Heterogeneous Responses to Signals and the Predictability of Forecast Errors

Symeon Taipliadis[†]

This Version: December 2025 [‡]

Abstract

Why do analysts underreact to news at the aggregate level yet overreact at the individual level? The literature explains underreaction with information frictions and overreaction with extrapolation from recent events. However, forecast revisions are weak predictors of forecast errors out-of-sample. I instead argue analysts overvalue private signals while underweighting public signals. To test this, I propose a novel method to disentangle responses to private signals from heterogeneous responses to public signals in high-frequency data. Using data from equity analysts' expectations of US firm revenues, I find that asymmetric responses to signals uncover behavioral bias that predicts forecast errors out-of-sample. (*JEL* C53, D84, E70, G17, G40)

Keywords: Rational Expectations, Forecast Errors, Reaction to signals, Behavioral bias, Out-of-sample prediction

[†]Department of Economics, University of Florida, 341 Matherly Hall, PO Box 117140, Gainesville, FL 32611 (email: s.taipliadis@ufl.edu).

[‡]First Version: March 2024. I thank Jonathan Adams, Eugenio Rojas, Min Fang, Gustavo Cortes, Mark Rush, Elias Dinopoulos, Steven Slutsky, Berkay Akyapi, Garrison Pollard, Douglas Turner, and participants at conferences for helpful feedback and suggestions for improvement.

1 Introduction

Full Information Rational Expectations (FIRE), the strict assumption that agents observe the entire information set and form rational beliefs, is fundamental to modern macroeconomic models. Yet, recent empirical examinations of the FIRE hypothesis provide mixed results, not reaching a consensus on whether the deviations consistently violate the full information (i.e., information frictions), the rational expectations (i.e., behavioral bias), or both assumptions (Angeletos et al., 2021, Bianchi et al., 2022, Eva and Winkler, 2023). More importantly, behavioral biases and information frictions may have different macroeconomic effects (Kohlhas and Walther, 2021, Coibion et al., 2022, Bianchi et al., 2021, Bordalo et al., 2018, Gennaioli et al., 2016).

At the core of the deviations from FIRE are the separate findings on underreaction to news at the consensus level in Coibion and Gorodnichenko (2015) (CG for short) and overreaction to news at the individual level in Bordalo et al. (2020b) (BGMS for short).[1] The literature conventionally uses information frictions to explain the underreaction and behavioral bias to explain the overreaction (Angeletos and Huo, 2021, Angeletos et al., 2021 and Kohlhas and Walther, 2021, among others).[2] However, while grounded in compelling theoretical foundation, this suggestion fails to explain why forecast revisions are weak predictors of forecast errors out-of-sample (Eva and Winkler, 2023).

An alternative hypothesis, which is the focus of this study, suggests analysts exhibit asymmetric responses to signals. In this direction are works from Broer and Kohlhas (2022), Bianchi et al. (2022), and Kučinskas and Peters (2024). At the core of this hypothesis is the interaction of information friction and behavioral bias. Because of information rigidities, agents are exposed to both public and private signals. Also, agents' expectations are driven by behavioral bias as their interpretation of these signals varies conditional on the type of the signal received. The bias drives them to underestimate the significance of public signals while overestimating the importance of private signals. In more detail, since public signals are widely disseminated among multiple agents, they are perceived as 'diluted' and an average agent discounts these signals' value. However, private signals are perceived as 'exclusive' driving an average agent to overvalue them.

I document, this hypothesis not only explains underreactions and overreactions from in-sample tests, but also reconciles mixed evidence from out-of-sample evaluations. First,

[1]CG regress analysts' forecast errors on forecast revisions from forecasts about macroeconomic variables at the consensus level and find that agents underreact to new information. BGMS run the same regression at the individual-analyst level and find that agents overreact to new information.

[2]A growing body of literature also focuses in this area. See for example, Rozsypal and Schlafmann (2023), Bordalo et al. (2022), Bordalo et al. (2018), Coibion et al. (2018) and Andolfatto et al. (2008) among others.

forecast revisions at the consensus level might effectively diversify revision components arising from private signals. Underreaction to news, as evidenced by the CG model, could therefore reflect analysts' tendency to underestimate the significance of public information. On the other hand, regressions at the individual level might expose overreaction to news given individuals' tendency to overrate the significance of private information. This channel may be a stronger predictor of forecast errors, obscuring underreaction to public signals.

Second, the mechanism explains why forecast revisions often fail to predict forecast errors out-of-sample. By not accounting for asymmetric responses to signals, any attempt to predict (and therefore diminish) forecast errors after controlling for analysts' information set is rather weak as the coefficient that identifies behavioral bias is not flexible to accommodate for different signal types that emerge in the testing sample. By decomposing forecast updates into revisions from different types of signals, the econometrician gets more accurate information of analysts' behavior without receiving excess information that analysts might have not been able to observe on real time when reported their forecasts. This helps account for asymmetries and provide more accurate predictions than analysts' reported forecasts if they are distorted by bias.

The key difference between the prevailing and this alternative hypothesis is summarized by two points. First, this hypothesis does not focus on agents' ability to process contemporaneous and lagged information in order to explain underreactions (i.e., sticky information). It instead emphasizes analysts' behavior to explain this observation. Analysts underestimate the degree to which these signals are informative about future outcomes. Second, analysts do not necessarily extrapolate all signals in their attention, but only those they believe are more informative as they perceive to have smaller noise. As a result, the presence of noisy information and analysts' non-rational prioritization of exclusive news sources jointly generates the bias seen in the data.

Definitions. — Throughout this study I define a public signal–received within a window of k periods–as the accumulated new information that was announced at any time from $t-k$ to t and is added to all individuals' information set (i.e., is publicly available). A private signal is the accumulated new information that an individual receives at any time from $t-k$ to t but this information is not available to every individual's information set. It follows that if all individuals' beliefs were rational, the public signal would reflect changes in consensus's expectations from $t-k$ to t while the private signal would be traced by deviations of an individual's beliefs from the consensus at time t.[3] Revisions from public (private) signals are an analyst's updates of their reported forecasts upon observing public (private) signals. Finally, I use the terms *correction revisions* and *forecast corrections* interchangeably to

[3]I discuss the identification of the signals in more detail in Section 2.2.

denote an analyst's forecast updates resulting from past disagreements with the *rational* consensus.[4] The forecast horizon, H, is defined as the number of quarters the forecast was reported ahead of the firm-revenue announcement and the revision window, k, is the number of days between the initial forecast announcement at $t - k$ and the updated forecast announcement at t.

Contribution. — The contribution of this paper is twofold. First, I propose an empirical methodology to decompose analysts' *overall* forecast revisions into components related to their reaction to public signals, private signals, and forecast corrections. This decomposition method allows for an empirical examination of heterogeneous reactions to different sources of signals when these signals are not directly observable in the data.

Second, I re-examine the predictability of forecast errors by forecast revisions in out-of-sample (OOS) tests. The evidence suggests that the weak OOS performance of the BGMS model, initially identified by Eva and Winkler (2023), is the result of analysts' heterogeneous responses to signals rather than the absence of behavioral bias. This finding is consistent with the rejection of both the *Full Information* and the *Rational Expectations* hypothesis and helps bridge the gap between evidence from two works studying forecast errors in OOS settings. The first, Bianchi et al. (2022), finds that individuals heavily weight their personal beliefs over publicly available information. The second, Eva and Winkler (2023), observes a weak performance of the BGMS model in OOS tests.

While prior studies focus on predictions for macroeconomic variables, I examine forecasts for firm revenues. The dataset employed in this study has certain advantages over those studied in other papers. First, high-frequency data ensures proper identification of different signal types analysts receive.[5] Second, analysts' forecasts are observed across three dimensions: time, forecast horizons, and firms. This allows for empirical estimation of individuals' responses to these signal types when the signals themselves are not directly observable. Additionally, the dataset enables the study of heterogeneity across analysts due to sufficiency in the number of observations per individual.

Identification Strategy. — Any inference of behavioral bias requires a valid identification of *ex ante* rational expectations. I start by noting that if analysts receive private signals on top of public signals, there cannot exist a universal agreement on a *rational* forecast, as this is conditional on the information set an analyst observes. Observing the *ex ante* rational

[4]The *rational* consensus is defined below in the identification strategy.

[5]Throughout this study I conventionally use the term 'high frequency' to refer to data on daily forecasts. Admittedly, one might associate high frequency with changes in values that occur several times per day. Here, analysts face no restrictions in reporting their revised expectations on firm revenues versus several datasets on expectations of macroeconomic variables where the survey occurs on monthly or quarterly basis—hence the term.

beliefs of every analyst separately would ideally be desirable, but it necessitates the proper identification of their information set. To my knowledge, only one prior study pursues this route, Bianchi et al. (2022), who control for analysts' information set with a machine learning model that utilizes a rich data environment. My identification strategy does not require such data to proxy for analysts' information set.

I instead argue that departures from rational expectations can be identified by comparing forecasts with a well-defined benchmark universal across analysts. I establish this benchmark by introducing a consensus estimate that is absent of any bias, which I term the *rational* consensus. The *rational* consensus can be conceptualized as a representative analyst who only observes public signals and revises expectations rationally. To demonstrate the main results, I estimate the *rational* consensus using a Kalman Filter algorithm, and I show that the results are robust to alternative specifications, for example the cross-analyst mean forecast estimated from all forecasts announced up to 7 days before time-t.

Following this definition, interday revisions of the *rational* consensus indicate how much an analyst should optimally revise her forecast had her beliefs not been influenced by behavioral bias and/or private signals. It follows that the sensitivity of her revision to revisions from the *rational* consensus can be estimated, and consequently, her *overall* forecast revisions can be decomposed into three parts: revisions from public signals, revisions from private signals, and correction revisions. I use these three factors to predict forecast errors using data from firm-revenue forecasts reported by professional equity analysts in the US. I test the predictability of analysts' forecast errors both in-sample and out-of-sample following the rolling-forward methodology suggested by Eva and Winkler (2023). Finally, I examine whether the behavioral bias is homogeneous across analysts or if they exhibit significant heterogeneity in the degree they react to news.

Main Findings. — I demonstrate that, consistent with evidence reported by Eva and Winkler (2023) who look at analysts' predictions of macroeconomic variables, out-of-sample (OOS) tests do not suggest the presence of bias when forecast errors are regressed on forecast updates. However, I find that this outcome arises due to analysts' heterogeneous responses to signals, which tend to offset one another. When forecast updates are decomposed into three factors—revisions from public signals, revisions from private signals, and correction revisions—OOS tests reveal strong evidence of behavioral bias in analysts' reported forecasts. Specifically, analysts tend to overreact to private signals, underreact to public signals, and exhibit persistence of over-/under-confidence by not properly correcting their past reports.

While revisions from public signals typically represent the largest part of *overall* revisions, underreaction to these signals weakly explains forecast errors. This underreaction is explained by analysts' tendency to underestimate the significance of adverse public news while when receiving positive public signals their responses are rather rational and in longer horizons can even be overreactive. In contrast, revisions from private signals and correction revisions, though smaller parts of *overall* revisions, exhibit strong predictive power for forecast errors, with overreaction to private signals being the most significant. Delving deeper into responses to private signals, while analysts tend to overreact to both positive and negative news they observe privately, their overreaction is more pronounced in positive news. Finally, instances of over-optimism are more pronounced than instances of over-pessimism in equity analysts' forecasts.

The data also provides some interesting stylized facts. Forecast updates typically occur from public signals with a frequency slightly over 43% on average, while private signals explain slightly under 30% of updates and corrections slightly over 26% of them. When a signal is informative of an upward revision of revenue expectations equivalent to 1 percent, an average analyst tends to underestimate the dynamics of this signal by 3 basis points if the signal is announced publicly or overestimate it by 8 basis points if this signal is a private announcement.

Additionally, I find significant cross-analyst heterogeneity of responses when analysts are studied in different groups based on their *overall* reaction coefficients. While all groups exhibit significant overreaction to private signals, this overreaction is more pronounced as we move toward the bottom of the distribution. A different pattern emerges with public signals, wherein the bottom of the distribution overreacts, while the top underreacts to these signals, consistent with evidence from Broer and Kohlhas (2022). Finally, while all groups display persistence in bias in their reports, analysts at the bottom of the distribution achieve more accurate correction revisions. Accounting for cross-analyst heterogeneity can further improve analysts' forecasts, but it adds little to the outstanding predictive power of the decomposed model.

