alpha G - (1-\alpha)L - C\rho]\bar{E}_{t-1}[x_{t-1}] + \\ & + LF_{t-1} + C\rho x_{t-1} + Cu_t + C_1 e_t \end{aligned} \tag{38}$$

where we used 30 to substitute

$$\alpha \bar{E}_t[F_{t-1}] = F_{t-1} - (1-\alpha)\bar{E}_{t-1}[x_{t-1}]$$

and we defined

$$C_1 \equiv \frac{K_{1,1} - \lambda(K_{2,1})}{1-\lambda}, \qquad C_2 \equiv \frac{K_{1,2} - \lambda K_{2,2}}{1-\lambda} \qquad \text{and} \qquad C = C_1 + C_2$$

Equation 38 must equal the second line of the perceived law of motion 31. The solution to the fixed point is given by $G = C\rho$, $m_{2,1} = C$, $m_{2,2} = C_1$ and $L = \rho - G$.

Given the law of motion of unobserved state 31 and the observable 33, the posterior variance of the forecast solves the following Ricatti equation

$$\begin{aligned} \Sigma &\equiv E[(Z_t - Z^i_{t,t-1})(Z_t - Z^i_{t,t-1})'] \\ \Sigma &= M(\Sigma - \Sigma H' \left(H\Sigma H' + \begin{bmatrix} 0 & 0 \\ 0 & \tau^{-1} \end{bmatrix} \right)^{-1} H\Sigma)M' + m \begin{bmatrix} \xi^{-1} & 0 \\ 0 & \nu^{-1} \end{bmatrix} m' \end{aligned} \tag{39}$$

and the Kalman filter is

$$K = \Sigma H' \left(H\Sigma H' + \begin{bmatrix} 0 & 0 \\ 0 & \tau^{-1} \end{bmatrix} \right)^{-1} \tag{40}$$

Finally, the individual posted forecast is

$$\begin{aligned}\hat{x}^i_{t,t} &= \xi E^i_t[Z_t] \\ &= \hat{x}^i_{t,t-1} + C_1(g_t - \hat{x}^i_{t,t-1}) + C_2(s^i_t - \hat{x}^i_{t,t-1})\end{aligned} \tag{41}$$

Note that the individual posted forecast updating in 41 is similar to individual Kalman Filter in 35, with C_1 and C_2 as "modified" gains in place of K_1 and K_2. In particular, if $\lambda = 0$, $C_1 = K_1$ and $C_2 = K_2$: with no strategic incentives, agents simply report their honest beliefs. However, when $\lambda > 0$, one can show that $C_1 < K_1$ and $C_2 > K_2$: agents underweight public information and overweight private information in their posted forecast.[12]

The posted forecast updating 41 mirrors the general framework 2 in section 2 with $G_1 = C_1$, $G_2 = C_2$ and $\tilde{E}_t[x_t] = \hat{x}_{t,t}$. Therefore the coefficient from regressions 5, 11, 3 and 12 follows from propositions 1-4.

4 Structural estimation

We now proceed to estimate our model to test its ability to match the data recover the honest beliefs of forecasters to compute the actual belief stickiness. First, for each series we estimate the autoregressive coefficient ρ and the fundamental disturbance variance $\sigma^2_u \equiv \xi^{-1}$ directly from the data. Then we use the simulated method of moments to estimate the remaining parameters of the model: the public noise variance $\sigma^2_e \equiv \nu^{-1}$, the private noise variance $\sigma^2_\eta \equiv \tau^{-1}$ and the strategic incentive parameter λ. We prefer the simulated method of moments to maximum likelihood as the simplicity of our model come at the cost of likely misspecification which would be problematic when using maximum likelihood.

The data moments we target are the mean cross sectional dispersion of forecast errors, the coefficient β_1 from the public information regression 12 and the posted gain C from regression 11. We choose these three moments as they are differently affected by the three parameters to be estimated and therefore provide good identification.[13]

[12] To see this, note that $K_{1,1} < K_{2,1}$: intuitively, the public signal is more informative about the average forecast than about the actual state, because of the average belief depends also on the public noise. On the other hand, $K_{1,2} > K_{2,2}$: intuitively, the private signal is less informative about the average forecast than about the actual state, as the average forecast also depends on the public noise.

[13] While larger strategic incentive parameter λ increases posted gain C, an increase in either private or public noise decreases it. On the other hand, the coefficient β_1 decreases in private noise and strategic incentives but increase in public noise (see proposition 7). Finally, strategic incentives and public noise always increase mean dispersion, while private noise initially increases it and then decreases it.

Table 9: Estimasted parameters

Variable	ρ (1)	$\frac{\sigma_e}{\sigma_u}$ (2)	$\frac{\sigma_\eta}{\sigma_u}$ (3)	λ (4)
Nominal GDP	0.93	1.48	1.70	0.74
GDP price index inflation	0.93	1.60	2.13	0.88
Real GDP	0.80	1.30	1.36	0.47
Consumer Price Index	0.78	1.38	1.60	0.61
Industrial production	0.85	1.28	1.86	0.68
Housing Start	0.85	1.38	1.81	0.70
Real Consumption	0.87	1.33	1.84	0.67
Real residential investment	0.89	1.56	1.74	0.49
Real nonresidential investment	0.89	2.37	1.28	0.25
Real state and local government consumption	0.89	1.32	2.79	0.90
Real federal government consumption	0.80	1.29	2.90	0.87
Ten-year Treasury rate	0.83	1.81	1.56	0.72
AAA Corporate Rate Bond	0.85	1.76	1.82	0.87

Table 9 reports the estimated parameters for each series, while table 10 reports targeted and untargeted moments in the model and in the data. While the model is able to match the targeted model in columns 1-6, it also does a good job in matching the untargeted moments in columns 7-12 for most of the series. The only exceptions are the CG coefficient for the financial series in column 7-8 and a general underestimation of the public information coefficient in column 11-12.

We use the model to recover the honest beliefs of forecasters behind the biased posted forecast, and compute the related moments. We compare them with the moments calculated on the posted forecast in table 11. First, columns 1-3 reports the weight on new information in posted forecast, in honest beliefs and their ratio. Hones beliefs are stickier than posted forecasts, as the latter overweight new information. The magnitude of the difference can be appreciated by looking at the difference between posted and honest consensus mean forecast error. Because of the individual overweight of new information, the average posted forecast is less sticky and therefore more accurate than the honest one. For some variable (nominal GDP, CPI, Housing Start, Ten-year and AAA bond rate) the honest error is more than 1.5 times the posted one. Finally, columns 7-8 compare honest and posted cross sectional dispersion of forecast errors, which is used as a proxy for uncertainty

Table 10: Moments in data and model

	Targeted moments						Untargeted moments					
	Mean Dispersion		C		β_1		β_{CG}		β_{BGMS}		β_2	
	Data	Model	Data	Model	Data	Model	Data	Model	Data	Model	Data	Model
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Nominal GDP	1.49	1.49	0.53	0.53	-0.54	-0.54	0.52	0.41	-0.25	-0.31	0.75	0.21
GDP price index inflation	0.33	0.33	0.49	0.49	-0.68	-0.68	0.29	0.50	-0.35	-0.44	0.81	0.31
Real GDP	0.92	0.92	0.56	0.56	-0.34	-0.34	0.65	0.33	-0.10	-0.15	0.57	0.13
Consumer Price Index	0.31	0.31	0.49	0.49	-0.48	-0.48	0.22	0.38	-0.30	-0.24	0.67	0.16
Industrial production	3.71	3.71	0.50	0.50	-0.59	-0.59	0.21	0.22	-0.30	-0.22	0.79	0.26
Housing Start	110.04	110.04	0.49	0.49	-0.58	-0.58	0.38	0.32	-0.28	-0.28	0.78	0.23
Real Consumption	0.51	0.51	0.49	0.49	-0.56	-0.56	0.31	0.25	-0.26	-0.23	0.80	0.23
Real residential investment	27.03	27.03	0.41	0.41	-0.37	-0.37	1.22	0.40	-0.08	-0.17	0.73	0.11
Real nonresidential investment	7.38	7.38	0.48	0.48	-0.12	-0.12	1.21	0.94	0.08	-0.10	0.65	0.01
Real state and local government consumption	1.41	1.41	0.47	0.47	-0.84	-0.84	0.63	0.17	-0.48	-0.41	0.91	0.45
Real federal government consumption	6.40	6.40	0.43	0.43	-0.83	-0.83	-0.23	0.12	-0.56	-0.35	0.93	0.37
Ten-year Treasury rate	0.17	0.17	0.51	0.51	-0.47	-0.47	-0.01	0.69	-0.22	-0.38	0.76	0.09
AAA Corporate Rate Bond	0.34	0.34	0.54	0.54	-0.61	-0.61	-0.03	0.62	-0.27	-0.48	0.83	0.18

in the literature (e.g. Kozeniauskas et al. 2018). The overweight of private information increase substantially the dispersion of forecasts, as the honest beliefs dispersion is less than half the posted one for most of the series.

Table 11: Posted and honest moments

	Gain			Consensus MSE			Dispersion		
	Posted	Honest	Ratio	Posted	Honest	Ratio	Posted	Honest	Ratio
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Nominal GDP	0.53	0.40	0.76	0.49	1.07	2.19	1.49	0.29	0.19
GDP price index inflation	0.49	0.32	0.66	0.05	0.14	2.92	0.33	0.02	0.06
Real GDP	0.56	0.49	0.88	0.78	1.14	1.47	0.92	0.41	0.44
Consumer Price Index	0.49	0.40	0.82	0.23	0.36	1.58	0.31	0.08	0.27
Industrial production	0.50	0.44	0.87	3.51	5.11	1.46	3.71	0.60	0.16
Housing Start	0.49	0.40	0.82	69.95	115.75	1.65	110.04	18.10	0.16
Real Consumption	0.49	0.42	0.86	0.46	0.68	1.49	0.51	0.09	0.18
Real residential investment	0.41	0.36	0.87	29.60	40.95	1.38	27.03	10.76	0.40
Real nonresidential investment	0.48	0.43	0.90	4.12	5.30	1.29	7.38	6.01	0.82
Real state and local government consumption	0.47	0.40	0.86	0.54	0.81	1.51	1.41	0.02	0.02
Real federal government consumption	0.43	0.39	0.90	5.96	7.49	1.26	6.40	0.14	0.02
Ten-year Treasury rate	0.51	0.33	0.64	0.04	0.11	2.55	0.17	0.05	0.27
AAA Corporate Rate Bond	0.54	0.29	0.54	0.04	0.14	3.75	0.34	0.04	0.11

References

BORDALO, P., N. GENNAIOLI, Y. MA, AND A. SHLEIFER (2018a): "Over-reaction in macroeconomic expectations," Tech. rep., National Bureau of Economic Research.

——— (2020): "Overreaction in macroeconomic expectations," *American Economic Review*, 110, 2748–82.

BORDALO, P., N. GENNAIOLI, AND A. SHLEIFER (2018b): "Diagnostic expectations and credit cycles," *The Journal of Finance*, 73, 199–227.

BROER, T. AND A. KOHLHAS (2018): "Forecaster (mis-) behavior," .

COIBION, O. AND Y. GORODNICHENKO (2012): "What can survey forecasts tell us about information rigidities?" *Journal of Political Economy*, 120, 116–159.

——— (2015): "Information rigidity and the expectations formation process: A simple framework and new facts," *American Economic Review*, 105, 2644–78.

CROUSHORE, D. D. (1997): "The Livingston survey: Still useful after all these years," *Business Review-Federal Reserve Bank of Philadelphia*, 2, 1.

CROWE, C. (2010): "Consensus forecasts and inefficient information aggregation," *IMF Working Papers*, 1–43.

DANIEL, K., D. HIRSHLEIFER, AND A. SUBRAHMANYAM (1998): "Investor psychology and security market under-and overreactions," *the Journal of Finance*, 53, 1839–1885.

DUPOR, B., T. KITAMURA, AND T. TSURUGA (2010): "Integrating Sticky Information and Sticky Prices," *Review of Economics and Statistics*, 92, 657–669.

EUROPEAN CENTRAL BANK (2014): "Results of the Second Special Questionnaire for Participants in the ECB Survey of Professional Forecasters," .