The paper is structured as follows. In Section 2 I introduce a benchmark model from Bordalo et al. (2020b) (BGMS) that implicitly assumes homogeneous responses to different types of signals and then describe a methodology to decompose forecast revisions which allows testing for heterogeneous responses. In Section 3 I describe the data used in this study. In Section 4 I estimate the bias identified from the decomposed model and compare with the results from the BGMS model when both models are tested in-sample (IS). In Section 5 I follow the methodology proposed by Eva and Winkler (2023) and show that the results are robust when the models are tested out-of-sample (OOS). In Section 6 I examine the cross-analyst heterogeneity of responses to signals. In Section 7 I provide a series of robustness checks and in Section 8 I conclude.

2 Methodology

In this section I describe the analytical process of decomposing forecast revisions. I first present the BGMS model of Bordalo et al. (2020b) as a reference test of the rational expectations hypothesis. I then proceed to the decomposition method by allowing analysts to exhibit cross-analyst and cross-signal heterogeneous responses to new information. This analysis considers revisions from public signals, private signals and correction revisions to provide insights on analysts' behavior. With these three factors available, I employ the decomposed model as a natural generalization of the BGMS model to test for the presence of bias and shed light on analysts' heterogeneous responses to different information sources. Figure 1 shows the timeline of a forecast announced by an individual analyst.

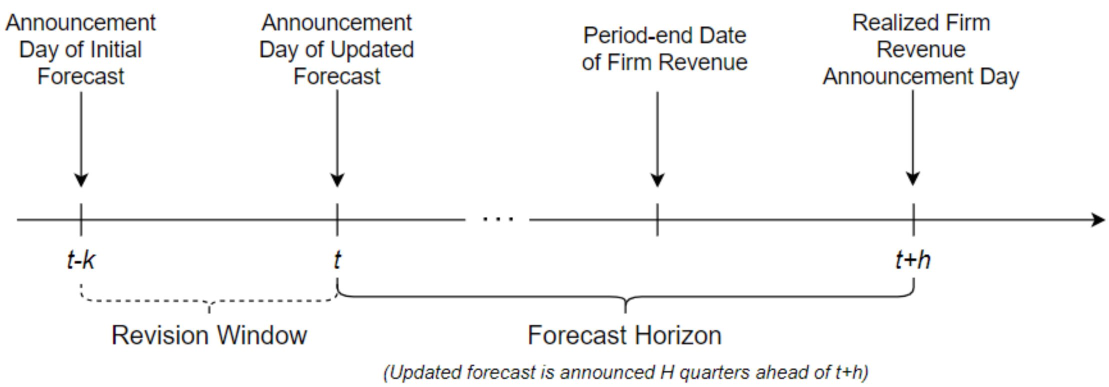

Figure 1: Timeline of a Forecast Announcement

Notes: Period-end date refers to the last day of the quarter firm revenues apply for. Capital letter H measures the forecast horizon in quarters. Small letter h measures the forecast horizon in days.

2.1 The BGMS Model

I begin by presenting the model of Bordalo et al. (2020b) (BGMS model) who run a pooled OLS regression of forecast errors on forecast revisions at the analyst level. I modify their empirical model to account for the fact that analysts in this setting provide forecasts of revenues for multiple firms at the same period, and therefore I analyze expectations at the time-analyst-firm level. Let $\tilde{x}^j_{t+h}$ denote firm j's end-of-quarter revenues (in logs) announced on day $t+h$.

Let $F^{i,j}_{t+h|t}$ be analyst i's forecast at day t (in logs) regarding firm j's revenues of the fiscal quarter announced at $t+h$. Analyst i's forecast update from day $t-k$ to t reflects either new information received within the past k day(s) or strategic behavior under which the analyst decides to report an update during that period or a combination of these. The forecast revision is given by:

$$\Delta_k F^{i,j}_{t+h} \equiv F^{i,j}_{t+h|t} - F^{i,j}_{t+h|t-k} \tag{1}$$

Unless otherwise stated, all regressions in this study include observations of which forecasts are revised within a window between 1 day and 4 weeks (inclusive). I henceforth neglect the term k in the notation to avoid confusion and denote the revision as $\Delta F^{i,j}_{t+h}$.[6] As documented by Bordalo et al. (2020b), these updates should not have explanatory power over (*ex post*) forecast errors under the assumption of rational expectations. To maintain the same interpretation of the estimated coefficients as with their work, I define the forecast errors by subtracting forecasts from the realized revenues (both expressed in logs). That is:

$$FE^{i,j}_{t+h} \equiv \tilde{x}^{j}_{t+h} - F^{i,j}_{t+h|t} \tag{2}$$

The following pooled (OLS) regression represents the baseline model (henceforth BGMS model) of this analysis on in-sample (IS) estimates and will set the ground for comparison with the decomposed model presented later in this section. The BGMS model is:

$$FE^{i,j}_{t+h} = \alpha^{\mathcal{B}} + \beta^{\mathcal{B}} \Delta F^{i,j}_{t+h} + \epsilon^{i,j}_{t+h} \quad \text{for } \ h \in H \tag{3}$$

where $j \in \{1, ..., J\}$ indexes the firm, $i \in \{1, ..., N\}$ the analyst, $t \in \{1, ..., T\}$ the day of the forecast announcement and $h \in H$ the forecast horizon (measured in days). I run separate regressions depending on the number of quarters, H, the forecast was announced ahead of the firm's realized revenue announcement. The regressions are grouped to horizons from 1 to 6 quarters ahead.[7]

As in their paper, the BGMS coefficient or coefficient of *overall* reaction to news, $\beta^{\mathcal{B}}$,

[6]There is no overlap on revisions as the most recent vintage forecast is taken into account on estimation of the revision. For example, if an analyst reported three different forecasts on revenues related to the same firm and period-end date, say on days $t = \{1, 14, 25\}$ then two revisions enter into the regressions; one from day 1 to day 14 and the second from day 14 to day 25. The forecast update from day 1 to day 25 will be omitted to mitigate concerns about residual autocorrelation.

[7]For example, when H is set to 2 quarters, the regression involves all forecasts reported as many days ahead of the firm-revenue announcement as the forecast horizon (in days) falls within a window starting 3 months beyond day t and ending 6 months beyond day t. The decision to set quarterly boundaries on horizons while forecast announcements are observed daily allows me to use a larger pool of observations as if I imposed stricter limitations on the forecast horizon (by requiring, for example, a horizon of 30 days ahead). In the latter case, only a smaller pool of analysts would meet this criterion.

tests if the *Rational Expectations* hypothesis holds. Formally,

$$H_0: \quad \beta^{\mathcal{B}} = 0, \ \text{Analysts exhibit rational reaction to news.}$$
$$H_1: \quad \beta^{\mathcal{B}} \neq 0, \ \text{Analysts exhibit non-rational reaction to news.}$$

A negative coefficient indicates that analysts overreact to news as when positive news induces them to update their expectations upward, they tend to announce forecasts beyond the level of actual revenues. Similarly, a positive coefficient is evidence of underreaction.

2.2 Decomposition of Forecast Updates

The BGMS coefficient captures the sensitivity of analysts' forecast errors to their *overall* reaction to news. However, if analysts demonstrate asymmetric responses to different signal types, forecast updates originating from one type of signal might counteract those from another type, potentially overshadowing the actual predictability of news reaction and leading to a non-informative coefficient estimate. In addition, forecast updates from $t-k$ to t as in (1) could signify corrections to inaccurate predictions submitted in the past—rather than conveying new information about fundamentals—or even strategic interactions with the consensus once measures of the latter become available. To estimate analysts' reactions to different signal sources, I start by decomposing forecast revisions according to:

$$\Delta F^{i,j}_{t+h} = \left(F^{i,j}_{t+h|t} - \bar{\mathbb{E}}^{j}_{t+h|t}\right) + \left(\bar{\mathbb{E}}^{j}_{t+h|t} - \bar{\mathbb{E}}^{j}_{t+h|t-k}\right) + \left(\bar{\mathbb{E}}^{j}_{t+h|t-k} - F^{i,j}_{t+h|t-k}\right) \tag{4}$$

I define $\bar{\mathbb{E}}^{j}_{t+h|t}$ as the *rational* consensus-the expectation the consensus would form at day t given all publicly available information in the absence of any form of bias. Specifically, the *rational* consensus does not suffer from aggregate behavioral bias (e.g., prevailing optimism/pessimism), and effectively diversifies any idiosyncratic behavioral bias (e.g., individuals' sentiments) or any component arising from individuals' strategic incentives. With these properties in mind, I estimate the *rational* consensus using the Kalman Filter, the details of which are provided in Appendix A.[8]

It follows that the second term on the RHS of equation (4) denotes how much the professional analyst would react to the public signal she observes from $t-k$ to t if she formed rational expectations.[9] The first term of the RHS captures inflated (depressed) reported

[8] In the robustness tests, I demonstrate that revisions from the rational consensus fail to predict forecast errors in out-of-sample tests, evidence in rejection of sticky information.

[9] Gemmi and Valchev (2025) subtract lagged individual forecasts from lagged consensus forecasts–consistent with the third term of (4)–to test for analysts' responses to public signals. This is a noteworthy difference with my work as their public signals correspond to individuals' observation of the formation of

forecast updates that can result from private signal observation and/or behavioral factors distorting rational expectations. These latter factors are assumed to cause *over-/under-*reaction to private and public news.

The third term of the RHS of equation (4) measures deviations from the *rational* consensus on day $t-k$. While at t the analyst might not acknowledge the consensus, she may eventually have a sense of the consensus as formed at $t-k$. The predictive ability of the latter component reflects the degree to which this analyst corrects or adjusts her historical forecasts. As the *rational* consensus—by definition—does not yield predictable forecast errors, the third term, which I name the *correction* revision, can only predict future forecast errors if the analyst is persistently over-/under-confident (positive coefficient), or if she exhibits aggressive correction to her past deviation from the consensus (negative coefficient). In the language of Gemmi and Valchev (2025), the positive coefficient is consistent with analysts' incentives to strategically diversify their reports (and thus stand out from the 'herd'), while the negative coefficient is consistent with their incentives to strategically coordinate (and thus follow the 'herd').[10]

The discussion mentioned above implies that the forecast updates from $t-k$ to t can be decomposed into forecast updates due to news about fundamentals (i.e., the first two terms), to which individuals might be reacting non-rationally, and updates due to news related to the consensus's formation (third term of (4)). Let the following process describe the sum of the first two terms:

$$F^{i,j}_{t+h|t} - \bar{\mathbb{E}}_{t+h|t-k} = \Delta F^{\mathcal{U},i,j}_{t+h} + \Delta F^{\mathcal{P},i,j}_{t+h} \tag{5}$$

$$\text{where,} \quad \Delta F^{\mathcal{U},i,j}_{t+h} \equiv \left(\alpha^u_i + \alpha^u_{i,t} + \alpha^u_{i,h}\right) \; \bar{s}^j_{t+h,k} \tag{6}$$

$$\text{and,} \quad \Delta F^{\mathcal{P},i,j}_{t+h} \equiv \left(\alpha^p_i + \alpha^p_{i,t} + \alpha^p_{i,h}\right) \; s^{i,j}_{t+h,k} \tag{7}$$

The term $\Delta F^{\mathcal{U},i,j}_{t+h}$ denotes analyst i's revision due to a public signal, while $\Delta F^{\mathcal{P},i,j}_{t+h}$ is her revision due to a private signal. Equation (5) therefore states that analyst i's current forecast diverges from the *rational* consensus's past forecast either because of her reaction to public signals, or because of her response to private signals, or a combination of both. The term $\bar{s}^j_{t+h,k} \equiv \bar{\mathbb{E}}^j_{t+h|t} - \bar{\mathbb{E}}^j_{t+h|t-k}$ in (6) is defined as the public signal, or equivalently the revision an analyst would need to make if she represented the *rational* consensus. Similarly, if $\mathbb{E}^{i,j}_{t+h|t}$

average opinion. In my work, public signals refer to news shocks about fundamentals that become common knowledge to all analysts. On the other hand, I name the third term as corrections to emphasize the fact that if this term explains forecast errors, it is attributed to poor correction efforts made by analysts when they revised their initial forecasts.