EYSTER, E., M. RABIN, AND D. VAYANOS (2019): "Financial markets where traders neglect the informational content of prices," *The Journal of Finance*, 74, 371–399.

GOLDSTEIN, N. (2021): "Tracking inattention," *Available at SSRN 3641353*.

KILEY, M. T. (2007): "A quantitative comparison of sticky-price and sticky-information models of price setting," *Journal of Money, Credit and Banking*, 39, 101–125.

KLENOW, P. J. AND J. L. WILLIS (2007): "Sticky information and sticky prices," *Journal of Monetary Economics*, 54, 79–99.

KNOTEK II, E. S. (2010): "A Tale of Two Rigidities: Sticky Prices in a Sticky-Information Environment," *Journal of Money, Credit and Banking*, 42, 1543–1564.

KORENOK, O. (2008): "Empirical comparison of sticky price and sticky information models," *Journal of Macroeconomics*, 30, 906–927.

KOZENIAUSKAS, N., A. ORLIK, AND L. VELDKAMP (2018): "What are uncertainty shocks?" *Journal of Monetary Economics*, 100, 1–15.

MANKIW, N. G. AND R. REIS (2002): "Sticky information versus sticky prices: a proposal to replace the New Keynesian Phillips curve," *The Quarterly Journal of Economics*, 117, 1295–1328.

MARINOVIC, I., M. OTTAVIANI, AND P. SORENSEN (2013): "Forecasters' objectives and strategies," in *Handbook of economic forecasting*, Elsevier, vol. 2, 690–720.

MORRIS, S. AND H. S. SHIN (2002): "Social value of public information," *american economic review*, 92, 1521–1534.

MORSE, A. AND A. VISSING-JORGENSEN (2020): "Information Transmission from the Federal Reserve to the Stock Market: Evidence from Governors' Calendars," Tech. rep., working paper.

OTTAVIANI, M. AND P. N. SØRENSEN (2006): "The strategy of professional forecasting," *Journal of Financial Economics*, 81, 441–466.

WOODFORD, M. (2001): "Imperfect common knowledge and the effects of monetary policy," Tech. rep., National Bureau of Economic Research.

A Variable definitions

The forecast data come from two different datasets: the Survey of Professional Forecasters, collected by the Federal Reserve Bank of Philadelphia, and the Tealbook/Greenbook of the Federal Reserve Board of Governors.

Survey of Professional Forecasters All surveys are collected around the 3rd week of the middle month in the quarter. In this section, x_t indicate the actual value and $F_t x_{t+h}$ the forecast provided in t about horizon h. All actual values of macroeconomic series (1-12) use the first release level, which are available to forecasters in the following quarter. We transform the macroeconomic level in year-over-year growth, following the literature.

1. NGDP
    - Variable: nominal GDP.
    - Question: The level of nominal GDP in the current quarter and the next 4 quarters.
    - Forecast: Nominal GDP growth from end of quarter $t-1$ to end of quarter $t+3$: $\frac{F_t x_{t+3}}{x_{t-1}} - 1$
    - Revision: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$
    - Actual: $\frac{x_{t+3}}{x_{t-1}} - 1$
2. RGDP
    - Variable: real GDP.
    - Question: The level of real GDP in the current quarter and the next 4 quarters.
    - Forecast: real GDP growth from end of quarter $t-1$ to end of quarter $t+3$: $\frac{F_t x_{t+3}}{x_{t-1}} - 1$
    - Revision: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$
    - Actual: $\frac{x_{t+3}}{x_{t-1}} - 1$
3. PGDP
    - Variable: GDP deflator.

- Question: The level of GDP deflator in the current quarter and the next 4 quarters.
- Forecast: GDP price deflator inflation from end of quarter $t-1$ to end of quarter $t+3$: $\frac{F_t x_{t+3}}{x_{t-1}} - 1$
- Revision: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$
- Actual: $\frac{x_{t+3}}{x_{t-1}} - 1$

4. CPI

- Variable: Consumer Price Index.
- Question: CPI growth rate in the current quarter and the next 4 quarters.
- Forecast: CPI inflation from end of quarter $t-1$ to end of quarter $t+3$: $F_t(z_t/4+1) * F_t(z_{t+1}/4+1) * F_t(z_{t+2}/4+1) * F_t(z_{t+3}/4+1)$, where z is the annualized quarterly CPI inflation in quarter t.
- Revision: $F_t(z_t/4+1) * F_t(z_{t+1}/4+1) * F_t(z_{t+2}/4+1) * F_t(z_{t+3}/4+1) - F_{t-1}(z_t/4+1) * F_{t-1}(z_{t+1}/4+1) * F_{t-1}(z_{t+2}/4+1) * F_{t-1}(z_{t+3}/4+1)$
- Actual: $\frac{x_{t+3}}{x_{t-1}} - 1$. Real time data is not available before 1994Q3. For actual periods prior to this date, we use data published in 1994Q3 to measure the actual outcome.

5. RCONSUM

- Variable: Real consumption.
- Question: The level of real consumption in the current quarter and the next 4 quarters.
- Forecast: GDP price deflator inflation from end of quarter $t-1$ to end of quarter $t+3$: $\frac{F_t x_{t+3}}{x_{t-1}} - 1$
- Revision: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$
- Actual: $\frac{x_{t+3}}{x_{t-1}} - 1$

6. INDPROD

- Variable: Industrial production index.

- Question: The average level of the industrial production index in the current quarter and the next 4 quarters.
- Forecast: Growth of the industrial production index from quarter $t-1$ to quarter $t+3$: $\frac{F_t x_{t+3}}{x_{t-1}} - 1$
- Revision: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$
- Actual: $\frac{x_{t+3}}{x_{t-1}} - 1$

7. RNRESIN

   - Variable: Real non-residential investment.
   - Question: The level of real non-residential investment in the current quarter and the next 4 quarters.
   - Forecast: Growth of real non-residential investment from quarter $t-1$ to quarter $t+3$: $\frac{F_t x_{t+3}}{x_{t-1}} - 1$
   - Revision: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$
   - Actual: $\frac{x_{t+3}}{x_{t-1}} - 1$

8. RRESIN

   - Variable: Real residential investment.
   - Question: The level of real residential investment in the current quarter and the next 4 quarters.
   - Forecast: Growth of real residential investment from quarter $t-1$ to quarter $t+3$: $\frac{F_t x_{t+3}}{x_{t-1}} - 1$
   - Revision: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$
   - Actual: $\frac{x_{t+3}}{x_{t-1}} - 1$

9. RGF

   - Variable: Real federal government consumption.
   - Question: The level of real federal government consumption in the current quarter and the next 4 quarters.

- Forecast: Growth of real federal government consumption from quarter $t-1$ to quarter $t+3$: $\frac{F_t x_{t+3}}{x_{t-1}} - 1$
- Revision: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$
- Actual: $\frac{x_{t+3}}{x_{t-1}} - 1$

10. RGSL

- Variable: Real state and local government consumption.
- Question: The level of real state and local government consumption in the current quarter and the next 4 quarters.
- Forecast: Growth of real state and local government consumption from quarter $t-1$ to quarter $t+3$: $\frac{F_t x_{t+3}}{x_{t-1}} - 1$
- Revision: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$
- Actual: $\frac{x_{t+3}}{x_{t-1}} - 1$

11. HOUSING

- Variable: Housing starts.
- Question: The level of housing starts in the current quarter and the next 4 quarters.
- Forecast: Growth of housing starts from quarter $t-1$ to quarter $t+3$: $\frac{F_t x_{t+3}}{x_{t-1}} - 1$
- Revision: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$
- Actual: $\frac{x_{t+3}}{x_{t-1}} - 1$

12. UNEMP

- Variable: Unemployment rate.
- Question: The level of average unemployment rate in the current quarter and the next 4 quarters.
- Forecast: Average quarterly unemployment rate in quarter $t+3$: $F_t x_{t+3}$
- Revision: $F_t x_{t+3} - F_{t-1} x_{t+3}$
- Actual: x_{t+3}

13. TB3M
    - Variable: 3-month Treasury rate.
    - Question: The level of average 3-month Treasury rate in the current quarter and next 4 quarters.
    - Forecast: Average quarterly 3-month Treasury rate in quarter $t+3$: $F_t x_{t+3}$
    - Revision: $F_t x_{t+3} - F_{t-1} x_{t+3}$
    - Actual: x_{t+3}
14. TN10Y
    - Variable: 10-year Treasury rate.
    - Question: The level of average 10-year Treasury rate in the current quarter and next 4 quarters.
    - Forecast: Average quarterly 10-year Treasury rate in quarter $t+3$: $F_t x_{t+3}$
    - Revision: $F_t x_{t+3} - F_{t-1} x_{t+3}$
    - Actual: x_{t+3}
15. AAA
    - Variable: AAA corporate bond rate.
    - Question: The level of average AAA corporate bond rate in the current quarter and next 4 quarters.
    - Forecast: Average quarterly AAA corporate bond rate in quarter $t+3$: $F_t x_{t+3}$
    - Revision: $F_t x_{t+3} - F_{t-1} x_{t+3}$
    - Actual: x_{t+3}

Tealbook/Greenbook projections The Tealbook/Greenbook is produced by the Research staff at the Federal Reserve Board of Governors before each meeting of the Federal Open Market Committee. We aggregate at quarterly frequency by considering only the last projection available in any quarter. The projections from the Tealbook/Greenbook are released to the public with a lag of five years. We consider the variables forecasted also in the SPF, meaning all the macroeconomic variables with the exception of real consumption.

For most variables (NGDP, PGDP, RGDP, CPI, INDPROD, RRESINV, RNRESIN, RGF, RGSL) the projections are provided n quarter-over-quarter growth, while HOUSING is provided in level. We transform m in tear-over-year growth in order to compare them with SPF forecasts. Similarly, we keep projectons for UNEMP in level. The actuals are the same as for the SPF data.

B Proofs

Proposition 1. Let $\hat{x}_{t,t+h} \equiv \tilde{E}_t[x_{t+h}]$ and $x_{t,t+h} \equiv E_t[x_{t+h}]$. Without loss of generality, assume $h = 0$. From 2

$$\begin{aligned}
\hat{x}^i_{t,t} &= \hat{x}^i_{t-1,t} + G(x_t - \hat{x}^i_{t-1,t}) + G_1 e_t + G_2 \eta^i_t \\
\hat{x}^i_{t,t} &= \hat{x}^i_{t-1,t} + Gx_t - G\hat{x}^i_{t,t} + G(\hat{x}^i_{t,t} - \hat{x}^i_{t-1,t}) + G_1 e_t + G_2 \eta^i_t \\
(1-G)(\hat{x}^i_{t,t} - \hat{x}^i_{t-1,t}) &= +G(x_t - \hat{x}^i_{t,t}) + G_1 e_t + G_2 \eta^i_t \\
\underbrace{(x_t - \hat{x}^i_{t,t})}_{fe^i_{t,t}} &= \frac{1-G}{G}\underbrace{(\hat{x}^i_{t,t} - \hat{x}^i_{t-1,t})}_{fr^i_{t,t}} - \frac{G_1}{G}e_t - \frac{G_2}{G}\eta^i_t
\end{aligned} \tag{42}$$

therefore by running BGMS regression 3, the regressor $fr^i_{t,t}$ is correlated with the unobservable error. The resulting $\hat{\beta}_{BGMS}$ is equal to:

$$\begin{aligned}
\beta_{BGMS} &= \frac{1-G}{G} + \frac{cov(G[x_t - \hat{x}^i_{t-1,t}] + G_1 e_t + G_2 \eta^i_t, -\frac{G_1}{G}e_t - \frac{G_2}{G}\eta^i_t)}{var(G[x_t - \hat{x}^i_{t-1,t}] + G_1 e_t + G_2 \eta^i_t)} \\
&= \frac{1-G}{G} - \frac{\frac{G_1^2}{G}\nu^{-1} + \frac{G_2^2}{G}\tau^{-1}}{G^2 var(x_t - \hat{x}^i_{t-1,t}) + G_1^2\nu^{-1} + G_2^2\tau^{-1}} \\
&= \frac{var(x_t - \hat{x}^i_{t-1,t}) - [Gvar(x_t - \hat{x}^i_{t-1,t}) + \frac{G_1^2}{G}\nu^{-1} + \frac{G_2^2}{G}\tau^{-1}]}{Gvar(x_t - \hat{x}^i_{t-1,t}) + \frac{G_1^2}{G}\nu^{-1} + \frac{G_2^2}{G}\tau^{-1}} \\
&= \frac{\hat{\Sigma} - [G\hat{\Sigma} + \frac{G_1^2}{G}\nu^{-1} + \frac{G_2^2}{G}\tau^{-1}]}{G\hat{\Sigma} + \frac{G_1^2}{G}\nu^{-1} + \frac{G_2^2}{G}\tau^{-1}}
\end{aligned} \tag{43}$$

with $\hat{\Sigma} \equiv var(x_t - \hat{x}^i_{t-1,t})$. Notice that $\hat{\Sigma} \neq \Sigma \equiv var(x_t - x^i_{t-1,t})$.

Consider the special case of x_t following an AR(1) process $x_t = \rho_i x_{t-1} + u_t$ with $u_t \sim$

$N(0, \xi^{-1})$ and assuming $\hat{x}_{t+1,t} = \rho \hat{x}_{t,t}$. In steady state

$$\begin{aligned} x_{t+1} - \rho \hat{x}^i_{t,t} =& \rho(x_t - \hat{x}^i_{t,t}) + u_{t+1} \\ \hat{\Sigma} =& \rho^2 \hat{\Phi} + \xi^{-1} \end{aligned} \tag{44}$$

where $\hat{\Phi} = var(x_t - \hat{x}^i_{t,t})$.

$$\begin{aligned} x_t - \hat{x}^i_{t,t} =& (1-G)[x_t - \rho \hat{x}^i_{t-1,t-1}] - G_1 e_t - G_2 \eta^i_t \\ \hat{\Phi} =& (1-G)^2 \hat{\Sigma} + G_1^2 \nu^{-1} + G_2^2 \tau^{-1} \end{aligned} \tag{45}$$

Substitute and solve for $\hat{\Sigma}$

$$\hat{\Sigma} = \frac{\rho^2[G_1^2 \nu^{-1} + G_2^2 \tau^{-1}] + \xi^{-1}}{1 - \rho^2(1-G)^2} \tag{46}$$

■

Corollary 1. With rational expectation, $G_1 = \frac{\nu}{\tau + \nu + \Sigma^{-1}}$ and $G_2 = \frac{\tau}{\tau + \nu + \Sigma^{-1}}$, with $\Sigma \equiv var(x_t - E^i_t[x_t])$. Define $\chi = \Sigma^{-1}$. From 43

$$\begin{aligned} \beta_{BGMS} &= \frac{\chi}{\nu+\tau} - \frac{\frac{1}{\chi+\nu+\tau}}{\frac{1}{(\chi+\nu+\tau)^2}[(\nu+\tau)^2\chi^{-1} + \nu + \tau]} \\ \beta_{BGMS} &= \frac{\chi}{\nu+\tau} - \frac{(\chi+\nu+\tau)\chi}{(\nu+\tau)^2 + \nu\chi + \tau\chi} \\ \beta_{BGMS} &= \frac{\chi}{\nu+\tau} - \frac{(\chi+\nu+\tau)\chi}{(\chi+\nu+\tau)(\nu+\tau)} \\ \beta_{BGMS} &= 0 \end{aligned} \tag{47}$$

■

Proposition 2. Let $\hat{x}_{t,t+h} \equiv \tilde{E}_t[x_{t+h}]$. Without loss of generality, assume $h = 0$. From 2

$$\underbrace{(x_t - \bar{\hat{x}}_{t,t})}_{\bar{fe}_{t,t}} = \frac{1-G}{G}\underbrace{(\bar{\hat{x}}_{t,t} - \bar{\hat{x}}_{t-1,t})}_{\bar{fr}_{t,t}} - \frac{G_1}{G} e_t \tag{48}$$

therefore by running CG regression 5, the regressor $\bar{fr}_{t,t}$ is correlated with the unobservable error. The resulting $\hat{\beta}_{CG}$ is equal to:

$$\begin{aligned}\beta_{CG} &= \frac{1-G}{G} + \frac{cov(G[x_t - \bar{\hat{x}}_{t-1,t}] + G_1 e_t, -\frac{G_1}{G} e_t)}{var(G[x_t - \bar{\hat{x}}_{t-1,t}] + G_1 e_t)} \\ &= \frac{1-G}{G} - \frac{\frac{G_1^2}{C}\nu^{-1}}{G^2 var(x_t - \bar{\hat{x}}_{t-1,t}) + G_1^2 \nu^{-1}} \\ &= \frac{var(x_t - \bar{\hat{x}}_{t-1,t}) - [G var(x_t - \bar{\hat{x}}_{t-1,t}) + \frac{G_1^2}{G}\nu^{-1}]}{G var(x_t - \bar{\hat{x}}_{t-1,t}) + \frac{G_1^2}{G}\nu^{-1}} \\ &= \frac{\bar{\hat{\Sigma}} - [G\bar{\hat{\Sigma}} + \frac{G_1^2}{G}\nu^{-1}]}{G\bar{\hat{\Sigma}} + \frac{G_1^2}{G}\nu^{-1}}\end{aligned} \tag{49}$$

with $\bar{\hat{\Sigma}} \equiv var(x_t - \bar{\hat{x}}_{t-1,t})$. Notice that $\bar{\hat{\Sigma}} \neq \bar{\Sigma} \equiv var(x_t - \bar{x}_{t-1,t})$.

Consider the special case of x_t following an AR(1) process $x_t = \rho_i x_{t-1} + u_t$ with $u_t \sim N(0, \xi^{-1})$ and assuming $\hat{x}_{t+1,t} = \rho \hat{x}_{t,t}$. In steady state

$$\begin{aligned}x_{t+1} - \rho\bar{\hat{x}}_{t,t} =& \rho(x_t - \bar{\hat{x}}_{t,t}) + u_{t+1} \\ \hat{\Sigma} =& \rho^2 \bar{\hat{\Phi}} + \xi^{-1}\end{aligned} \tag{50}$$

where $\bar{\hat{\Phi}} = var(x_t - \hat{x}_{t,t})$.

$$\begin{aligned}x_t - \bar{\hat{x}}_{t,t} =& (1-G)[x_t - \rho\bar{\hat{x}}_{t-1,t-1}] - G_1 e_t \\ \bar{\hat{\Phi}} =& (1-G)^2 \bar{\hat{\Sigma}} + G_1^2 \nu^{-1}\end{aligned} \tag{51}$$

Substitute and solve for $\hat{\Sigma}$

$$\bar{\hat{\Sigma}} = \frac{\rho^2[C_1^2 \nu^{-1}] + \xi^{-1}}{1 - \rho^2(1-G)^2} \tag{52}$$

■

Corollary 2. With rational expectation, $G_1 = \frac{\nu}{\tau+\nu+\Sigma^{-1}}$ and $G_2 = \frac{\tau}{\tau+\nu+\Sigma^{-1}}$, with $\Sigma \equiv var(x_t - E_t^i[x_t])$ and $\bar{\Sigma} \equiv var(x_t - \bar{E}_t[x_t])$. From 49, it follows that

- if $\nu = 0$, $G_1 = 0$ and $\beta_{CG} = \frac{1-G_2}{G_2}$. Moreover, if $\tau \to \infty$, $G_2 = 1$ and $\beta_{CG} = 0$.
- if $\tau = 0$, $\Sigma = \bar{\Sigma}$, $G = G_1$ and $\frac{1-G}{G} = \frac{\Sigma^{-1}}{\nu^{-1}}$. Therefore $\beta_{CG} = 0$.

■

Corollary 3. From 49, $\frac{1}{1+\beta_{CG}}$ is given by

$$\begin{aligned}\frac{1}{1+\beta_{CG}} &= \frac{1}{1+\frac{1-G}{G}-\frac{\frac{G_1^2}{G}\nu^{-1}}{G^2\hat{\Sigma}+G_1^2\nu^{-1}}} \\ &= G\left(\frac{G^2\hat{\bar{\Sigma}}+G_1^2\nu^{-1}}{G^2\hat{\bar{\Sigma}}}\right) > 0\end{aligned} \tag{53}$$

which is equal to G if $G_1^2 = 0$. Subtracting the actual gain G

$$\begin{aligned}G\left(\frac{G^2\hat{\bar{\Sigma}}+G_1^2\nu^{-1}}{G^2\hat{\bar{\Sigma}}}-1\right) \\ G\left(\frac{G_1^2\nu^{-1}}{G^2\hat{\bar{\Sigma}}}\right) > 0\end{aligned} \tag{54}$$

■

Proposition 3. Let $\hat{x}_{t,t+h} \equiv \tilde{E}_t[x_{t+h}]$. Without loss of generality, assume $h = 0$. From 2

$$\begin{aligned}\hat{x}^i_{t,t} - \hat{x}^i_{t-1,t} &= G(x_t - \hat{x}^i_{t-1,t}) + G_1 e_t + G_2\eta^i_t \\ &= G_2(x_t - \hat{x}^i_{t-1,t}) + G_1(g_t - \hat{x}^i_{t-1,t}) + G_2\eta^i_t \\ &= G_2(x_t - \hat{x}^i_{t,t}) + G_2(\hat{x}^i_{t,t} - \hat{x}^i_{t-1,t}) + G_1(g_t - \hat{x}^i_{t-1,t}) + G_2\eta^i_t \\ \underbrace{(x_t - \hat{x}^i_{t,t})}_{fe^i_{t,t}} &= \frac{1-G_2}{G_2}\underbrace{(\hat{x}^i_{t,t} - \hat{x}^i_{t-1,t})}_{fr^i_{t,t}} - \frac{G_1}{G_2}\underbrace{(g_t - \hat{x}^i_{t-1,t})}_{pi^i_{t,t}} - \eta^i_t\end{aligned} \tag{55}$$

Write regression 12 as

$$fe^i_{t,t} = X\beta + err^i_t \tag{56}$$

where $X = [fr^i_{t,t} \quad pi^i_{t,t}]$ and $\beta = \begin{bmatrix}\beta_1 \\ \beta_2\end{bmatrix}$.

$$\hat{\beta} = \beta + \Sigma^{-1}_{XX}\Sigma_{Xu} \tag{57}$$

where

$$\begin{aligned}
\Sigma_{XX} &= \begin{bmatrix} var(fr^i_{t,t}) & cov(fr^i_{t,t}, pi^i_{t,t}) \\ cov(fr^i_{t,t}, pi^i_{t,t}) & var(pi^i_{t,t}) \end{bmatrix} \\
\Sigma^{-1}_{XX} &= \frac{1}{var(fr^i_{t,t})var(pi^i_{t,t}) - cov(fr^i_{t,t}, pi^i_{t,t})^2} \begin{bmatrix} var(pi^i_{t,t}) & -cov(fr^i_{t,t}, pi^i_{t,t}) \\ -cov(fr^i_{t,t}, pi^i_{t,t}) & var(fr^i_{t,t}) \end{bmatrix} \\
\Sigma_{Xu} &= \begin{bmatrix} cov(fr^i_{t,t}, err^i) \\ cov(pi^i_{t,t}, err^i) \end{bmatrix} \\
\hat{\beta} &= \beta + \Sigma^{-1}_{XX}\Sigma_{Xu} = \beta + \begin{bmatrix} \frac{var(pi^i_{t,t})cov(fr^i_{t,t}, err)}{var(fr^i_{t,t})var(pi^i_{t,t}) - cov(fr^i_{t,t}, pi^i_{t,t})^2} \\ \frac{-cov(fr^i_{t,t}, pi^i_{t,t})cov(fr^i_{t,t}, err)}{var(fr^i_{t,t})var(pi^i_{t,t}) - cov(fr^i_{t,t}, pi^i_{t,t})^2} \end{bmatrix}
\end{aligned} \tag{58}$$

and

$$\begin{aligned}
var(fr^i_{t,t}) &= [G^2\hat{\Sigma} + G_1^2\nu^{-1} + G_2^2\tau^{-1}] \\
var(pi^i_{t,t}) &= \hat{\Sigma} + \nu^{-1} \\
cov(fr^i_{t,t}, pi^i_{t,t}) &= [G\hat{\Sigma} + G_1\nu^{-1}] \\
cov(fr^i_{t,t}, err^i) &= -G_2\tau^{-1} \\
cov(pi^i_{t,t}, err^i) &= 0
\end{aligned} \tag{59}$$

where $\hat{\chi} = \hat{\Sigma}^{-1}$ therefore

$$\begin{aligned}
\hat{\beta}_1 &= \frac{1 - G_2}{G_2} + \frac{var(pi^i_{t,t})cov(fr^i_{t,t}, err)}{var(fr^i_{t,t})var(pi^i_{t,t}) - cov(fr^i_{t,t}, pi^i_{t,t})^2} \\
&= \frac{1 - G_2}{G_2} - \frac{(\hat{\Sigma} + \nu^{-1})G_2\frac{1}{\tau}}{(\hat{\Sigma} + \nu^{-1})(G^2\hat{\Sigma} + G_1^2\nu^{-1} + G_2^2\tau^{-1}) - (G\hat{\Sigma} + G_1\nu^{-1})^2}
\end{aligned} \tag{60}$$

and

$$\begin{aligned}
\hat{\beta}_2 &= -\frac{G_1}{G_2} + \frac{-cov(fr^i_{t,t}, pi^i_{t,t})cov(fr^i_{t,t}, err)}{var(fr^i_{t,t})var(pi^i_{t,t}) - cov(fr^i_{t,t}, pi^i_{t,t})^2} \\
&= -\frac{G_1}{G_2} + \frac{(G\hat{\Sigma} + G_1\nu^{-1})G_2\frac{1}{\tau}}{(\hat{\Sigma} + \nu^{-1})(G^2\hat{\Sigma} + G_1^2\nu^{-1} + G_2^2\tau^{-1}) - (G\hat{\Sigma} + G_1\nu^{-1})^2}
\end{aligned} \tag{61}$$