[10]Another explanation of the predictive ability of the correction term—which is not the focus of this study—is related to analysts' strategic behavior to, for instance, stay overoptimistic as sell-side analysts when the companies participate in M&As.

represents analyst i's rational beliefs conditional on her information set, the term $s^{i,j}_{t+h,k} \equiv \mathbb{E}^{i,j}_{t+h|t} - \bar{\mathbb{E}}^{j}_{t+h|t}$ in (7) is the private signal.[11] The parameters inside the parentheses with an upper letter u and p are related to analyst i's reaction to public and private signals, respectively. The first parameter characterizes her tendency to *over-/under*-estimate the significance of the signal. The second parameter captures time-specific factors that induce her to inflate or depress her reported forecast. The third parameter characterizes the degree by which the proximity of the forecast period-end date might trigger her to pay closer attention to the signal or not.[12]

Notice that if this analyst tends to reveal rational responses to public and private signals, then $\alpha^u_i = \alpha^p_i = 1$ because she neither tends to overreact nor underreact to any signal source. Furthermore, if during periods of overall optimism or pessimism she maintains her rational reaction then $\alpha^u_{i,t} = \alpha^p_{i,t} = 0$ as she keeps responding to the signals as much as needed. Finally, $\alpha^u_{i,h} = \alpha^p_{i,h} = 0$ implies that the timing she receives the signal compared to the proximity of the firm-revenue announcement date does not influence her revision decision. If all three conditions are met, then the three terms of (4) are interpreted as her rational response to private signals (first term), her rational response to public signals (second term), and her correction of potential mismatch with the consensus due to false private signals she received in the past (third term).

The identification strategy of this paper is to isolate every analyst's revisions that originate from the reception of public signals from revisions due to private signals to estimate how each of them predicts forecast errors. There are two challenges to achieving this. First, a valid measurement of the *rational* expectation given the publicly available information of day t is needed to infer the public signal. Second, an estimation of the parameters of (6) for every analyst is needed to infer her forecast updates from public signals. Once these two issues are resolved, the revision from private signals can be estimated as the difference between the LHS and the first term of the RHS of (5) without the need to identify either the private signal or the parameters of (7).

Again, the first issue is resolved by estimating the *rational* consensus with Kalman Filter. The Kalman Filter algorithm takes as inputs the cross-analyst average forecast reported on day t (ADF) and estimates the (hidden) rational expectations conditional on publicly available information of day t. This process is performed for forecasts on all combinations of firm and period-end date separately. Since individual reported forecasts are likely contaminated by bias, so will the ADF, especially on days when only a few - or even one - analysts report.

[11] The private signal is not observed in the data, but as I discuss below, it is not needed when estimating revisions from private signals. Later in the results I show how one can infer these signals as well and draw their distribution.

[12] The forecast period-end date denotes the quarter and year for which the prediction is made.

Part of this bias is assumed to distort all individuals' expectations—and is therefore aggregate. The remainder is assumed to be idiosyncratic bias, but if the pool of analysts reported on the same day is small, this component might have not been diversified when computing the ADF.

Both components of bias are not observed in the data, but their variances are parameterized as following. The aggregate bias is assumed to vary consistent with the comovement of the daily growth rate of the S&P 500 volatility index (VIX) and the forecast errors made by the ADF. The idiosyncratic bias is assumed to vary consistent with the variance of individual forecasters' dispersion from the ADF and approximates zero when the number of forecasters reported on day t tends to infinity. Finally, the public signals are assumed to vary consistent with the variance of the log inter-day changes of the ADF. The aforementioned variances determine the Kalman gain which in turn updates the *rational* consensus. Once the time series of the *rational* consensus has been estimated for revenue forecasts on a specific firm and period-end date, the public signals are backed out as a first-order change of the *rational* consensus from (any) day $t-k$ to t. The Kalman Filter method is presented in detail in Appendix A.

The nature of the data used in this study allows me to overcome the second challenge as well. Here, analysts report their revenue forecasts on different firms, announcement periods, and horizons. Given that the parameters of (6) are not firm-dependent, they can be estimated with an OLS regression (as I show in Equation (8)). To ensure sufficient statistical power of the regressions, I modify the time-specific ($\alpha^u_{i,t}$) and horizon-specific ($\alpha^u_{i,h}$) parameters of (6) to $\alpha^u_{i,q}$ and $\alpha^u_{i,H}$ respectively, such that q denotes the quarter-year that the announcement day t belongs, and H measures the forecast horizon in quarters. In other words, the model allows for cross-analyst heterogeneous reaction to public signals, while it also allows for cross-period and cross-horizon (measured in quarters) variation in their coefficients of reaction, but it assumes these parameters do not change within each quarter.[13] I simultaneously estimate the parameters of (6) and revisions from private signals by running for every analyst, i, an OLS regression:

$$F^{i,j}_{t+h|t} - \hat{\mathbb{E}}^j_{t+h|t-k} = \hat{a}^u_i \ \ \hat{s}^j_{t+h,k} + \sum_q \hat{a}^u_{i,q} \left(\mathbb{I}_{t\in q} \ \ \hat{s}^j_{t+h,k}\right) + \sum_H \hat{a}^u_{i,H} \left(\mathbb{I}_{h\in H} \ \ \hat{s}^j_{t+h,k}\right) + \widehat{\Delta F}^{\mathcal{P},i,j}_{t+h} \quad (8)$$

The variables $\hat{\mathbb{E}}^j_{t+h|t-k}$ and $\hat{s}^j_{t+h,k} = \hat{\mathbb{E}}^j_{t+h|t} - \hat{\mathbb{E}}^j_{t+h|t-k}$ are denoted with hats as they are estimations following the Kalman Filter algorithm. The dummy variables $\mathbb{I}_{x\in X}$ equal to 1 if

[13] While there is a trade-off between the statistical power of the regressions and the accuracy of the estimated parameters, I expect this ad-hoc assumption does not compromise the significance of the results. The results are consistent when the frequency is set to month-varying parameters.

$x \in X$ and 0 otherwise. $\widehat{\Delta F}^{\mathcal{P},i,j}_{t+h}$ are the residuals of (8) and serve as the estimated revisions originating from private signals (see Equation (7)). The fitted values of (8) are the estimated revisions from public signals (see Equation (6)).

The novelty of the aforementioned method to infer revisions from public signals and revisions from private signals is supported by the data that provides sufficient observations per analyst to infer their individual coefficients of reaction to news, as well as the formulation of the *rational* consensus as a reference point to infer public signals. Contrary to works that proxy revisions from public signals by first-order changes of the mean forecasts, here, estimates of public signals are cleared out of aggregate behavioral bias and idiosyncratic behavioral bias. My work also distinguishes between news signals that become common knowledge and signals related to what other analysts believe—the latter captured by the third term of Equation (4).

2.3 The Decomposed Model

Having estimated the three factors of forecast updates from Section 2.2—namely revisions from public signals, revisions from private signals and correction revisions—the BGMS model of (3) can be extended to allow for cross-signal heterogeneity in analysts' reactions. The decomposed model is a pooled OLS regression of the form:

$$FE^{i,j}_{t+h} = \alpha_0 + \beta^{\mathcal{U}} \ \widehat{\Delta F}^{\mathcal{U},i,j}_{t+h} + \beta^{\mathcal{P}} \ \widehat{\Delta F}^{\mathcal{P},i,j}_{t+h} + \beta^{\mathcal{C}} \ \widehat{\Delta F}^{\mathcal{C},i,j}_{t+h} + \varepsilon^{i,j}_{t+h} \tag{9}$$

where the first two factors are the estimated revisions from public and private signals respectively and the third factor $\widehat{\Delta F}^{\mathcal{C},i,j}_{t+h} \equiv \hat{\mathbb{E}}^{j}_{t+h|t-k} - F^{i,j}_{t+h|t-k}$ is the correction revision (i.e., the third term of the RHS of (4)). By definition, the summation of all three factors is equal to analyst i's *overall* revision. That is, $\widehat{\Delta F}^{\mathcal{U},i,j}_{t+h} + \widehat{\Delta F}^{\mathcal{P},i,j}_{t+h} + \widehat{\Delta F}^{\mathcal{C},i,j}_{t+h} = \Delta F^{i,j}_{t+h}$. This identity has a great implication since if analysts exhibit a homogeneous reaction to private and public signals as well as signals regarding the consensus formation, then the coefficient estimates must equal to the BGMS coefficient; that is, $\beta^{\mathcal{U}} = \beta^{\mathcal{P}} = \beta^{\mathcal{C}} = \beta^{\mathcal{B}}$. The decomposed model contributes to the literature that studies the predictability of forecast errors from forecast revisions as it relaxes the strict assumption of a homogeneous coefficient of reaction to news (Bordalo et al., 2020b).

It is possible that, in surveys of professional analysts, forecast errors may capture overoptimism because of agency (Bordalo et al., 2023). I assume such bias would tend to inflate their forecasts overall, but not necessarily trigger their reaction to news signals. In that sense, the addition of the *correction* revision as a separate factor in this model should control for bias not attributed to behavioral factors, and estimates of $\beta^{\mathcal{U}}$ and $\beta^{\mathcal{P}}$ can arguably serve as joint tests against the *Full Information* and *Rational Expectations* hypotheses. Formally,

$$H_0 : \beta^{\mathcal{U}} = \beta^{\mathcal{P}} = 0, \qquad \text{Rational reaction to news.}$$
$$H_1 : \beta^{\mathcal{U}} = \beta^{\mathcal{P}} \neq 0, \qquad \text{Non-rational reaction, not associated with noisy information.}$$
$$H_2 : \beta^{\mathcal{U}} \neq \beta^{\mathcal{P}} \neq 0, \qquad \text{Non-rational reaction associated with noisy information.}$$

Notice that the first alternative hypothesis does not suggest the absence of noisy information or rational inattention (see Woodford (2001) and Sims (2003)). It merely suggests that (if any) evidence of irrational updating is homogeneous across signal types. The second alternative suggests that the degree of non-rational response varies conditional on the type of the signal received.

3 The Data

I obtain forecasts by professional analysts from the *Detail History* of the Institutional Brokers' Estimate System (IBES) concerning end-of-quarter revenues of publicly traded companies in the US. All forecasts are point estimates and I retain forecasts and realized revenues reported in USD currency.[14] I maintain observations with strictly positive forecasts and actual revenues, both of which are transformed into logarithms before computing forecast errors and forecast revisions.

Conventional practice in macroeconomics literature is to compute monthly or quarterly revisions in the regressions of the CG and BGMS tests based on the frequency at which analysts report their expectations of macroeconomic variables (Karnaukh and Vokata (2022) and Bordalo et al. (2020b)).[15] The mixed-frequency data used in this study reports the entire history of analysts' forecasts tracking the time a forecast was submitted or revised by an analyst. This allows me to exploit higher-frequency revisions yielding granular information on responses to public and private signals.

If an analyst reported more than once a day revenue forecasts regarding the same firm and same horizon, I maintain their most recent vintage forecast of that day. This approach to exploit high frequency does not substantially increase the number of observations in the regressions compared to keeping analysts' most recent vintage forecast per month. That is because the majority of professional analysts do not update their initial beliefs as frequently.

[14] I exclude international firms to minimize the impact of exchange rate conversion on forecast errors. To ensure consistency, I also exclude forecasts that were either submitted anonymously or submitted after the realized revenues were reported by the firm.

[15] For example, SPF data asks panelists to report their expectations in quarterly basis. Blue Chip data tracks analysts' expectations every month.

However, this approach ensures that if more than one news signals are received within the same month, they will not be omitted exposing the results to sampling bias.[16]

While most studies in the finance literature focus on expectations of earnings per share (EPS), the decision to study expectations on gross revenues is made for two reasons. First, the data does not suffer from stock split issues which would contaminate analysts' errors as the accuracy of EPS forecasts also depends on analysts' expectations of the future number of common stocks. Second, tax-related regulations do not affect forecasts of gross revenues. Additionally, while IBES provides analysts' expectations on several variables, revenue forecasts exhibit good cross-firm coverage.