■

Corollary 4. With rational expectation, $G_1 = \frac{\nu}{\tau+\nu+\Sigma^{-1}}$ and $G_2 = \frac{\tau}{\tau+\nu+\Sigma^{-1}}$, with

$\Sigma \equiv var(x_t - E_t^i[x_t])$. Define $\chi = \Sigma^{-1}$.

From 60

$$
\begin{aligned}
\beta_1 &= \frac{\chi+\nu}{\tau} - \frac{(\chi^{-1}+\nu^{-1})G_2\frac{1}{\tau}}{(\chi^{-1}+\nu^{-1})(G^2\chi^{-1}+G_1^2\nu^{-1}+G_2^2\tau^{-1})-(G\chi^{-1}+G_1\nu^{-1})^2} \\
&= \frac{\chi+\nu}{\tau} - \frac{(\frac{1}{\chi+\nu+\tau})(\frac{\nu+\chi}{\nu\chi})}{(\frac{1}{\chi+\nu+\tau})^2[(\frac{\nu+\chi}{\nu\chi})(\frac{(\nu+\tau)^2}{\chi}+\nu+\tau)-((\frac{(\nu+\tau)}{\chi})+1)^2]} \\
&= \frac{\chi+\nu}{\tau} - \frac{(\frac{\nu+\chi}{\nu\chi})(\chi+\nu+\tau)}{(\frac{\nu+\chi}{\nu\chi})(\nu+\tau)(\chi+\nu+\tau)\frac{1}{\chi}-(\chi+\nu+\tau)^2\frac{1}{\chi^2}} \\
&= \frac{\chi+\nu}{\tau} - \frac{(\nu+\chi)}{(\chi+\nu+\tau)-(\nu+\chi)} \\
&= 0
\end{aligned} \tag{62}
$$

While from 61

$$
\begin{aligned}
\hat{\beta}_2 &= -\frac{\nu}{\tau} + \frac{(G\chi^{-1}+G_1\nu^{-1})G_2\frac{1}{\tau}}{(\chi^{-1}+\nu^{-1})(G^2\chi^{-1}+G_1^2\nu^{-1}+G_2^2\tau^{-1})-(G\chi^{-1}+G_1\nu^{-1})^2} \\
&= -\frac{\nu}{\tau} + \frac{(\frac{1}{\chi+\nu+\tau})^2(\frac{\nu+\tau}{\chi}+1)}{(\frac{1}{\chi+\nu+\tau})^2[(\frac{\nu+\chi}{\nu\chi})(\frac{(\nu+\tau)^2}{\chi}+\nu+\tau)-((\frac{(\nu+\tau)}{\chi})+1)^2]} \\
&= -\frac{\nu}{\tau} + \frac{(\chi+\nu+\tau)\frac{1}{\chi}}{(\frac{\nu+\chi}{\nu\chi})(\nu+\tau)(\chi+\nu+\tau)\frac{1}{\chi}-(\chi+\nu+\tau)^2\frac{1}{\chi^2}} \\
&= -\frac{\nu}{\tau} + \frac{\nu\chi}{(\nu+\chi)(\nu+\tau)-(\nu+\tau+\chi)\nu} \\
&= 0
\end{aligned} \tag{63}
$$

■

Proposition 4. From 21, using $\delta = \delta_1 + \delta_2$

$$
\begin{aligned}
\hat{x}^i &= \mu + \delta(x-\mu) + \delta_1 e + \delta_2 \eta^i \\
\hat{x}^i &= \mu + \delta x - \delta\hat{x}^i + \delta(\hat{x}^i-\mu) + \delta_1 e + \delta_2\eta^i \\
(1-\delta)(\hat{x}^i-\mu) &= +\delta(x-\hat{x}^i) + \delta_1 e + \delta_2\eta^i \\
(\underbrace{x-\hat{x}^i}_{fe^i}) &= \frac{1-\delta}{\delta}(\underbrace{\hat{x}^i-\mu}_{fr^i}) - \frac{\delta_1}{\delta}e - \frac{\delta_2}{\delta}\eta^i
\end{aligned} \tag{64}
$$

therefore by running BGMS regression 3, the regressor $fr^i = \hat{x}^i - \mu$ is correlated with the unobservable error. The resulting $\hat{\beta}_{BGMS}$ is equal to:

$$\begin{aligned}
\hat{\beta}_{BGMS} &= \frac{1-\delta}{\delta} + \frac{cov(\hat{x}^i - \mu, -\frac{\delta_1}{\delta}e - \frac{\delta_2}{\delta}\eta^i)}{var(\hat{x}^i - \mu)} \\
&= \frac{1-\delta}{\delta} + \frac{cov(\delta(x_t - \mu) + \delta_1 e + \delta_2 \eta^i, -\frac{\delta_1}{\delta}e - \frac{\delta_2}{\delta}\eta^i)}{var(\delta(x_t - \mu) + \delta_1 e + \delta_2 \eta^i)} \\
&= \frac{1-\delta}{\delta} + \frac{-\frac{\delta_1^2}{\delta}\nu^{-1} - \frac{\delta_2^2}{\delta}\tau^{-1}}{\delta^2\chi^{-1} + \delta_1^2\nu^{-1} + \delta_2^2\tau^{-1}}
\end{aligned} \tag{65}$$

substitute for δ_1 and δ_2

$$\hat{\beta}_{BGMS} = \frac{(1-\lambda)(1-\gamma_1-\gamma_2)}{(1-\lambda)\gamma_1+\gamma_2} - \frac{\frac{(1-\lambda)^2}{(1-\lambda)+\lambda\gamma_2}\frac{\gamma_1^2}{(1-\lambda)\gamma_1+\gamma_2}\nu^{-1} + \frac{1}{(1-\lambda)+\lambda\gamma_2}\frac{\gamma_2^2}{(1-\lambda)\gamma_1+\gamma_2}\tau^{-1}}{\frac{1}{[(1-\lambda)+\lambda\gamma_2]^2}([(1-\lambda)\gamma_1+\gamma_2]^2\chi^{-1} + (1-\lambda)^2\gamma_1^2\nu^{-1} + \gamma_2^2\tau^{-1})} \tag{66}$$

use definition of γ_1 and γ_2

$$\begin{aligned}
\hat{\beta}_{BGMS} &= \frac{(1-\lambda)\chi}{(1-\lambda)\nu+\tau} - \frac{\frac{1}{(1-\lambda)\nu+\tau}[(1-\lambda)^2\nu+\tau]}{\frac{1}{(1-\lambda)(\nu+\chi)+\tau}([[(1-\lambda)\nu+\tau]^2\chi^{-1} + (1-\lambda)^2\nu+\tau)} \\
&= \frac{(1-\lambda)\chi}{(1-\lambda)\nu+\tau} - \frac{[(1-\lambda)(\nu+\chi)+\tau][(1-\lambda)^2\nu+\tau]\chi}{[(1-\lambda)\nu+\tau]([[(1-\lambda)\nu+\tau]^2 + [(1-\lambda)^2\nu+\tau]\chi)} \\
&= \frac{\chi\{(1-\lambda)[(1-\lambda)\nu+\tau]^2 - [(1-\lambda)\nu+\tau][(1-\lambda)^2\nu+\tau]\}}{[(1-\lambda)\nu+\tau]([[(1-\lambda)\nu+\tau]^2 + [(1-\lambda)^2\nu+\tau]\chi)} \\
&= \frac{-\lambda\tau\chi[(1-\lambda)\nu+\tau]}{[(1-\lambda)\nu+\tau]([[(1-\lambda)\nu+\tau]^2 + [(1-\lambda)^2\nu+\tau]\chi)} \\
& \frac{-\lambda\tau\chi}{([[(1-\lambda)\nu+\tau]^2 + [(1-\lambda)^2\nu+\tau]\chi)} < 0
\end{aligned} \tag{67}$$

which is negative as long as $0 < \lambda < 1$. ■

Proposition 5. From 21

$$\begin{aligned}
\bar{\hat{x}} &= \mu + \delta(x-\mu) + \delta_1 e \\
\bar{\hat{x}} &= \mu + \delta x - \delta\bar{\hat{x}} + \delta(\bar{\hat{x}} - \mu) + \delta_1 e \\
(1-\delta)(\bar{\hat{x}} - \mu) &= +\delta(x - \bar{\hat{x}}) + \delta_1 e \\
(\underbrace{x - \bar{\hat{x}}}_{fe^i}) &= \frac{1-\delta}{\delta}(\underbrace{\bar{\hat{x}} - \mu}_{fr^i}) - \frac{\delta_1}{\delta}e
\end{aligned} \tag{68}$$

therefore by running CG regression 5, the regressor $\bar{fr} = \bar{\hat{x}} - \mu$ is correlated with the unobservable error. The resulting $\hat{\beta}_{CG}$ is equal to:

$$\begin{aligned}\hat{\beta}_{CG} &= \frac{1-\delta}{\delta} + \frac{cov(\bar{\hat{x}} - \mu, -\frac{\delta_1}{\delta}e)}{var(\bar{\hat{x}} - \mu)} \\ &= \frac{1-\delta}{\delta} + \frac{cov(\delta(x_t - \mu) + \delta_1 e, -\frac{\delta_1}{\delta}e)}{var(\delta(x_t - \mu) + \delta_1 e)} \\ &= \frac{1-\delta}{\delta} + \frac{-\frac{\delta_1^2}{\delta}\nu^{-1}}{\delta^2\chi^{-1} + \delta_1^2\nu^{-1}}\end{aligned} \tag{69}$$

substitute for δ_1 and δ_2

$$\hat{\beta}_{CG} = \frac{(1-\lambda)(1-\gamma_1-\gamma_2)}{(1-\lambda)\gamma_1+\gamma_2} - \frac{\frac{(1-\lambda)^2}{(1-\lambda)+\lambda\gamma_2}\frac{\gamma_1^2}{(1-\lambda)\gamma_1+\gamma_2}\nu^{-1}}{\frac{1}{[(1-\lambda)+\lambda\gamma_2]^2}([(1-\lambda)\gamma_1+\gamma_2]^2\chi^{-1} + (1-\lambda)^2\gamma_1^2\nu^{-1})} \tag{70}$$

use definition of γ_1 and γ_2

$$\begin{aligned}\hat{\beta}_{CG} &= \frac{(1-\lambda)\chi}{(1-\lambda)\nu+\tau} - \frac{\frac{1}{(1-\lambda)\nu+\tau}[(1-\lambda)^2\nu]}{\frac{1}{(1-\lambda)(\nu+\chi)+\tau}([(1-\lambda)\nu+\tau]^2\chi^{-1} + (1-\lambda)^2\nu)} \\ &= \frac{(1-\lambda)\chi}{(1-\lambda)\nu+\tau} - \frac{[(1-\lambda)(\nu+\chi)+\tau](1-\lambda)^2\nu\chi}{[(1-\lambda)\nu+\tau]([(1-\lambda)\nu+\tau]^2 + (1-\lambda)^2\nu\chi)} \\ &= \frac{(1-\lambda)\tau\chi[(1-\lambda)\nu+\tau]}{[(1-\lambda)\nu+\tau]([(1-\lambda)\nu+\tau]^2 + (1-\lambda)^2\nu\chi)} \\ &\frac{(1-\lambda)\tau\chi}{([(1-\lambda)\nu+\tau]^2 + (1-\lambda)^2\nu\chi)} > 0\end{aligned} \tag{71}$$

which is positive as long as $0 < \lambda < 1$. If $\lambda = 1$, it is zero. ■

Proposition 6. From 21

$$\begin{aligned}x^i &= \mu + \delta_1(y-\mu) + \delta_2(x-\mu) + \delta_2\eta^i \\ x^i &= \mu + \delta_1(y-\mu) + \delta_2 x_t - \delta_2 x^i + \delta_2(\hat{x}^i - \mu) + \delta_2\eta^i \\ (1-\delta_2)(\hat{x}^i - \mu) &= \delta_1(y-\mu) + \delta_2(x - \hat{x}^i) + \delta_2\eta^i \\ (\underbrace{x - \hat{x}^i}_{fe^i}) &= \frac{1-\delta_2}{\delta_2}(\underbrace{\hat{x}^i - \mu}_{fr^i}) - \frac{\delta_1}{\delta_2}(\underbrace{g - \mu}_{pi}) - \eta^i\end{aligned} \tag{72}$$

write regression 12 as

$$fe^i = X\beta + err^i \tag{73}$$

where $X = [fr^i \quad pi]$ and $\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$.

$$\hat{\beta} = \beta + \Sigma_{XX}^{-1}\Sigma_{Xu} \tag{74}$$

where

$$\begin{aligned}
\Sigma_{XX} &= \begin{bmatrix} var(fr^i) & cov(fr^i, pi) \\ cov(fr^i, pi) & var(pi) \end{bmatrix} \\
\Sigma_{XX}^{-1} &= \frac{1}{var(fr^i)var(pi) - cov(fr^i, pi)^2} \begin{bmatrix} var(pi) & -cov(fr^i, pi) \\ -cov(fr^i, pi) & var(fr^i) \end{bmatrix} \\
\Sigma_{Xu} &= \begin{bmatrix} cov(fr^i, err^i) \\ cov(pi, err^i) \end{bmatrix} \\
\hat{\beta} = \beta + \Sigma_{XX}^{-1}\Sigma_{Xu} &= \beta + \begin{bmatrix} \frac{var(pi)cov(fr,err)}{var(fr)var(pi)-cov(fr,pi)^2} \\ \frac{-cov(fr,pi)cov(fr,err)}{var(fr)var(pi)-cov(fr,pi)^2} \end{bmatrix}
\end{aligned} \tag{75}$$

and

$$\begin{aligned}
var(fr^i) &= \delta^2\chi^{-1} + \delta_1^2\nu^{-1} + \delta_2^2\tau^{-1} \\
var(pi) &= \chi^{-1} + \nu^{-1} \\
cov(fr^i, pi) &= \delta\chi^{-1} + \delta_1\nu^{-1} \\
cov(fr^i, err^i) &= -\delta_2\tau^{-1} \\
cov(pi, err^i) &= 0
\end{aligned} \tag{76}$$

$$
\begin{aligned}
\hat{\beta}_1 &= \frac{1-\delta_2}{\delta_2} + \frac{var(pi)cov(fr,err)}{var(fr)var(pi) - cov(fr,pi)^2} \\
&= (1-\lambda)\frac{1-\gamma_2}{\gamma_2} - \frac{(\chi^{-1}+\nu^{-1})\delta_2\tau^{-1}}{(\chi^{-1}+\nu^{-1})(\delta^2\chi^{-1}+\delta_1^2\nu^{-1}+\delta_2^2\tau^{-1}) - (\delta\chi^{-1}+\delta_1\nu^{-1})^2} \\
&= (1-\lambda)\frac{1-\gamma_2}{\gamma_2} - \\
&\frac{\frac{\chi+\nu}{\chi\nu}\gamma_2\tau^{-1}}{\frac{\chi+\nu}{\chi\nu}\frac{1}{(1-\lambda)+\lambda\gamma_2}([(1-\lambda)\gamma_1+\gamma_2]^2\chi^{-1}+(1-\lambda)^2\gamma_1^2\nu^{-1}+\gamma_2^2\tau^{-1}) - \frac{1}{(1-\lambda)+\lambda\gamma_2}([(1-\lambda)\gamma_1+\gamma_2]\chi^{-1}+(1-\lambda)\gamma_1}
\end{aligned}
\tag{77}
$$

use definition of γ_1 and γ_2

$$
\begin{aligned}
\hat{\beta}_1 &= (1-\lambda)\frac{\nu+\chi}{\tau} - \frac{(\chi+\nu)[(1-\lambda)(\nu+\chi)+\tau]\chi}{(\chi+\nu)([(1-\lambda)\nu+\tau]^2+[(1-\lambda)^2\nu+\tau]\chi) - ([(1-\lambda)\nu+\tau]+(1-\lambda)^2\nu\chi)^2\nu} \\
&= (1-\lambda)\frac{\nu+\chi}{\tau} - \frac{(\chi+\nu)[(1-\lambda)(\nu+\chi)+\tau]\chi}{\chi\tau(\tau+\nu+\chi)} \\
&= \frac{-\lambda(\nu+\chi)\chi\tau}{\chi\tau(\tau+\nu+\chi)} \\
&= \frac{-\lambda(\nu+\chi)}{(\tau+\nu+\chi)}
\end{aligned}
\tag{78}
$$

negative as long as $0 < \lambda < 1$.

$$
\begin{aligned}
\hat{\beta}_2 &= -\frac{\delta_1}{\delta_2} + \frac{-cov(fr,pi)cov(fr,err)}{var(fr)var(pi) - cov(fr,pi)^2} \\
&= -\frac{\delta_1}{\delta_2} + \frac{(\delta\chi^{-1}+\delta_1\nu^{-1})\delta_2\frac{1}{\tau}}{(\chi^{-1}+\nu^{-1})(\delta^2\chi^{-1}+\delta_1^2\nu^{-1}+\delta_2^2\tau^{-1}) - (\delta\chi^{-1}+\delta_1\nu^{-1})^2} \\
&= -(1-\lambda)\frac{\gamma_1}{\gamma_2} - \\
&\frac{[(1-\lambda)\nu+\tau]\chi^{-1}+(1-\lambda)}{\frac{\chi+\nu}{\chi\nu}([(1-\lambda)\gamma_1+\gamma_2]^2\chi^{-1}+(1-\lambda)^2\gamma_1^2\nu^{-1}+\gamma_2^2\tau^{-1}) - ([(1-\lambda)\gamma_1+\gamma_2]\chi^{-1}+(1-\lambda)\gamma_1\nu^{-1})^2}
\end{aligned}
\tag{79}
$$

use definition of γ_1 and γ_2

$$\begin{aligned}\hat{\beta}_2 &= -(1-\lambda)\frac{\nu}{\tau} + \frac{([(1-\lambda)\nu+\tau]+(1-\lambda)\chi)\chi\nu}{(\chi+\nu)([(1-\lambda)\nu+\tau]^2+[(1-\lambda)^2\nu+\tau]\chi)-([(1-\lambda)\nu+\tau]+(1-\lambda)^2\nu\chi)^2\nu} \\ &= -(1-\lambda)\frac{\nu}{\tau} + \frac{([(1-\lambda)\nu+\tau]+(1-\lambda)\chi)\chi\nu}{\chi\tau(\tau+\nu+\chi)} \\ &= \frac{\lambda\nu\chi\tau}{\chi\tau(\tau+\nu+\chi)} \\ &= \frac{\lambda\nu}{(\tau+\nu+\chi)}\end{aligned} \tag{80}$$

positive as long as $0 < \lambda < 1$. ■

C Different public signal measure

In addition to our baseline measure in section 2, we use the current value of the forecasted series as an additional possible proxy for public signal . In particular, assume that the observable series y agents are asked to forecast depends on a latent unobservable factor x and some noise e. Moreover, agents receive some private noisy signal on it s_t^i.

$$\begin{aligned} y_t &= x_t + e_t \\ x_t &= \rho x_{t-1} + u_t \\ s_t^i &= x_t + \eta_t^i \end{aligned} \tag{81}$$

with u_t, e_t and η^i normally distributed with zero mean and $\rho < 1$. The observable contemporaneous y_t is a public noisy signal about the underlying fundamental x_t. This structure is consistent with CG and BGMS econometric specification as long as $\tilde{E}_t[y_{t+h}] = \tilde{E}_t[x_{t+h}]$.

To measure the contemporaneous public signal for financial series, we use the average value of the series in the same quarter up to the survey date, which is the second month of the quarter. On the other hand, macroeconomic series are released with some lag, therefore we use the first release of the previous period value, which is available at the time of the forecast. To capture the surprise component in the public information, we compute the difference between the public signal and individual prior about the signal. In this case $pi_{t,t+h} = y_t - E_{t-1}^i[y_t]$. For macroeconomic variables, we compare contemporaneous release of lagged value with lagged nowcasting. In thi case $pi_{t,t+h} = y_{t-1} - E_{t-1}^i[y_{t-1}]$.

We run regression 12 using this different measure of public information. Panel A of Table 12 reports the panel data regressions at 3 quarters horizon with individual fixed effects and the median from individual regressions. The tables displays consistent $\beta_{GV,1} < 0$ and $\beta_{GV,2} > 0$ across variables, though less consistently than in table 2 in section 2. The reason being that the measure of public signal considered here doesn't refer direction to horizon $h = 3$, but to horizon $h = 0$ or even ($h = -1$ for macro variable) and it is therefore less informative about longer horizons. Panel B of Table 12 reports the same regression using a shorter horizon $h = 2$ and shows that the result are much more consistent and significant.[14] Figure **??** shows the coefficients graphically.

[14] At both horizons forecasts about the Consumer Price Index seem to overreact to this measure of public information instead of underreacting. However, in unreported result we show that if we consider the actual

Table 12: Private and public information: alternative measure of public information

Panel A: 3 quarters horizon

	Revision				Public signal			
	β_1	SE	p-value	Median	β_2	SE	p-value	Median
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nominal GDP	-0.25	0.08	0.00	-0.18	-0.05	0.12	0.69	-0.13
GDP price index inflation	-0.40	0.04	0.00	-0.40	0.39	0.15	0.01	0.30
Real GDP	-0.10	0.08	0.22	0.06	0.02	0.18	0.90	-0.09
Consumer Price Index	-0.19	0.08	0.03	-0.14	-0.56	0.28	0.06	-0.52
Industrial production	-0.30	0.14	0.03	-0.35	0.08	0.14	0.57	0.11
Housing Start	-0.09	0.09	0.36	-0.13	0.57	0.13	0.00	0.37
Real Consumption	-0.30	0.12	0.01	-0.25	0.27	0.13	0.06	0.15
Real residential investment	-0.09	0.10	0.39	-0.07	0.57	0.18	0.00	0.48
Real nonresidential investment	0.06	0.14	0.65	0.18	0.20	0.22	0.38	0.14
Real state and local government consumption	-0.53	0.05	0.00	-0.53	0.12	0.10	0.24	0.17
Real federal government consumption	-0.47	0.04	0.00	-0.39	0.28	0.09	0.00	0.19
Unemployment rate	0.26	0.16	0.10	0.18	-0.39	0.25	0.12	-0.44
Three-month Treasury rate	-0.26	0.10	0.02	-0.31	0.93	0.26	0.00	1.30
Ten-year Treasury rate	-0.63	0.05	0.00	-0.64	0.61	0.11	0.00	0.62
AAA Corporate Rate Bond	-0.69	0.04	0.00	-0.78	0.80	0.10	0.00	0.75