I apply winsorization to the data by trimming the upper and lower 1% of observations in the distribution of the log difference between the forecast and the actual revenue. To mitigate concerns regarding measurement errors in the data, I introduce an additional requirement for every firm and period-end date. All forecasts announced within the same month must fall within 10 times the interquartile range away from their median, unless the absolute difference between the identified outlier forecast and the actual revenues is smaller than twice the absolute difference between the median forecast and the actual revenues. To the best of my knowledge, this requirement effectively isolates bad predictions from measurement errors.

All data surviving the winsorization process are used in the identification of the *rational* consensus, but I set a minimum requirement of 30 observations per individual analyst to ensure their inclusion in the forecast decomposition process.[17] Additionally, unless otherwise specified, the revision window in this study is set between 1 day to 4 weeks (i.e., 28 days) in all regressions.[18] Finally, I retain observations with a forecast horizon between 1 to 6 quarters ahead and run different regressions based on their horizon.[19]

Table O1 in the Online Appendix reports the summary statistics of the data where forecast revisions are available. The number of observations depends on the horizon considered with a minimum of $113,884$ observations on forecasts with horizon 6 quarters, and a maximum of $421,785$ observations on forecasts with horizon 1 quarter. Similarly, the number of companies varies from $5,634$ to $7,798$, and the number of analysts varies from $3,364$ to

[16]The results remain robust when I keep analysts' most recent vintage forecast every month (not reported in this paper).

[17]The vast majority of analysts meet this criterion as they can report forecasts on different firms and horizons multiple times. This minimum threshold was chosen to ensure the proper assignment of analysts to different groups when testing for heterogeneity.

[18]The results remain robust when the revision window is adjusted to include forecasts updated between 4 to 8 weeks and 8 to 12 weeks.

[19]Forecasts with a horizon of zero quarters ahead are excluded from the regressions due to insufficient data and to avoid inferring outcomes driven by insider information. Additionally, there is an insufficient number of observations on forecasts that exceed 6 quarters ahead.

4, 465. There are on average 366 observations per analyst, and an average analyst reports forecasts for about 20 companies. While on average analysts report a little over than 18 revenue forecasts for a specific firm, they only revise their initial reports about 2 times. Finally, the pool of forecasters who report their expectations on same-firm same-quarter revenues is slightly less than 4 on average.

Figure 2 provides insights into the composition of forecast revisions. The first bin of every histogram shows the frequency at which a factor represents close to 0% of overall revisions while the last bin shows the frequency at which this factor explains almost 100% of overall revisions. Notably, revisions from public signals constitute the largest component of overall revisions; the upper left histogram shows that public signals that cover at most 50% of overall revision appear in lower frequency compared to the rest two factors while public signals that cover the vast majority of overall revision appear in higher frequency. On the other hand, both histograms of revisions from private signals and correction revisions are right skewed as these factors typically cover only a smaller part of revisions. In numbers, revisions from public signals explain on average slightly over 43%, while revisions from private signals slightly under 30% and corrections slightly over 26% of overall revisions, respectively.

4 Evidence from In-Sample Tests

In this section, I present the findings from in-sample (IS) regressions. To establish a baseline, I first examine the BGMS model outlined in Equation (3). Next, I explore the decomposed model in Equation (9). This allows me to understand analysts' responses to different types of information.

4.1 In-Sample Evidence from the BGMS Model

Figure 3 presents the estimated coefficients resulting from a regression of forecast errors on *overall* forecast revisions at both the consensus level (CG model) and individual-analyst level (BGMS model; Equation (3)). The analysis employs the CG methodology on the *Summary History* and the BGMS methodology on the *Detail History* from the IBES dataset. The *Summary History* provides a consensus forecast on the third Thursday of every month by summarizing individual forecasts.[20]

The two models present some challenges for direct comparison due to differences in the revision windows. First, individuals might adjust their initial forecasts within a short time-

[20]The CG coefficients, albeit not being the primary focus of this study, were included for reference, demonstrating the consistency of the data on firm-revenue forecasts with data on expectations of macroeconomic variables explored in existing literature.

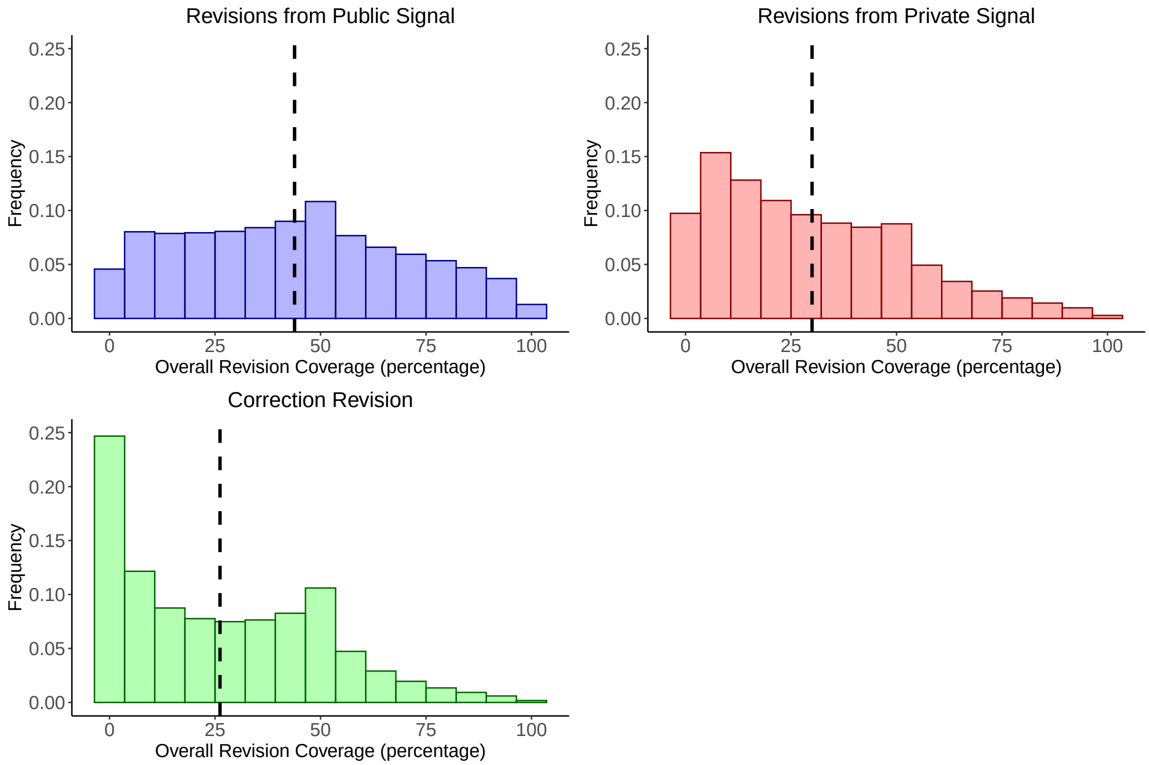

Figure 2: Composition of forecast updates

Notes: The histograms show the frequency of the percentage coverage of forecast revision by each of the three factors of the decomposed model. For every observation in the sample, the coverage ratio of any factor is calculated as the absolute value of that factor divided by the sum of absolute values of the three factors. The ratio is multiplied by 100 to denote percentage values. Taking the absolute value of the factors ensures that the sum of coverage ratio of all three factors adds up to 100. Each of the three factors from every observation is then classified to 1 of the 15 bins according to the percentage of coverage (from 0% to 100%, as shown in the horizontal axis). The vertical dashed lines point the mean coverage ratio.

frame, while the *Summary History* aggregates forecasts reported over the past month. Second, the *Summary History* is the rounding forecasts to two decimals, in contrast to the four-decimal rounding in the *Detail History*. Third, a brokerage house might opt for excluding analysts' forecasts from the *Detail History* file; a common caveat in studies that use this data.

Despite these disparities, the data exhibit consistency with findings on various macroeconomic variables in existing literature. Following the CG methodology, significant underreaction to news on forecasts with horizons ranging from 1 quarter to 1 year ahead is observed, in line with Coibion and Gorodnichenko (2015).[21] When employing the BGMS methodology, however, most horizon coefficients are not statistically significant, except for forecasts with

[21] Forecasts with horizons of 5 and 6 quarters ahead had limited observations and were excluded from the analysis.

horizons of 1 and 6 quarters ahead where the coefficients have a negative sign. The negative coefficients align with the results of Bordalo et al. (2020b), indicating overreaction to news by professional analysts at the individual level. Table O2 in the Online Appendix provides detailed coefficients.

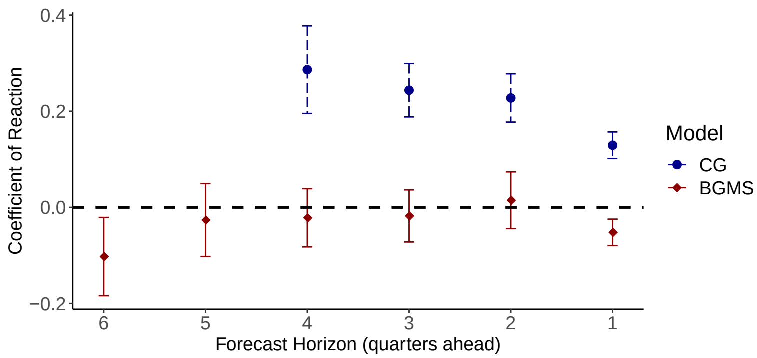

Figure 3: Estimated coefficients from the CG and BGMS models

Notes: The plot shows the estimated BGMS coefficient from a regression of forecast errors on (overall) forecast revisions as shown in equation 3 using observations from the IBES Detail File. For comparison, the estimated CG coefficient is shown from a regression of forecast errors on forecast revisions at the consensus level using observations from the IBES Summary File. Separate regressions are run for different forecast horizons (from 1 to 6 quarters ahead of firm-revenue announcement.) Standard errors are double-clustered (quarter of forecast announcement-individual analyst) in the BGMS model, and clustered at the quarter of forecast announcement in the CG model. The estimated coefficients are shown in Table O2.

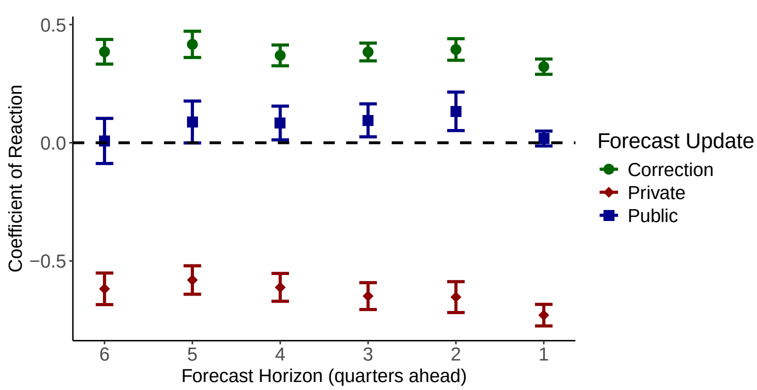

Figure 4: Estimated coefficients from the Decomposed model

Notes: The plot shows the estimated coefficients of the decomposed model from a regression of forecast errors on three factors, namely, revisions from public signals, revisions from private signals, and forecast correction (past disagreement with consensus) as shown in equation 9. Separate regressions are run for different forecast horizons (from 1 to 6 quarters ahead of firm-revenue announcement.) Double-clustered standard errors are used (quarter of forecast announcement-individual analyst). The estimated coefficients are shown in Table O3.

4.2 In-Sample Evidence from the Decomposed Model

Figure 4 illustrates the estimated coefficients derived from regressions of the decomposed model outlined in (9) across various horizons. Detailed results are presented in Table O3. The IS estimates provide evidence of a pronounced overreaction to private signals, signified by a negative and statistically significant coefficient. This implies that when analysts receive a positive private signal prompting them to announce a more optimistic forecast than the *rational* consensus, the negative coefficient predicts a negative forecast error. In other words, the forecast tends to be overly optimistic compared to the *ex post* revenues. Conversely, when analysts encounter unfavorable private news and adjust their forecasts to fall below the *rational* consensus, the forecast error tends to be positive. Both scenarios underscore the presence of overreaction to private news.