Panel B: 2 quarters horizon

	Revision				Public signal			
	β_1	SE	p-value	Median	β_2	SE	p-value	Median
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nominal GDP	-0.14	0.09	0.11	-0.10	0.12	0.12	0.34	0.04
GDP price index inflation	-0.41	0.04	0.00	-0.38	0.46	0.12	0.00	0.34
Real GDP	-0.13	0.10	0.21	0.07	0.22	0.17	0.22	-0.04
Consumer Price Index	-0.07	0.14	0.59	-0.12	-0.50	0.34	0.16	-0.54
Industrial production	-0.19	0.17	0.26	-0.15	0.35	0.22	0.12	0.32
Housing Start	0.03	0.06	0.67	-0.04	0.29	0.11	0.01	0.27
Real Consumption	-0.25	0.11	0.02	-0.21	0.21	0.13	0.11	0.14
Real residential investment	-0.09	0.09	0.32	-0.12	0.41	0.14	0.00	0.41
Real nonresidential investment	0.13	0.12	0.28	0.17	-0.02	0.20	0.94	-0.11
Real state and local government consumption	-0.40	0.04	0.00	-0.36	0.20	0.10	0.05	0.25
Real federal government consumption	-0.42	0.05	0.00	-0.33	0.29	0.10	0.00	0.08
Unemployment rate	0.22	0.12	0.06	0.20	-0.30	0.18	0.10	-0.28
Three-month Treasury rate	-0.33	0.14	0.02	-0.43	0.78	0.30	0.01	1.04
Ten-year Treasury rate	-0.80	0.06	0.00	-0.92	0.75	0.12	0.00	0.76
AAA Corporate Rate Bond	-0.77	0.05	0.00	-0.83	0.92	0.07	0.00	0.88

Notes: this table reports the coefficients of regression 12 (individual forecast errors on individual revisions and public information). Columns 1 to 3 show coefficient β_1 (forecast revision) from the panel regression with individual fixed effect, with standard errors and corresponding p-values. Standard errors are robust and clustered by forecaster. Column 4 shows the median coefficient of the same regression at the individual level. Columns 5 to 7 show coefficient β_2 (public information) from the panel regression with individual fixed effect, with standard errors and corresponding p-values. Standard errors are robust and clustered by forecaster. Column 8 shows the median coefficient of the same regression at the individual level. Panel A uses forecast at 3 quarters horizon and panel B uses forecast at 2 quarters horizon.

D Empirical evidence with AR2

Table 13: Motivating evidence: BGMS regressions with 2 lags

	$fr^i_{t+2,t}$				$fr^i_{t+1,t}$			
	$\beta_{BGMS,1}$	SE	p-value	Median	$\beta_{BGMS,2}$	SE	p-value	Median
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	
Nominal GDP	-0.24	0.14	0.10	-0.19	0.15	0.15	0.30	0.11
GDP price index inflation	-0.36	0.09	0.00	-0.38	0.27	0.13	0.04	0.36
Real GDP	-0.11	0.15	0.46	0.10	0.14	0.19	0.45	-0.01
Consumer Price Index	-0.89	0.18	0.00	-1.20	0.70	0.28	0.02	1.05
Industrial production	-0.30	0.19	0.12	-0.21	0.25	0.22	0.27	0.09
Housing Start	-0.26	0.15	0.09	-0.32	0.67	0.17	0.00	0.68
Real Consumption	-0.34	0.19	0.08	-0.35	0.34	0.20	0.10	0.29
Real residential investment	-0.54	0.19	0.01	-0.29	0.83	0.20	0.00	0.62
Real nonresidential investment	0.26	0.31	0.42	0.57	-0.05	0.38	0.91	-0.27
Real state and local government consumption	-0.10	0.13	0.46	-0.20	-0.06	0.16	0.70	0.01
Real federal government consumption	-0.46	0.13	0.00	-0.44	0.34	0.14	0.02	0.25
Unemployment rate	0.14	0.22	0.51	0.00	0.21	0.20	0.29	0.26
Three-month Treasury rate	-0.35	0.12	0.01	-0.58	0.75	0.21	0.00	1.25
Ten-year Treasury rate	-0.97	0.13	0.00	-0.96	0.85	0.16	0.00	0.76
AAA Corporate Rate Bond	-0.68	0.14	0.00	-1.08	0.54	0.20	0.01	0.84

value of GDP deflator as a public signal for consumer inflation (highly correlated with CPI), the forecasts underreact to it.

Table 14: Motivating evidence: public information regressions with 2 lags

	$fr^i_{t+2,t}$				$fr^i_{t+1,t}$				$pi^i_{t+2,t}$				$pi^i_{t+1,t}$			
	$\beta_{fr,1}$	SE	p-value	Median	$\beta_{fr,2}$	SE	p-value	Median	$\beta_{pi,1}$	SE	p-value	Median	$\beta_{pi,2}$	SE	p-value	Median
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Nominal GDP	-0.71	0.21	0.00	-0.55	1.12	0.14	0.00	1.25	0.38	0.21	0.09	0.16	-0.48	0.17	0.01	-0.67
GDP price index inflation	-0.85	0.13	0.00	-0.82	1.17	0.11	0.00	1.21	0.48	0.18	0.01	0.46	-0.57	0.14	0.00	-0.74
Real GDP	-0.62	0.25	0.02	-0.38	1.28	0.22	0.00	1.08	0.50	0.31	0.12	0.31	-0.88	0.33	0.01	-0.65
Consumer Price Index	-1.55	0.27	0.00	-1.53	1.55	0.26	0.00	1.85	1.27	0.39	0.00	1.43	-1.14	0.34	0.00	-1.61
Industrial production	-0.72	0.25	0.01	-0.66	1.08	0.16	0.00	0.78	0.49	0.26	0.07	0.22	-0.52	0.19	0.01	-0.22
Housing Start	-0.75	0.20	0.00	-0.78	1.03	0.14	0.00	1.03	0.97	0.23	0.00	0.84	-0.69	0.18	0.00	-0.63
Real Consumption	-0.72	0.25	0.01	-0.89	1.02	0.21	0.00	0.89	0.51	0.24	0.04	0.57	-0.40	0.23	0.10	-0.29
Real residential investment	-1.03	0.24	0.00	-0.79	1.45	0.21	0.00	1.47	1.17	0.27	0.00	0.90	-1.18	0.24	0.00	-0.73
Real nonresidential investment	0.08	0.38	0.84	0.26	0.63	0.32	0.06	0.95	-0.01	0.45	0.99	-0.34	-0.18	0.43	0.69	-0.66
Real state and local government consumption	-0.46	0.15	0.00	-0.29	0.95	0.19	0.00	0.98	-0.01	0.18	0.98	-0.15	-0.26	0.21	0.23	-0.32
Real federal government consumption	-1.10	0.16	0.00	-0.90	1.35	0.12	0.00	1.07	0.55	0.17	0.00	0.36	-0.62	0.15	0.00	-0.45
Unemployment rate	-0.15	0.30	0.62	-0.33	0.83	0.21	0.00	0.93	0.43	0.27	0.13	0.65	-0.55	0.20	0.01	-0.74
Three-month Treasury rate	-1.13	0.20	0.00	-1.75	1.64	0.19	0.00	2.06	1.43	0.33	0.00	1.98	-1.47	0.35	0.00	-2.31
Ten-year Treasury rate	-1.72	0.19	0.00	-1.68	1.91	0.17	0.00	1.98	1.39	0.25	0.00	1.45	-1.42	0.23	0.00	-1.53
AAA Corporate Rate Bond	-1.50	0.13	0.00	-1.59	1.64	0.15	0.00	1.47	1.10	0.24	0.00	1.09	-1.03	0.25	0.00	-0.90

E Comparison between public signals

Table 15: Comparison between public signals: 3 quarters horizon

	Revision				Public signal: consensus				Public signal: realization			
	β_1	SE	p-value	Median	β_2	SE	p-value	Median	β_3	SE	p-value	Median
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Nominal GDP	-0.54	0.12	0.00	-0.46	0.77	0.06	0.00	0.79	-0.16	0.12	0.19	-0.15
GDP price index inflation	-0.71	0.04	0.00	-0.66	0.79	0.04	0.00	0.85	0.27	0.14	0.07	0.22
Real GDP	-0.33	0.12	0.01	-0.18	0.57	0.09	0.00	0.62	-0.03	0.19	0.86	-0.12
Consumer Price Index	-0.36	0.12	0.01	-0.30	0.71	0.06	0.00	0.77	-0.59	0.30	0.06	-0.56
Industrial production	-0.61	0.14	0.00	-0.70	0.80	0.08	0.00	0.82	0.02	0.15	0.92	0.14
Housing Start	-0.41	0.13	0.00	-0.41	0.71	0.06	0.00	0.68	0.42	0.13	0.00	0.31
Real Consumption	-0.59	0.15	0.00	-0.58	0.80	0.08	0.00	0.84	0.23	0.13	0.10	0.14
Real residential investment	-0.36	0.14	0.01	-0.32	0.70	0.08	0.00	0.68	0.51	0.17	0.01	0.34
Real nonresidential investment	-0.12	0.18	0.50	-0.12	0.64	0.09	0.00	0.50	0.13	0.22	0.56	0.06
Real state and local government consumption	-0.84	0.05	0.00	-0.79	0.93	0.03	0.00	0.85	-0.03	0.11	0.81	-0.02
Real federal government consumption	-0.83	0.03	0.00	-0.75	0.89	0.04	0.00	0.83	0.09	0.09	0.30	0.05
Unemployment rate	0.11	0.20	0.58	0.00	0.54	0.11	0.00	0.50	-0.63	0.26	0.02	-0.55
Three-month Treasury rate	-0.44	0.11	0.00	-0.55	0.53	0.11	0.00	0.49	0.93	0.26	0.00	1.26
Ten-year Treasury rate	-0.86	0.07	0.00	-0.83	0.74	0.05	0.00	0.80	0.59	0.12	0.00	0.60
AAA Corporate Rate Bond	-0.90	0.06	0.00	-0.96	0.70	0.07	0.00	0.77	0.66	0.11	0.00	0.58

Notes: this table reports the coefficients of regression 12 (individual forecast errors on individual revisions and public information). Columns 1 to 3 show coefficient β_1 (forecast revision) from the panel regression with individual fixed effect, with standard errors and corresponding p-values. Standard errors are robust and clustered by forecaster. Column 4 shows the median coefficient of the same regression at the individual level. Columns 5 to 7 show coefficient β_2 (public information) from the panel regression with individual fixed effect, with standard errors and corresponding p-values. Standard errors are robust and clustered by forecaster. Column 8 shows the median coefficient of the same regression at the individual level. Panel A uses forecast at 3 quarters horizon and panel B uses forecast at 2 quarters horizon.