On the other hand, when statistically significant, the coefficient associated with public signals is positive, suggesting an underreaction to such signals. When analysts adjust their forecasts upwards after encountering positive news that is publicly available, the positive coefficient predicts a positive forecast error, indicating that their revision tends to be less optimistic than expected. Similarly, downward revisions resulting from negative news predict negative forecast errors, implying that analysts did not lower their initial forecasts as much as anticipated—an indication of underreaction. If analysts had adjusted their forecasts in line with the *rational* consensus upon receiving public signals, this factor would fail to predict forecast errors.

The correction coefficient is consistently positive and statistically significant for all horizons, indicating that analysts exhibit persistent optimism or pessimism. For instance, if the individual's initial forecast was more optimistic than the *rational* consensus, the positive correction coefficient predicts a negative error on the individual's updated forecast. This implies the analyst remained overconfident even after her revision as she did not appropriately correct her initial report. Similarly, if the analyst announced a less optimistic initial forecast than the *rational* consensus, the positive coefficient predicts that her updated report will remain less optimistic compared to the *ex post* firm revenues.

Following the main IS regression estimates, a crucial question arises: do analysts demonstrate a more or less pronounced underreaction to public signals depending on whether these signals are considered *good* or *bad* news? To explore this, I add an interaction term of revisions from public signals with a dummy variable, $\mathbb{I}^-$, which is equal to 1 if the signal causes a negative revision and 0 otherwise. This interaction term is added to the decomposed model as a fourth factor, indexing with $x = \{\mathcal{U}\}$ for the public signal:

$$FE^{i,j}_{t+h} = \alpha_0 + \sum_f \beta^f \ \widehat{\Delta F}^{f,i,j}_{t+h} + \beta^{-(x)} \ \mathbb{I}^-_x \ \widehat{\Delta F}^{x,i,j}_{t+h} + \varepsilon^{i,j}_{t+h} \ \text{ with } \ f \in \{\mathcal{U}, \mathcal{P}, \mathcal{C}\} \tag{10}$$

The outcomes are presented in Table O4 and indicate that analysts' reaction to public signals is not symmetric across the different scenarios (adverse vs. favorable news). Aside from forecasts with a horizon of 1 quarter ahead, where no underreaction is evident, analysts appear to underreact to negative public signals, maintaining an optimistic outlook in the face of adverse news. However, they do not underreact to favorable public signals. Indeed, the coefficient of public signals now becomes insignificant across all horizons or even turns negative for horizons of 4 and 6 quarters ahead (with significance at the 10% and 1% levels, respectively). The positive and statistically significant coefficient (at the 1% level) as initially shown in Table O3 now emerges in the interaction term. It is noteworthy that the in-sample forecasting performance of the model only shows a marginal improvement with the inclusion of the fourth factor, as the adjusted R-squared increases by less than 0.2 pp in all regressions.

I repeat the exercise of regression (10) by setting $x = \{\mathcal{P}\}$ to test if analysts' overreaction coefficient to private signals changes significantly under favorable vs. adverse private signals. Table O5 suggests that analysts exhibit a more profound overreaction to positive private news. While in all horizons the coefficient to private signals remains negative and statistically significant at the 1% level, the interaction term is positive and statistically significant at the 5% level or beyond (with the exception of horizon 1 when it is only significant at the 10% level). Importantly, the magnitude of the coefficient of the interaction term is smaller than the coefficient of revisions from private signals. When the two coefficients are summed, the results suggest that analysts also overreact to *adverse* private news, but in a less pronounced manner. Once again, the inclusion of the interaction term only marginally raises the adjusted R-Squared relative to the main model.

Notice that the interpretation of the correction revision coefficient differs from that of the coefficients associated with the first two factors. The positive coefficient suggests that analysts are persistently overconfident or underconfident by placing a relatively higher weight on their initial forecast compared to the news they subsequently receive. But is this evidence in favor of persistent optimism or pessimism? To answer this, I repeat the same exercise by extending the decomposed model with the inclusion of an interaction term by setting, this time, $x = \{\mathcal{C}\}$ in (10). The results are shown in Table O6. Except for forecasts with a horizon of 1 quarter, the positive and statistically significant interaction term reveals that instances of overoptimism are more persistent than instances of overpessimism. To see this, when the coefficient of the interaction term is added to the correction revision coefficient, the effect of a negative correction revision to the forecast error becomes stronger.

Table 1: Regression Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A					Forecast Horizon: 1 Quarter Ahead				
Public	-0.0485***				0.0980***	-0.0627***	0.0179	-0.0177**	0.0043
	(0.0165)				(0.0172)	(0.0154)	(0.0160)	(0.0072)	(0.0079)
Private		-0.9038***		-0.7344***		-0.9061***	-0.7293***	-0.7486***	-0.7243***
		(0.0324)		(0.0229)		(0.0326)	(0.0233)	(0.0194)	(0.0200)
Correction			0.5700***	0.3135***	0.6058***		0.3218***	0.3098***	0.3022***
			(0.0298)	(0.0141)	(0.0320)		(0.0165)	(0.0131)	(0.0138)
Constant	0.0056**	0.0044*	0.0044*	0.0038	0.0051**	0.0039	0.0040		
	(0.0024)	(0.0025)	(0.0024)	(0.0025)	(0.0023)	(0.0025)	(0.0025)		
Effects	No	No	No	No	No	No	No	Time	Analyst
Observations	421,785	421,785	421,785	421,785	421,785	421,785	421,785	421,785	421,785
Adjusted R^2	0.0009	0.1285	0.0791	0.1479	0.0823	0.1300	0.1481	0.1549	0.1327
Panel B					Forecast Horizon: 3 Quarters Ahead				
Public	-0.0344				0.1879***	-0.0296	0.0949***	0.0224**	0.0447***
	(0.0320)				(0.0333)	(0.0323)	(0.0356)	(0.0108)	(0.0112)
Private		-0.8755***		-0.6756***		-0.8752***	-0.6486***	-0.6907***	-0.6696***
		(0.0292)		(0.0233)		(0.0290)	(0.0290)	(0.0161)	(0.0166)
Correction			0.6072***	0.3371***	0.6788***		0.3840***	0.3529***	0.3489***
			(0.0245)	(0.0153)	(0.0235)		(0.0192)	(0.0144)	(0.0153)
Constant	−0.0167**	−0.0182***	−0.0182***	−0.0187***	−0.0177***	−0.0183***	−0.0184***		
	(0.0067)	(0.0068)	(0.0067)	(0.0068)	(0.0066)	(0.0068)	(0.0067)		
Effects	No	No	No	No	No	No	No	Time	Analyst
Observations	314,946	314,946	314,946	314,946	314,946	314,946	314,946	314,946	314,946
Adjusted R^2	0.0002	0.0616	0.0439	0.0719	0.0482	0.0617	0.0729	0.0823	0.0591

Notes: The table shows the estimated coefficients of the decomposed model from a regression of forecast errors on three factors, namely, revisions from public signals, revisions from private signals, and forecast correction (past disagreement with consensus) as shown in Equation (9). Separate regressions are run for different forecast horizons (from 1 to 6 quarters ahead of firm-revenue announcement.) Panel A reports the regression results from observations with forecast horizon of 1 quarter ahead while Panel B reports the results from observations with a horizon of 3 quarters ahead of the firm-revenue announcement date. Standard errors are double clustered in the pooled (OLS) regressions in columns (1)-(7) at the quarter of forecast announcement and (individual) analyst level. Columns (8) and (9) use time-fixed effects and analyst-fixed effects respectively and heteroscedasticity-consistent standard errors are reported. Time-fixed effects are at the month-year level. Significance levels: 10% (*), 5% (**) and 1% (***).

So far, the results suggest that the decomposed model provides superior information–when evaluated IS–that the BGMS coefficient cannot capture. This is confirmed by the sign, size and significance of the three factors and the elevated R-Squared compared to the baseline model of *overall* revisions. But do the three factors contribute equally to forecast error predictability? The answer is no. While Figure 2 shows that the largest factor of forecast updates is typically revisions from public signals, Table 1 conveys weak explanatory power of this factor on forecast errors. Indeed, column 1 of Panel A shows a small R-squared when this factor is evaluated alone on forecasts with a horizon of 1 quarter while Panel B confirms the small R-squared value on observations with a forecast horizon of 3 quarters.[22]

Contrary to public signals, private signals appear to be the primary driver of forecast errors, as evidenced by the relatively large R-squared (see column 2). The statistical significance of private signals remains robust even with the addition of the remaining two factors and the inclusion of time-fixed effects (column 8), analyst-fixed effects (column 9), and time-analyst fixed effects (not reported). It is noteworthy that while its statistical significance is retained, its economic impact increases when the correction term is excluded, possibly due to the correction term's control over past private signals received.

Similarly, the correction term exhibits a greater magnitude when private signals are excluded from the regression, and its statistical significance holds in all specifications. Notably, the in-sample predictive ability of all factors becomes stronger at shorter horizons. Lastly, the results from the time-fixed effects models (column 8) should be interpreted with caution due to the presence of the well-known Nickel bias in such data (Nickell, 1981). These results are presented to test the robustness of the main model (i.e., the pooled OLS regression of column 7).[23]

In the Online Appendix, I add a fourth factor, namely past forecast errors to control for feedback mechanisms and I show that their inclusion primarily affects analysts' reaction to public signals and not private signals. This result is of no surprise as past forecast errors are observed by analysts when firms announce their ex-post revenues (which are public signals).

5 Out-of-Sample Evaluation

An important consideration when assessing models like those presented in (3) and (9) is the potential oversight of the information set available to the analyst by the time she announced her expectations. This arises because the coefficients are estimated using the entire sample, which includes data not known by analysts at time t. Unless all analysts were aware of the distribution of these parameters, the econometrician might draw false conclusions regarding the presence of behavioral bias. To infer robust evidence of bias, a behavioral model that estimates coefficients using a training sample should be able to produce more accurate forecasts than those reported by analysts in a succeeding-period evaluation sample. This is

[22]Consistent results across various horizons are omitted from the table for clarity.

[23]To mitigate this bias, the time-fixed effects are specified at the month-year level of the forecast announcement, as opposed to the daily level, which is the frequency at which the data is observed.

because the behavioral model effectively conditions on the information observed by analysts and accounts for any bias distorting their expectations (see Eva and Winkler, 2023, Afrouzi et al., 2019, Chen et al., 2019, and Welch and Goyal, 2008).

In this section, I assess the out-of-sample (OOS) performance of the decomposed model following the methodology proposed by Eva and Winkler (2023). To ascertain the value-added of the decomposed model in the literature, I start by evaluating the OOS performance of a *behavioral* model that addresses the bias identified by the BGMS coefficient. Subsequently, I assess the OOS performance of a *behavioral* model that accounts for the bias identified by the coefficients of the three decomposed factors. While the *behavioral* model related to BGMS fails to outperform professional analysts' forecasts out-of-sample, the decomposed *behavioral* model effectively outperforms analysts, providing confirmation of the evidence shown from the IS regressions.

5.1 Out-of-Sample Evaluation of the BGMS Model

I adopt the methodology proposed by Eva and Winkler (2023) for OOS evaluation of the BGMS model. Utilizing a rolling-forward approach to estimate the BGMS coefficient, I assess the performance of a *behavioral* model that takes into account the bias identified by the BGMS coefficient for a testing period spanning from January 2005 to June 2023. Specifically, for every quarter q within the testing period, I estimate the BGMS model's coefficient β_q using all forecasts, for which revenues were announced up to 1 quarter ahead of quarter q. Subsequently, I employ the estimated coefficient to generate *ex ante* expectations for a *behavioral* model, aiming to correct for any reported bias. I then compare the performance of these expectations with a *rational* model that assumes professional analysts' reported forecasts are unbiased. Following the methodology outlined by Eva and Winkler (2023), the comparison of the two models relies on the difference in their accumulated sum squared errors (SSE).

Specifically, for every quarter q in the evaluation period from $Q1$ 2005 to $Q2$ 2023, the first stage involves estimating the $\hat{\beta}_q$ coefficient from:

$$FE^{i,j}_{t+h} = \hat{\alpha}_q + \hat{\beta}_q \ \Delta F^{i,j}_{t+h} + \epsilon^{i,j}_{t,h} \quad \text{with} \ \ t \in [Q1 \ \ 1998, q-1] \tag{11}$$

Importantly, in these regressions I drop all forecasts for which a company announced their realized revenues any time beyond quarter $q-1$ to avoid feeding the OOS predictions—formed from forecasts announced at quarter q—with excess information to which analysts could not have access. The $\hat{\beta}_q$ is then used in the second stage to infer the *ex ante* expectations formed by a *behavioral* model that aims to correct for any bias captured by the coefficient estimated in the first stage. Specifically, the *behavioral* model corrects analysts' forecasts that were announced any day of quarter q, and the *ex ante* expectations are defined as:[24]

$$\hat{x}^{i,j}_{t+h|t} \equiv F^{i,j}_{t+h|t} + \hat{\beta}_q \ \ \Delta F^{i,j}_{t+h} \quad \text{with } \ t \in q \tag{12}$$

Every quarter, q, the sum of squared (forecast) errors from the *rational* and *behavioral* models, respectively, are:

$$SSE^R_q = \sum_t \sum_h \sum_i \sum_j (\tilde{x}^j_{t+h} - F^{i,j}_{t+h|t})^2 \quad \forall t \in [Q1 \ \ 2005, q] \tag{13}$$

$$SSE^B_q = \sum_t \sum_h \sum_i \sum_j (\tilde{x}^j_{t+h} - \hat{x}^{i,j}_{t+h|t})^2 \quad \forall t \in [Q1 \ \ 2005, q] \tag{14}$$

The *behavioral* model incorporates a time-varying coefficient, $\hat{\beta}_q$, and every quarter, q, this updated coefficient is estimated using an expanding training sample that only adds information that is already known by quarter q. Finally, at any quarter q, the ability of the *behavioral* model to outperform OOS the *rational* model is tested by the following statistic:

$$\Delta SSE_q = \frac{SSE^R_q - SSE^B_q}{SSE^R_T} \quad \text{where } \ T = \{Q2 \ \ 2023\} \tag{15}$$

A rising ΔSSE_q indicates that at the current testing period adjusting for bias essentially minimizes forecast errors. The level at which ΔSSE_q stands is also important, as any level beyond zero reveals an overall outperformance of the *behavioral* model compared to the *rational* model (and vice versa), by taking into consideration all forecasts of the testing sample up to quarter q. As per Eva and Winkler (2023), the statistic is normalized by the *rational* model's final-period sum of squared errors, revealing the percentage of improvement or dis-improvement in sum squared errors over the entire evaluation period at $T = \{Q2 \ \ 2023\}$ if reports were adjusted for the observed bias.[25]

Notice that the model is parsimonious in nature as it only accounts for bias captured from analysts' responses to news and is therefore free of overabundant information that analysts could potentially not process if they encountered limitations in processing information due to cognitive factors which in turn would bias the results towards showing non-rational reaction to news (Da Silveira et al., 2020, Afrouzi et al., 2023, and Gabaix and Laibson, 2017).

[24]Consistent with Eva and Winkler (2023), I do not include the intercept from the first stage as its significance could emerge from the small size of the time series, and its inclusion would not contribute to any robust conclusion regarding any identified bias in the data.

[25]When the statistic is measured IS, a single coefficient of reaction is used at every testing period that is estimated using the entire sample and therefore not controlling for analysts' information set. The IS statistic is added for consistency, but is not the focus of this Section.

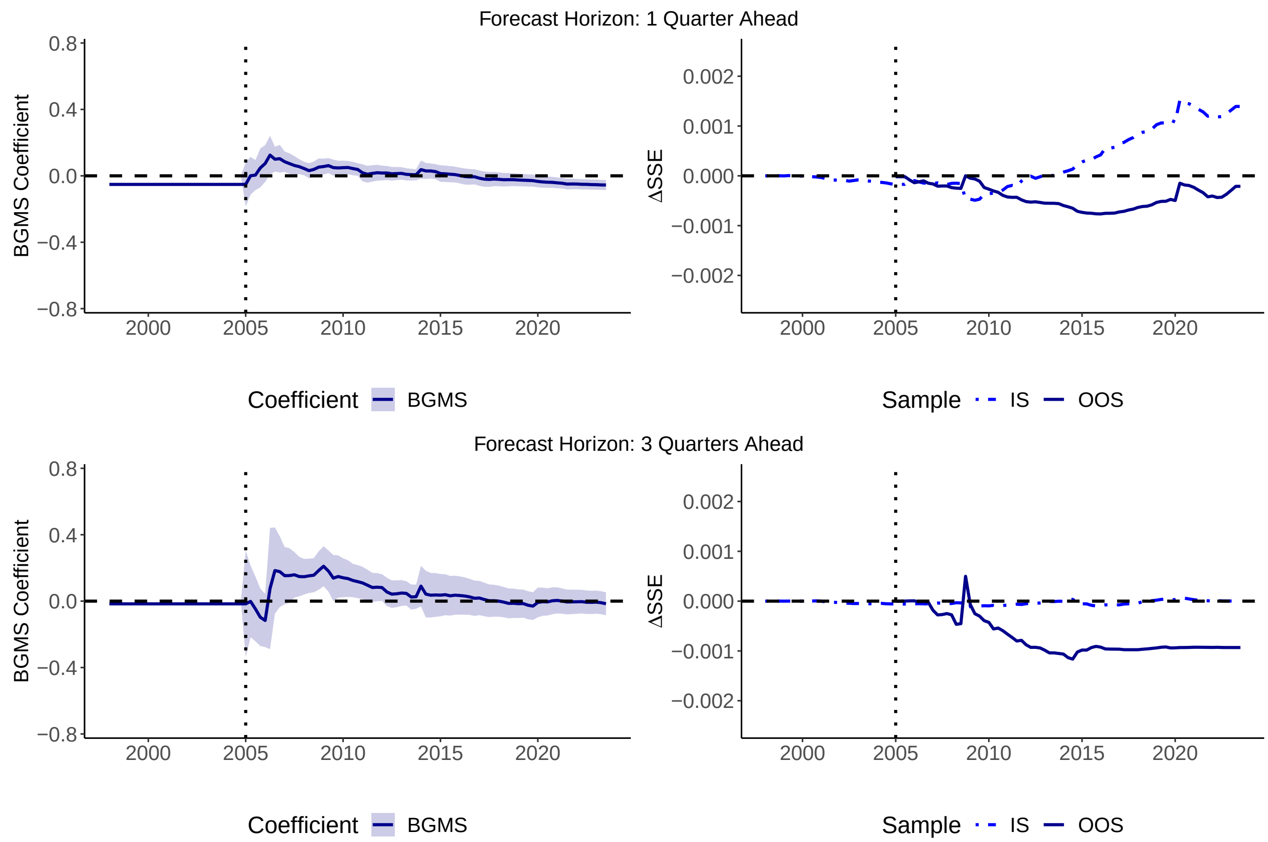

Figure 5: OOS Predictability of Forecast Errors from the BGMS Model

Notes: The plots on the LHS show the BGMS coefficient from a regression of Forecast Errors on Forecast Revisions at the individual (analyst) level using high-frequency data. Every calendar quarter q in the testing period uses a coefficient that is estimated with a rolling forward methodology (see main text) on revenues that were released up until quarter $q-1$. The shaded areas are 95% confidence intervals using double-clustered standard errors at the forecast announcement period (calendar quarter) and analyst. The plots on the RHS show the In-Sample (IS) and Out-of-Sample (OOS) performance of the ΔSSE statistic (see main text). The vertical dotted lines show the end of the training period of the rolling forward methodology and the beginning of the testing period.

Figure 5 illustrates that the *behavioral* model produced less accurate expectations compared to the *rational* model (see plots on the RHS). While the figure reports the results on forecasts with horizon 1 and 3 quarters ahead, consistent results are reported for forecasts with horizon 1-6 quarters as shown in the Online Appendix. Importantly, the ΔSSE statistic maintains a negative value for almost all quarters of the evaluation period when measured out-of-sample, and it remains negative at the endpoint of the evaluation period (except for the sample of forecasts with a horizon of 6 quarters). The BGMS coefficient is also statistically insignificant for most periods, except for quarters around the Global Financial Crisis of 2007-9 (see plots on the LHS). The results suggest that the BGMS coefficient fails to identify bias in forecasts of firm revenues when accounting for professional analysts' information sets in out-of-sample evaluations.

Next, I demonstrate that, in contrast to the BGMS coefficient, the decomposed coefficients result in a significant improvement in the formation of *ex ante* expectations implying successful identification of bias.

5.2 Out-of-Sample Evaluation of the Decomposed Model

As with the BGMS model, the robustness of any evidence from the IS evaluation of the decomposed model depends on its ability to provide smaller expectation errors than the *rational* model when evaluated OOS. To assess this, the rolling-forward methodology of (11) is adapted by replacing *overall* revisions with the three factors of the decomposed model. Specifically, for every quarter q in the evaluation period from $Q1\ \ 2005$ to $Q2\ \ 2023$, the first stage involves estimating the coefficients $\beta_q^{\mathcal{U}}$, $\beta_q^{\mathcal{P}}$ and $\beta_q^{\mathcal{C}}$ from:

$$FE_{t+h}^{i,j} = \hat{\alpha}_q + \sum_f \hat{\beta}_q^f \ \ \Delta F_{t+h}^{f,i,j} + \epsilon_{t,h}^{i,j} \quad \text{with } \ t \in [Q1\ \ 1998, q-1] \quad \text{and } \ f \in \{\mathcal{U}, \mathcal{P}, \mathcal{C}\} \tag{16}$$

Now, contrary to Equation (12), the *ex ante* expectations of the *behavioral* model are shaped by adjusting for the bias associated with any of the three factors:

$$\hat{x}_{t+h|t}^{i,j} \equiv F_{t+h|t}^{i,j} + \sum_f \hat{\beta}_q^f \ \ \Delta F_{t+h}^{f,i,j} \quad \text{with } \ t \in q \ \ \text{and } \ f \in \{\mathcal{U}, \mathcal{P}, \mathcal{C}\} \tag{17}$$

Finally, as in (15), the ΔSSE_q statistic is used to examine whether the *behavioral* model incorporating the factors of the decomposed model can outperform analysts' reported forecasts out-of-sample.

Figure 6 presents the outcomes of the OOS evaluation on forecasts with horizons of 1 and 3 quarters, with all results for horizons of 1-6 quarters are displayed in the Online Appendix. The ΔSSE_t statistic consistently rises as the sample rolls forward. Notably, in the upper right plot for forecasts with a 1-quarter horizon, the adjustment for behavioral bias linked to any of the three factors leads to a 14.86 percent reduction in the sum of squared errors by the end of the testing period. This improvement in SSE diminishes for longer forecast horizons but remains positive. For instance, correcting for bias in reports would result in a 7.12 percent reduction in SSE for forecasts with a 3-quarter horizon and a 4.05 percent improvement for forecasts with a 6-quarter horizon. Therefore, while the findings on longer horizons affirm the bias, the data indicate a more pronounced bias in shorter-horizon forecasts.