Table 16: Comparison between public signals: 2 quarters horizon

	Revision				Public signal: consensus				Public signal: realization			
	β_1	SE	p-value	Median	β_2	SE	p-value	Median	β_3	SE	p-value	Median
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Nominal GDP	-0.35	0.11	0.00	-0.31	0.63	0.05	0.00	0.61	0.00	0.13	0.98	0.05
GDP price index inflation	-0.66	0.04	0.00	-0.62	0.68	0.05	0.00	0.65	0.35	0.12	0.01	0.26
Real GDP	-0.29	0.13	0.03	-0.05	0.53	0.09	0.00	0.47	0.15	0.18	0.43	-0.09
Consumer Price Index	-0.18	0.16	0.27	-0.23	0.59	0.08	0.00	0.64	-0.62	0.36	0.10	-0.57
Industrial production	-0.39	0.18	0.03	-0.43	0.55	0.08	0.00	0.50	0.41	0.22	0.07	0.34
Housing Start	-0.21	0.08	0.02	-0.20	0.52	0.06	0.00	0.53	0.25	0.12	0.03	0.20
Real Consumption	-0.49	0.13	0.00	-0.42	0.66	0.06	0.00	0.69	0.25	0.13	0.06	0.16
Real residential investment	-0.28	0.12	0.02	-0.26	0.49	0.07	0.00	0.46	0.40	0.14	0.01	0.37
Real nonresidential investment	0.02	0.13	0.91	0.06	0.45	0.06	0.00	0.44	-0.05	0.19	0.80	-0.11
Real state and local government consumption	-0.70	0.06	0.00	-0.60	0.77	0.05	0.00	0.69	0.15	0.09	0.10	0.16
Real federal government consumption	-0.79	0.04	0.00	-0.64	0.78	0.04	0.00	0.71	0.23	0.10	0.02	0.08
Unemployment rate	0.10	0.15	0.51	0.11	0.48	0.09	0.00	0.48	-0.52	0.18	0.01	-0.44
Three-month Treasury rate	-0.44	0.12	0.00	-0.53	0.46	0.13	0.00	0.35	0.76	0.29	0.01	1.05
Ten-year Treasury rate	-0.97	0.07	0.00	-1.04	0.67	0.07	0.00	0.76	0.70	0.13	0.00	0.69
AAA Corporate Rate Bond	-0.88	0.05	0.00	-0.95	0.49	0.07	0.00	0.57	0.80	0.09	0.00	0.73

Notes: this table reports the coefficients of regression 12 (individual forecast errors on individual revisions and public information). Columns 1 to 3 show coefficient β_1 (forecast revision) from the panel regression with individual fixed effect, with standard errors and corresponding p-values. Standard errors are robust and clustered by forecaster. Column 4 shows the median coefficient of the same regression at the individual level. Columns 5 to 7 show coefficient β_2 (public information) from the panel regression with individual fixed effect, with standard errors and corresponding p-values. Standard errors are robust and clustered by forecaster. Column 8 shows the median coefficient of the same regression at the individual level. Panel A uses forecast at 3 quarters horizon and panel B uses forecast at 2 quarters horizon.

F Survey anonimity

The forecast data used in this paper are from the Survey of Professional Forecasters, compiled by the Federal Reserve of Philadelphia. Even if this particular survey is anonymous, we argue that it can nonetheless be affected by strategic incentives as well. In particular, we argue that the survey provided by forecasters to anonymous surveys appear to be the same as the one provided to other non-anonymous survey. This has been noted before in the forecasting literature: ”According to industry experts, forecasters often seem to submit to the anonymous surveys the same forecasts they have already prepared for public (i.e. non-anonymous) release. There are two reasons for this. First, it might not be convenient for the forecasters to change their report, unless they have a strict incentive to do so. Second, the forecasters might be concerned that their strategic behavior could be uncovered by the editor of the anonymous survey.” (Marinovic et al., 2013)

Two observations support this claim. First, Bordalo et al. (2020) establish fact 1 and 2 in section 2 by using both the SPF data and the Blue Chip data, which are not anonymous. They show that the two series provide very similar results, which is in line with the hypothesis of forecasters provided similar forecast to both surveys. Second, in a survey by the European Central Bank supplementary to their Survey of Professional Forecasters, respondents are asked explicitly ”When responding to the SPF, what forecast do you provide?”. In 2013, more than 80% of the panelists responded ”the last available, while in 2008 more than 90% gave the same answer (European Central Bank, 2014). It is also important to note that this is a conservative estimate of agents compiling a new forecast exclusively for the ECB survey, as the new forecast provided might be compiled to be used for other non-anonymous surveys as well.

G Dynamic model with AR(2)

We consider here a dynamic setting with a fundamental AR(2) process

$$
\begin{aligned}
x_t &= \rho_1 x_{t-1} + \rho_2 x_{t-2} + u_t \\
\begin{bmatrix} x_t \\ x_{t-1} \end{bmatrix} &= \begin{bmatrix} \rho_1 & \rho_2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ x_{t-2} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_t \\ e_t \end{bmatrix} \\
\bar{X}_t &= A\bar{X}_{t-1} + a \begin{bmatrix} u_t \\ e_t \end{bmatrix}
\end{aligned} \tag{82}
$$

With $u \sim N(0, \nu^{-1})$.

Each agent receive a private signal s_t^i and a public signal g_t

$$
\begin{aligned}
s_t^i &= x_t + \eta_t^i \\
g_t &= x_t + e_t
\end{aligned} \tag{83}
$$

with $\eta_t^i \sim N(0, \tau^{-1})$, $e_t \sim N(0, \nu^{-1})$. In matrix form

$$
V_t^i \equiv \begin{bmatrix} g_t \\ s_t^i \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \bar{X}_t + \begin{bmatrix} 0 \\ \eta_t^i \end{bmatrix} \tag{84}
$$

Honest beliefs Agents form beliefs about x at horizon h: $E_t^i[\bar{X}_{t+h}]$. The honest posterior belief about $\bar{X}$ is given by the Kalman filter

$$
E_t^i[\bar{X}_t] = AE_{t-1}^i[\bar{X}_{t-1}] + K(V_t^i - E_{t-1}^i[V_t])
$$

With the first line yields the posterior $E_t^i[x_t] \equiv x_{t,t}^i$

$$
x_{t,t}^i = x_{t,t-1}^i + K_{1,1}(g_t - x_{t,t-1}^i) + K_{1,2}(s_t^i - x_{t,t-1}^i)
$$

where the Kalman gains are

$$K_{1,1} = \frac{\nu}{\Sigma^{-1} + \nu + \tau}$$
$$K_{1,2} = \frac{\tau}{\Sigma^{-1} + \nu + \tau} \tag{85}$$

and the posterior forecast error variance

$$\Sigma \equiv E[(x_t - x^i_{t,t-1})(x_t - x^i_{t,t-1})'] \tag{86}$$

Strategic interactions As in the previous section, the strategic substitutability in agents objective function leads them to report

$$\begin{bmatrix} \hat{x}^i_{t,t} \\ \hat{x}^i_{t-1,t} \end{bmatrix} = \begin{bmatrix} \frac{1}{1-\lambda} E^i_t[x_t] - \frac{\lambda}{1-\lambda} E^i[\bar{\hat{x}}_{t,t}] \\ \frac{1}{1-\lambda} E^i_t[x_t - 1] - \frac{\lambda}{1-\lambda} E^i[\bar{\hat{x}}_{t-1,t}] \end{bmatrix}$$
$$F^i_t = \begin{bmatrix} \frac{1}{1-\lambda} & 0 & -\frac{\lambda}{1-\lambda} & 0 \\ 0 & \frac{1}{1-\lambda} & 0 & -\frac{\lambda}{1-\lambda} \end{bmatrix} E^i_t \begin{bmatrix} \bar{X}_t \\ F_t \end{bmatrix} \tag{87}$$

where $\hat{x}^i_{t+h,t}$ is the forecast provided by individual i in t about realization in $t+h$, and $\bar{\hat{x}}_{t+h,t} = \int^i \hat{x}^i_{t+h,t} di$ is the average of forecasts provided in t about realization in $t+h$. Define $F_t \equiv \begin{bmatrix} \bar{\hat{x}}_{t,t} \\ \bar{\hat{x}}_{t-1,t} \end{bmatrix}$ and $F^i_t \equiv \begin{bmatrix} \hat{x}^i_{t,t} \\ \hat{x}^i_{t-1,t} \end{bmatrix}$. If $\lambda = 0$, agents report their true beliefs. With $1 > \lambda > 0$, agents not only want to be accurate, but also to stand out with respect to the average forecast.

We average $\hat{x}^i_{t,t}$ across agents and use repeated substitution in 87 to express the reported average forecast as

$$F_t = -\frac{1}{1-\lambda} \sum_{k=0}^{\infty} \left(\frac{\lambda}{1-\lambda} \right)^k \bar{E}^{(k)}[\bar{X}_t] = \frac{1}{1-\lambda} \bar{E}_t \bar{X}_t - \frac{\lambda}{1-\lambda} F_t \tag{88}$$

We guess and verify the law of motion for F_t and the other unobserved state variables. In particular, we conjecture that the state vector evolves according to[15]

$$Z \equiv \begin{bmatrix} \bar{X}_t \\ F_t \\ w_t \end{bmatrix} = MZ_{t-1} + m \begin{bmatrix} u_t \\ e_t \end{bmatrix} \tag{89}$$

Where

$$M = \begin{bmatrix} \underset{2x2}{A} & \underset{2x2}{0} & \underset{2x1}{0} \\ \underset{2x2}{G} & \underset{2x2}{L} & \underset{2x1}{0} \\ \underset{1x2}{0} & \underset{1x2}{0} & \underset{1x1}{0} \end{bmatrix} \quad \text{and} \quad m = \begin{bmatrix} \underset{2x2}{a} \\ \underset{2x2}{\mu} \\ 0 \quad 1 \end{bmatrix} \tag{90}$$

the observable variables are the two signals about x_t

$$V_t^i \equiv \begin{bmatrix} g_t \\ s_t^i \end{bmatrix} = HZ_t + \begin{bmatrix} 0 \\ \eta_t^i \end{bmatrix} \tag{91}$$

where

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{92}$$

Agents use their conjecture law of motion 89 and the observables 91 to infer the state using the individual Kalman filter. The posterior estimate of the state vector by agent i is

$$\begin{aligned} E_t^i[Z_t] &= ME_{t-1}^i[Z_{t-1}] + K(V_t^i - E_{t-1}^i[V_t]) \\ &= (I - KH)ME_{t-1}^i[Z_{t-1}] + KHMZ_{t-1} + KHm \begin{bmatrix} u_t \\ e_t \end{bmatrix} + K \begin{bmatrix} 0 \\ \eta_t^i \end{bmatrix} \end{aligned} \tag{93}$$

Where K is the Kalman gain. Average 93 to find the consensus believe on the state vector.

$$\bar{E}_t[Z_t] = (I - KH)M\bar{E}_{t-1}[Z_{t-1}] + KHMZ_{t-1} + KHm \begin{bmatrix} u_t \\ e_t \end{bmatrix} \tag{94}$$

[15] w_t takes care of the correlation between public signal and higher order beliefs F_t

From the definition on F_t in 30 it follows that

$$\begin{aligned} F_t &= \begin{bmatrix} \frac{1}{1-\lambda} & 0 & -\frac{\lambda}{1-\lambda} & 0 \\ 0 & \frac{1}{1-\lambda} & 0 & -\frac{\lambda}{1-\lambda} \end{bmatrix} \bar{E}_t[Z_t] \equiv \xi \bar{E}_t[Z_t] \\ &= \xi(I - KH)M\bar{E}_{t-1}[Z_{t-1}] + \xi KHMZ_{t-1} + \xi KHm \begin{bmatrix} u_t \\ e_t \end{bmatrix} \end{aligned} \tag{95}$$

Compute (i) $\xi M\bar{E}_{t-1}[Z_{t-1}]$, (ii) $HM\bar{E}_{t-1}[Z_{t-1}]$, (iii) HMZ_{t-1}, (iv) Hm.

1. write ξ as a vector of matrices

$$\xi \equiv \left[\begin{bmatrix} \frac{1}{1-\lambda} & 0 \\ 0 & \frac{1}{1-\lambda} \end{bmatrix} \begin{bmatrix} -\frac{\lambda}{1-\lambda} & 0 \\ 0 & -\frac{\lambda}{1-\lambda} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right] \equiv \begin{bmatrix} \frac{1}{1-\lambda}I & -\frac{\lambda}{1-\lambda}I & 0 \end{bmatrix} \tag{96}$$

Then

$$\begin{aligned} \xi M\bar{E}_{t-1}[Z_{t-1}] &= \begin{bmatrix} \frac{1}{1-\lambda}I & -\frac{\lambda}{1-\lambda}I & 0 \end{bmatrix} \begin{bmatrix} A & 0 & 0 \\ G & L & 0 \\ 0 & 0 & 0 \end{bmatrix} \bar{E}_{t-1}[Z_{t-1}] \\ &= \begin{bmatrix} \frac{1}{1-\lambda}A - \frac{\lambda}{1-\lambda}G & -\frac{\lambda}{1-\lambda}L & 0 \end{bmatrix} \bar{E}_{t-1} \begin{bmatrix} \bar{X}_{t-1} \\ F_{t-1} \\ w_{t-1} \end{bmatrix} \\ &= \left(\frac{1}{1-\lambda}A - \frac{\lambda}{1-\lambda}G \right) \bar{E}_{t-1}[\bar{X}_{t-1}] - \frac{\lambda}{1-\lambda} L\bar{E}_{t-1}[F_{t-1}] \end{aligned} \tag{97}$$

2. write H as a vector of matrices $H \equiv \left[\begin{bmatrix} H_{1,1} & H_{1,2} \\ H_{1,1} & H_{1,2} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right] \equiv [H_1 \quad H_2 \quad H_3]$.