The main takeaway from the OOS analysis is that the IS estimates of the decomposed

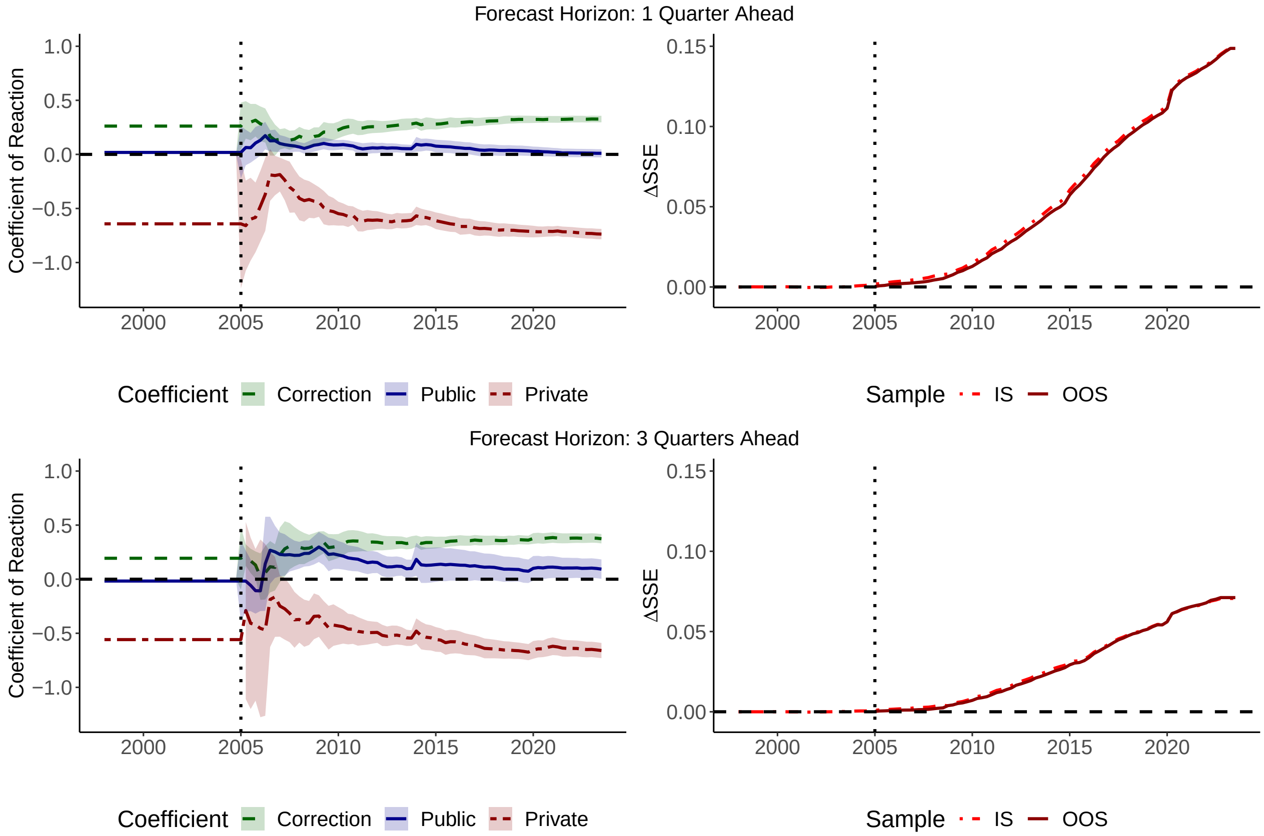

Figure 6: OOS Predictability of Forecast Errors from the Decomposed Model

Notes: The plots on the LHS show the three coefficients of the decomposed model from a regression of Forecast Errors on Forecast Revisions due to public signals, private signals, and the correction factor, at the individual (analyst) level using high-frequency data. Every calendar quarter q in the testing period uses coefficients that are estimated with a rolling forward methodology (see main text) on revenues that were released up until quarter $q-1$. The shaded areas are 95% confidence intervals using double-clustered standard errors at the announcement period (calendar quarter) and analyst. The plots on the RHS show the In-Sample (IS) and Out-of-Sample (OOS) performance of the ΔSSE statistic (see main text). The vertical dotted lines show the end of the training period of the rolling forward methodology and the beginning of the testing period.

model remain robust when accounting for professional analysts' information set and evaluating their bias OOS. In contrast, the BGMS coefficient fails to improve forecast errors OOS, reflecting the large offsetting effects of heterogeneous responses to different signal types. Notably, while *overall* revisions are predominantly associated with revisions from public signals, the smaller components—revisions from private signals and correction revisions—appear to contain superior information on forecast errors due to substantial bias. While the latter two factors exhibit a coefficient of reaction of comparable magnitude but different sign, the explanatory power of one factor over forecast errors offsets the explanatory power of the other, plummeting the power of the BGMS model while the bias is still present.

A critical implication of the out-of-sample analysis is that one can back out all analysts' rational expectations conditional on their information set from the *behavioral* model assuming that the decomposed factors effectively explain the entire predictable component of forecast errors. Equation (17) is then used as the rational expectations conditional on their information set which in turn is used to infer the private signals. That is:

$$\text{if } \hat{x}^{i,j}_{t+h|t} \approx \mathbb{E}^{i,j}_{t+h|t} \text{ then, } s^{i,j}_{t+h,k} \approx \hat{x}^{i,j}_{t+h|t} - \bar{\mathbb{E}}^{j}_{t+h|t} \quad (18)$$

In words, the private signal denotes an individual's departure from the *rational* consensus that is not explained by any bias. This departure is therefore explained by the excess information included in analyst i's information set that is excluded from the *rational* consensus's information set.

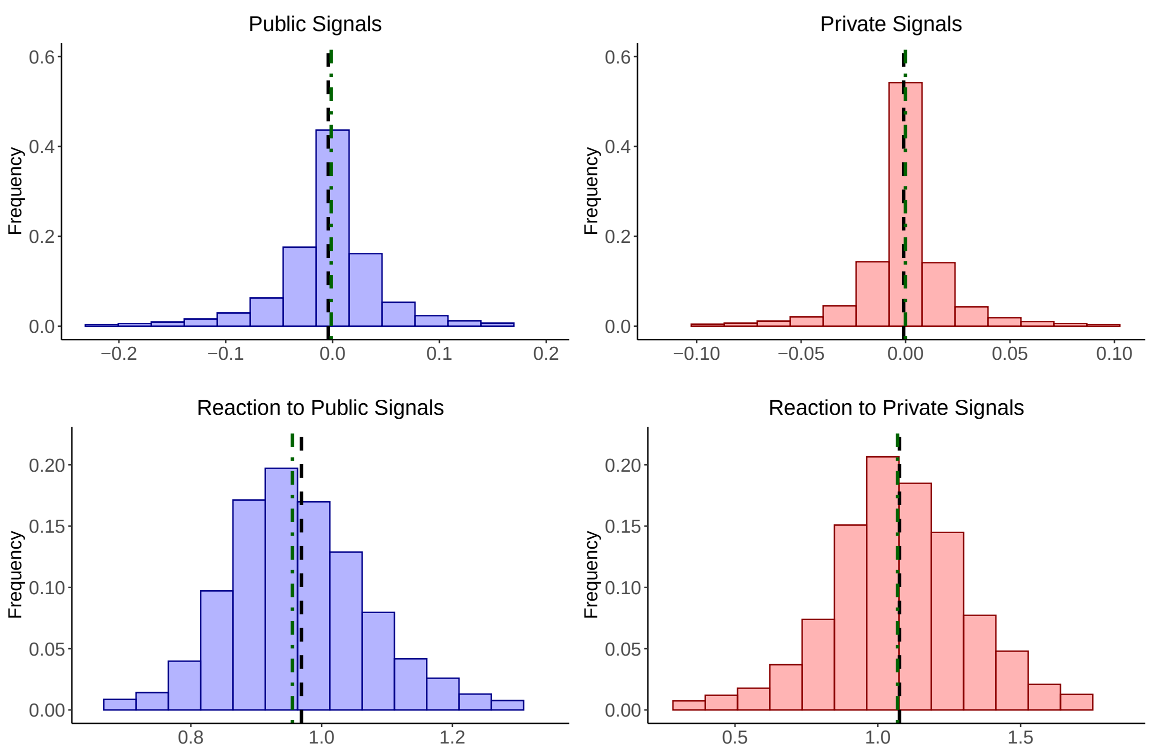

Figure 7: Distribution of Signals and Individuals' Tendency to Respond

Notes: The upper left (right) plot shows the distribution of public (private) signals for the entire testing sample from January 2005 to March 2023. Public signals are defined as the log deviations of the *rational* consensus estimates from Kalman Filter across the revision window an analyst revised. Private signals are the log deviations of the individual rational expectations (proxied by the *ex ante* expectations of the *behavioral* model) and the *rational* consensus. The lower left (right) plot shows the distribution of analysts' tendency to respond to public (private) signals. Analysts' response coefficients are estimated by an OLS regression of (19) for every analyst. All histograms are winzorized at the top and bottom 1 percent. Black dashed lines show mean values while green dot-dashed lines show median values.

Figure 7 plots the distribution of public and private signals (upper left and upper right plots respectively) as well as analysts' tendency to respond to them (lower left and lower right plots) for the entire testing sample from 2005 to 2023. Both signals are normally distributed around a mean zero value and the public signals have longer tails compared to private signals. The distribution of analysts' tendency to respond to these signals is inferred by an OLS regression on analyst i's observations:

$$F^{i,j}_{t+h|t} - \hat{\mathbb{E}}^{j}_{t+h|t-k} = \alpha_i + \phi_i \;\; \hat{s}^{j}_{t+h,k} + \psi_i \;\; \hat{s}^{i,j}_{t+h,k} + \eta^{i,j}_{t+h,k} \tag{19}$$

where ϕ_i denotes analyst i's tendency to respond to public signals and ψ_i is her tendency to respond to private signals. An individual who tends to report rational forecasts exhibits $\phi_i = \psi_i = 1$ by exploiting all available information inherent to the signals received. A coefficient greater than one implies overreaction while less than one underreaction. The lower plots of Figure 7, confirm the main results of this paper as the distribution of analysts reveals a mean coefficient of ϕ_i equal to 0.97 and a mean coefficient of ψ_i equal to 1.08. When a signal is informative of an upward revision of revenue expectations equivalent to 1 percent, an average analyst tends to underestimate the dynamics of this signal by 3 basis points or overestimate it by 8 basis points depending on whether this signal is a public announcement or private.

6 Cross-Analyst Heterogeneity

The decomposed model unveils substantial heterogeneity in analysts' responses to different types of signals. Notably, analysts tend to overestimate the significance of private signals, while frequently neglecting the relevance of public signals. Moreover, their persistent divergence from the *rational* consensus, as evidenced by past disagreements, suggests a consistent pattern of overconfidence. But do these facts describe well the entire pool of analysts, or do they exhibit asymmetric responses as we move away from the median of the distribution? To test this, all analysts are categorized into one of five groups, each representing a distinct quintile in the distribution of the BGMS coefficient. Subsequently, I examine potential significant differences in the coefficient of reaction across these different groups. Specifically, in the first stage, I run for every analyst i an OLS regression of the form:[26]

$$FE^{i,j}_{t+h} = \alpha^i + \beta^i \Delta F^{i,j}_{t+h} + \epsilon^{i,j}_{t+h} \ \forall \ h, j \tag{20}$$

All analysts are then sorted into one of the five groups that characterize the distribution of β^i. Every group represents a different quintile: 'Bottom' (bottom 20 percent), 'Low-Mid' (20-40 percentile), 'Mid' (40-60 percentile), 'High-Mid' (60-80 percentile), and 'Top' (top 20 percent). Importantly, there is no group re-arrangement as this practice would tend to group observations with significant overreaction at the bottom of the distribution and observations with significant underreaction at the top of the distribution without necessarily reflecting persistent behavioral bias from analysts' perspective. Since the allocation is decided on analysts' tendency to over-/under-react, any significant differences in responses across groups serves as evidence supporting the rejection of the null hypothesis of cross-analyst homogeneity of responses. In the BGMS setting, this is tested by an OLS regression of the form:

$$FE^{i,j}_{t+h} = \sum_g \mathbb{I}_{i \in g} \ \ \alpha_g + \sum_g \mathbb{I}_{i \in g} \ \ \beta^{\mathcal{B}}_g \ \ \Delta F^{i,j}_{t+h} + \epsilon^{i,j}_{t+h} \tag{21}$$

where g indexes the five analyst groups. The null hypothesis is that $\beta^{\mathcal{B}}_g = \beta^{\mathcal{B}} \ \ \forall g$. The upper left plot of Figure 8 demonstrates that analysts in the top 20% of the BGMS coefficient distribution tend to underreact to news, while those in the bottom 20% tend to overreact. The group around the median does not show predictable forecast errors from their *overall* revisions, a consistent observation across all horizons.