Then

$$
\begin{aligned}
HM\bar{E}_{t-1}[Z_{t-1}] &= [H_1 \quad H_2 \quad H_3] \begin{bmatrix} A & 0 & 0 \\ G & L & 0 \\ 0 & 0 & 0 \end{bmatrix} \bar{E}_{t-1}[Z_{t-1}] \\
&= [H_1 A \quad 0 \quad 0] \bar{E}_{t-1} \begin{bmatrix} \bar{X}_{t-1} \\ F_{t-1} \\ w_{t-1} \end{bmatrix} \\
&= H_1 A \bar{E}_{t-1}[\bar{X}_{t-1}]
\end{aligned} \tag{98}
$$

3. Similarly,

$$
HMZ_{t-1} = H_1 A \bar{X}_{t-1} \tag{99}
$$

4. Similarly

$$
Hm = H_1 a + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \tag{100}
$$

Substitute back in the posted KF, using that $-\frac{\lambda}{1-\lambda}\bar{E}_{t-1}[F_{t-1}] = F_{t-1} - \frac{1}{1-\lambda}\bar{E}_{t-1}[X_{t-1}]$. After some algebra, one gets

$$
\begin{aligned}
F_t = \left(\frac{1}{1-\lambda} A + -\frac{\lambda}{1-\lambda} G - \frac{1}{1-\lambda} L - \xi K H_1 A \right) \bar{E}_{t-1}[\bar{X}_{t-1}] + \xi K H_1 A \bar{X}_{t-1} + L F_{t-1} \\
+ \xi K \left(H_1 a + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \right) \begin{bmatrix} u_t \\ e_t \end{bmatrix}
\end{aligned} \tag{101}
$$

Equation 101 must equal the second line (a 2x1 vector) of the perceived law of motion 89. The solution to the fixed point is given by $G = \xi K H_1 A$, $\mu = \xi K \left(H_1 a + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \right)$ and $L = A - G$.

In particular, define

$$
\begin{aligned}
C_1 &\equiv \frac{K_{1,1} - \lambda(K_{3,1})}{1-\lambda}, \qquad C_2 \equiv \frac{K_{1,2} - \lambda K_{3,2}}{1-\lambda} \qquad \text{and} \qquad C = C_1 + C_2 \\
D_1 &\equiv \frac{K_{2,1} - \lambda(K_{4,1})}{1-\lambda}, \qquad D_2 \equiv \frac{K_{2,2} - \lambda K_{4,2}}{1-\lambda} \qquad \text{and} \qquad D = D_1 + D_2
\end{aligned}
$$

Then $G = \begin{bmatrix} \rho_1 C & \rho_2 C \\ \rho_1 D & \rho_2 D \end{bmatrix}$, $\mu = \begin{bmatrix} C & C_1 \\ D & 0 \end{bmatrix}$ and $L = \begin{bmatrix} \rho_1(1-C) & \rho_2(1-C) \\ 1-\rho_1 D & -\rho_2 D \end{bmatrix}$.

Given the law of motion of unobserved state 31 and the observable 33, the posterior variance of the forecast solves the following Ricatti equation

$$
\begin{aligned}
&\Sigma \equiv E[(Z_t - Z^i_{t,t-1})(Z_t - Z^i_{t,t-1})'] \\
&\Sigma = M(\Sigma - \Sigma H' \left(H\Sigma H' + \begin{bmatrix} 0 & 0 \\ 0 & \tau^{-1} \end{bmatrix} \right)^{-1} H\Sigma)M' + m \begin{bmatrix} \xi^{-1} & 0 \\ 0 & \nu^{-1} \end{bmatrix} m'
\end{aligned}
\tag{102}
$$

and the Kalman filter is

$$
K = \Sigma H' \left(H\Sigma H' + \begin{bmatrix} 0 & 0 \\ 0 & \tau^{-1} \end{bmatrix} \right)^{-1} \tag{103}
$$

Step 6: derive the action of individual With the model's solution, one can obtain the individual forecast as

$$
\begin{aligned}
F_t^i &= \xi \quad E_t^i[Z_t] \\
&= \xi(I-KH)ME_{t-1}^i[Z_{t-1}] + \xi KHMZ_{t-1} + \xi KHm \begin{bmatrix} u_t \\ e_t \end{bmatrix} \\
&= A \underbrace{\left(\frac{1}{1-\lambda}E_{t-1}^i[\bar{X}_{t-1}] - \frac{\lambda}{1-\lambda}E_{t-1}^i[F_{t-1}]\right)}_{F_{t-1}^i} + [-\frac{\lambda}{1-\lambda}G - G]E_{t-1}^i[x_{t-1}] + \frac{\lambda}{1-\lambda}GE_{t-1}^i[F_{t-1}] \\
&\quad - \xi KH_1AE_{t-1}^i[\bar{X}_{t-1}] + \xi KH_1A\bar{X}_{t-1} + \xi KH_1a \begin{bmatrix} u_t \\ e_t \end{bmatrix} + \xi K \begin{bmatrix} e_t \\ 0 \end{bmatrix} + \xi K \begin{bmatrix} 0 \\ \eta_t^i \end{bmatrix} \\
F_t^i - AF_{t-1}^i &= -\xi KH_1AF_{t-1}^i + \xi KH_1\bar{X}_t + \xi K \begin{bmatrix} e_t \\ 0 \end{bmatrix} + \xi K \begin{bmatrix} 0 \\ \eta_t^i \end{bmatrix} \\
F_t^i - AF_{t-1}^i &= \xi K \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} (\bar{X}_t - AF_{t-1}^i) + \xi K \begin{bmatrix} 1 \\ 0 \end{bmatrix} e_t + \xi K \begin{bmatrix} 0 \\ 1 \end{bmatrix} \eta_t^i \\
F_t^i - AF_{t-1}^i &= \begin{bmatrix} C \\ D \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} (\bar{X}_t - AF_{t-1}^i) + \begin{bmatrix} C_1 \\ D_1 \end{bmatrix} e_t + \begin{bmatrix} C_2 \\ D_2 \end{bmatrix} \eta_t^i
\end{aligned} \tag{104}
$$

consider the first line

$$\hat{x}_{t,t}^i - \hat{x}_{t,t-1}^i = C[x_t - \hat{x}_{t,t-1}^i] + C_1e_t + C_2\eta_t^i \tag{105}$$

Which is similar to the basic framework in section 2. Consider the second line

$$\hat{x}_{t-1,t}^i - \hat{x}_{t-1,t-1}^i = D[x_{t-1} - \hat{x}_{t-1,t-1}^i] + D_1e_t + D_2\eta_t^i \tag{106}$$

H Structural estimation at: 2 quarters horizon

Table 17: Estimasted parameters

Variable	ρ (1)	$\frac{\sigma_e}{\sigma_u}$ (2)	$\frac{\sigma_\eta}{\sigma_u}$ (3)	λ (4)
Nominal GDP	0.93	1.51	1.31	0.61
GDP price index inflation	0.90	1.10	1.08	0.32
Real GDP	0.80	1.20	1.19	0.38
Consumer Price Index	0.97	1.14	1.22	0.56
Industrial production	0.85	1.41	1.16	0.29
Housing Start	0.85	2.12	1.20	0.31
Real Consumption	0.73	1.05	1.32	0.39
Real residential investment	0.89	1.44	1.20	0.23
Real nonresidential investment	0.88	3.16	1.04	0.14
Real state and local government consumption	0.74	1.07	1.67	0.73
Real federal government consumption	0.77	1.11	1.61	0.69
Unemployment rate	0.97	3.15	1.03	-0.28
Three-month Treasury rate	0.94	3.16	1.03	0.06
Ten-year Treasury rate	0.83	1.48	1.39	0.69
AAA Corporate Rate Bond	0.85	2.13	1.53	0.83

Table 18: Moments in data and model

	Targeted moments						Untargeted moments					
	Mean Dispersion		C		β_1		β_{CG}		β_{BGMS}		β_2	
	Data	Model	Data	Model	Data	Model	Data	Model	Data	Model	Data	Model
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Nominal GDP	0.94	0.94	0.61	0.61	-0.35	-0.35	0.20	0.43	-0.11	-0.22	0.62	0.11
GDP price index inflation	0.34	0.34	0.70	0.70	-0.20	-0.20	0.48	0.20	-0.02	-0.06	0.44	0.10
Real GDP	0.69	0.69	0.63	0.63	-0.25	-0.25	0.33	0.28	-0.07	-0.10	0.54	0.11
Consumer Price Index	0.27	0.27	0.70	0.70	-0.38	-0.38	-0.05	0.21	-0.24	-0.14	0.51	0.20
Industrial production	2.49	2.49	0.59	0.59	-0.16	-0.16	0.33	0.41	-0.01	-0.09	0.49	0.05
Housing Start	75.76	75.76	0.53	0.53	-0.15	-0.15	0.91	0.75	0.12	-0.13	0.54	0.01
Real Consumption	0.34	0.34	0.63	0.63	-0.31	-0.31	0.12	0.16	-0.11	-0.07	0.61	0.16
Real residential investment	16.69	16.69	0.56	0.56	-0.13	-0.13	0.56	0.42	0.07	-0.07	0.49	0.04
Real nonresidential investment	5.02	5.02	0.61	0.59	-0.02	-0.05	0.53	0.67	0.10	-0.05	0.41	0.00
Real state and local government consumption	0.92	0.92	0.61	0.61	-0.65	-0.65	0.05	0.15	-0.24	-0.23	0.77	0.35
Real federal government consumption	4.40	4.40	0.60	0.60	-0.60	-0.60	-0.21	0.18	-0.27	-0.21	0.77	0.31
Unemployment rate	0.09	0.06	0.56	0.55	0.09	0.08	0.59	0.78	0.20	0.08	0.39	0.00
Three-month Treasury rate	0.21	0.12	0.63	0.60	0.02	-0.02	0.40	0.66	0.14	-0.02	0.48	0.00
Ten-year Treasury rate	0.12	0.12	0.60	0.60	-0.46	-0.46	-0.09	0.45	-0.24	-0.31	0.71	0.14
AAA Corporate Rate Bond	0.25	0.25	0.61	0.61	-0.49	-0.49	0.05	0.58	-0.22	-0.44	0.70	0.07

Table 19: Posted and honest moments

	Gain			Consensus MSE			Dispersion		
	Posted	Honest	Ratio	Posted	Honest	Ratio	Posted	Honest	Ratio
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Nominal GDP	0.61	0.49	0.80	0.27	0.60	2.22	0.94	0.41	0.44
GDP price index inflation	0.70	0.66	0.94	0.27	0.37	1.36	0.34	0.22	0.65
Real GDP	0.63	0.57	0.91	0.56	0.80	1.41	0.69	0.39	0.57
Consumer Price Index	0.70	0.62	0.89	0.13	0.23	1.86	0.27	0.10	0.39
Industrial production	0.59	0.54	0.92	1.71	2.29	1.34	2.49	1.79	0.72
Housing Start	0.53	0.47	0.87	35.99	50.84	1.41	75.76	57.87	0.76
Real Consumption	0.63	0.59	0.95	0.57	0.71	1.24	0.34	0.16	0.48
Real residential investment	0.56	0.53	0.94	13.81	17.09	1.24	16.69	12.95	0.78
Real nonresidential investment	0.59	0.57	0.95	2.08	2.43	1.17	5.02	4.65	0.93
Real state and local government consumption	0.61	0.55	0.89	0.73	1.10	1.51	0.92	0.11	0.12
Real federal government consumption	0.60	0.53	0.88	3.60	5.48	1.52	4.40	0.69	0.16
Unemployment rate	0.55	0.60	1.08	0.04	0.03	0.78	0.06	0.07	1.12
Three-month Treasury rate	0.60	0.59	0.98	0.05	0.05	1.07	0.12	0.11	0.97
Ten-year Treasury rate	0.60	0.44	0.73	0.03	0.08	2.48	0.12	0.04	0.29
AAA Corporate Rate Bond	0.61	0.31	0.51	0.02	0.10	4.58	0.25	0.06	0.24