Table O7 presents the coefficient estimates along with two statistics for testing cross-group heterogeneity. A Wald one-sided test, under the null hypothesis that the Top group does not have a coefficient greater than the Bottom group, is rejected at the 1% level across all horizons. Similarly, a Wald one-sided test, under the null hypothesis that the High-Mid group does not exhibit a greater coefficient than the Low-Mid group, is rejected at the 1% level for all horizons as well. These results indicate not only significant heterogeneity in analysts' responses but also suggest that this heterogeneity is not limited to the extreme groups alone.

All but the upper left plot of Figure 8 assess the cross-group coefficients of the decomposed

[26] I include forecasts of different horizons in this step. Since I set a minimum number of 30 observations per analyst and due to the large number of observations per analyst for the majority of them, the estimated coefficients should not suffer from sampling bias. I confirm the results by characterizing them according to their t-statistic with robust standard errors that are clustered on the forecast announcement period.

model. That is tested by modifying the regression model of (21) to:

$$FE^{i,j}_{t+h} = \sum_g \mathbb{I}_{i\in g} \ \alpha_g + \sum_g \mathbb{I}_{i\in g}\beta^{\mathcal{U}}_g \Delta F^{\mathcal{U},i,j}_{t+h} + \sum_g \mathbb{I}_{i\in g}\beta^{\mathcal{P}}_g \Delta F^{\mathcal{P},i,j}_{t+h} + \sum_g \mathbb{I}_{i\in g}\beta^{\mathcal{C}}_g \Delta F^{\mathcal{C},i,j}_{t+h} + \epsilon^{i,j}_{t+h} \quad (22)$$

Significantly, the upper right plot illustrates that analysts display heterogeneous responses to public signals: While the Top group tends to underreact, the Bottom group tends to overreact to public news. This pattern is not evident with private signals, as all groups tend to overreact. However, the overreaction is more pronounced as we move toward the bottom of the distribution (see lower left plot). Finally, as the lower right plot reveals, past disagreement with the consensus predicts positive errors in all five groups, indicating that analysts do not adequately correct their forecasts. Analysts persist more in reporting biased forecasts as we move toward the top of the distribution, while they tend to achieve better forecast corrections as we move toward the bottom. The cross-analyst heterogeneity of responses becomes less pronounced in forecasts with smaller horizons, as observed in all three factors. Analytical results from the IS regressions are presented in Table O8.

While the IS evaluation reveals significant cross-analyst heterogeneity, the credibility of this observation hinges on empirical support from an OOS evaluation. To rigorously assess the presence of cross-analyst heterogeneity OOS, I modify the rolling-forward methodology presented in Section 5. For every quarter q in the period $Q1$ 2005 to $Q2$ 2023, I run an OLS regression to estimate the time-varying coefficients $\beta^{\mathcal{U}}_{q,g}$, $\beta^{\mathcal{P}}_{q,g}$, and $\beta^{\mathcal{C}}_{q,g}$ for all five groups g:

$$\begin{gathered} FE^{i,j}_{t+h} = \sum_g \mathbb{I}_{i\in g}\hat{\alpha}_{q,g} + \sum_g \mathbb{I}_{i\in g}\left\{\sum_f \hat{\beta}^f_{q,g} \ \Delta F^{f,i,j}_{t+h}\right\} + \epsilon^{i,j}_{t,h} \\ \text{with } \ t \in [Q1 \ \ 1998, q-1] \ \ , \ f \in \{\mathcal{U},\mathcal{P},\mathcal{C}\} \ \text{ and } \ g \in G \end{gathered} \quad (23)$$

In contrast to the standardized ΔSSE statistic that tests if the *behavioral* model outperforms the *rational* model, I modify the statistic (as shown in Equation (27)) to test if the *behavioral* model that allows for cross-group heterogeneity outperforms the *behavioral* model that assumes homogeneous cross-group reaction coefficients. The *ex ante* expectations of the *heterogeneous behavioral* model are:

$$\dot{x}^{i,j}_{t+h|t} \equiv F^{i,j}_{t+h|t} + \sum_g \mathbb{I}_{i\in g}\left\{\sum_f \hat{\beta}^f_{q,g} \ \Delta F^{f,i,j}_{t+h}\right\} \ \text{ with } \ t\in q \ \ , \ f \in \{\mathcal{U},\mathcal{P},\mathcal{C}\} \quad (24)$$

where the coefficients $\hat{\beta}^f_{q,g}$ are estimated from the rolling-forward methodology of (23). The *ex ante* expectations of the *homogeneous behavioral* model are those estimated in (17) which

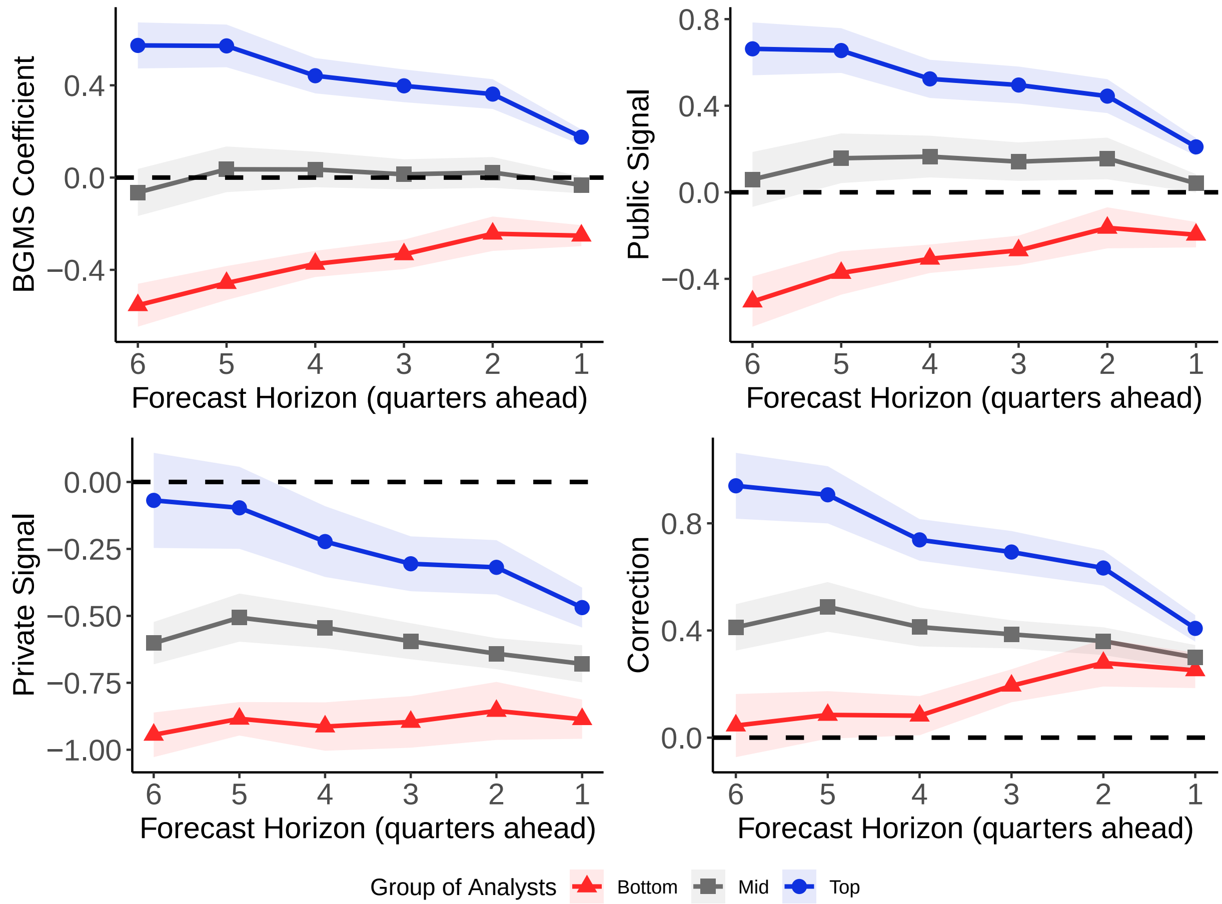

Figure 8: Cross-analyst heterogeneity of response to signals

Notes: The figure plots the coefficients from the BGMS model (upper left) and the decomposed model (remaining 3 plots) over different forecast horizons for three out of the five quintile groups of analysts. The 'Top' group includes analysts whose individual BGMS coefficient is located at the top 20 percent of the distribution. The 'Bottom' group includes analysts whose individual BGMS coefficient is located at the bottom 20 percent of the distribution. The 'Mid' group includes analysts whose individual BGMS coefficient is located at 20 percent around the median estimate. The regression for the BGMS model is shown in (21) and for the decomposed model in (22). The shaded areas are 95% confidence intervals using double-clustered standard errors at the forecast announcement period (calendar quarter) and analyst. Tables O7 and O8 in the Appendix report these coefficients for the two models respectively.

I re-write here for convenience:

$$\hat{x}^{i,j}_{t+h|t} \equiv F^{i,j}_{t+h|t} + \sum_f \hat{\beta}^f_q \;\; \Delta F^{f,i,j}_{t+h} \quad \text{with } \; t \in q \;\; \text{and} \;\; f \in \{\mathcal{U}, \mathcal{P}, \mathcal{C}\}$$

where the coefficients $\hat{\beta}^f_q$ are estimated from the rolling-forward methodology described in (16). Every quarter of the testing sample, the sum of squared errors for the two *behavioral* models is computed as:

$$SSE_q^G = \sum_t \sum_h \sum_i \sum_j (\tilde{x}_{t+h}^{i,j} - \dot{x}_{t+h|t}^{i,j})^2 \quad \forall t \in [Q1\ \ 2005, q] \tag{25}$$

$$SSE_q^B = \sum_t \sum_h \sum_i \sum_j (\tilde{x}_{t+h}^{i,j} - \hat{x}_{t+h|t}^{i,j})^2 \quad \forall t \in [Q1\ \ 2005, q] \tag{26}$$

Finally, the ΔSSE_q is modified to test for the presence of cross-analyst heterogeneity:

$$\widetilde{\Delta SSE}_q = \frac{SSE_q^B - SSE_q^G}{SSE_T^B} \quad \text{where } T = \{Q2\ \ 2023\} \tag{27}$$

Here, a rising $\widetilde{\Delta SSE}_q$ coefficient is evidence in favor of cross-analyst heterogeneity, while a diminishing coefficient implies that a universal coefficient across groups is sufficient to describe reaction to the three factors of the decomposed model. In other words, the $\widetilde{\Delta SSE}_q$ statistic now tests if the decomposed model's performance (as shown Section 5) can be further improved OOS by assuming that analysts' responses are driven by intrinsic factors (e.g., their temper and risk-aversion).

Figure 9 shows the results for forecasts with a horizon of 1 quarter ahead. Results on different horizons are shown in the Online Appendix. To ease the plots' interpretation, only the two extreme groups are presented (top and bottom quintiles). The upper left plot shows that after 2010, analysts who belong at the bottom of the distribution tend to overreact to public signals while analysts who belong to the top of the distribution tend to underreact to these signals. All groups tend to overreact to private signals (see upper right plot). However, as we move toward the bottom of the distribution analysts exhibit a more pronounced overreaction for most periods in the testing sample. Similarly, the correction coefficient is positive for all groups (see lower left plot). As we move toward the bottom of the distribution, the coefficient becomes less pronounced for most quarters.

The initial quarters of the testing period exhibit instabilities in the coefficients of all three factors. This instability is likely due to the small number of observations in the initial training period. As the training window expands, the coefficients stabilize, revealing the characteristics identified in the IS evaluation, and the cross-group differences become more apparent. Finally, the lower right plot shows that the $\widetilde{\Delta SSE}$ statistic exhibits an upward trend during the entire testing period when examined IS. When it is evaluated OOS, the statistic collapses during the initial quarters presumably due to the small number of observations (as already mentioned), but after 2010 the statistic gradually improves and becomes greater than zero by the end of the testing period.

When the analysis is repeated on observations of longer horizons, the $\widetilde{\Delta SSE}$ statistic is either negative or positive depending on the horizon examined (see the Online Appendix). For horizons of 2 and 3 quarters ahead, the statistic remains below the threshold of zero, whereas for horizons greater than 3 quarters ahead, it consistently exceeds this threshold. In all horizons, however, the statistic exhibits an upward trend after 2010, indicating that when the statistic fails to reject the null hypothesis, that might be due to insufficient data in the initial training period.

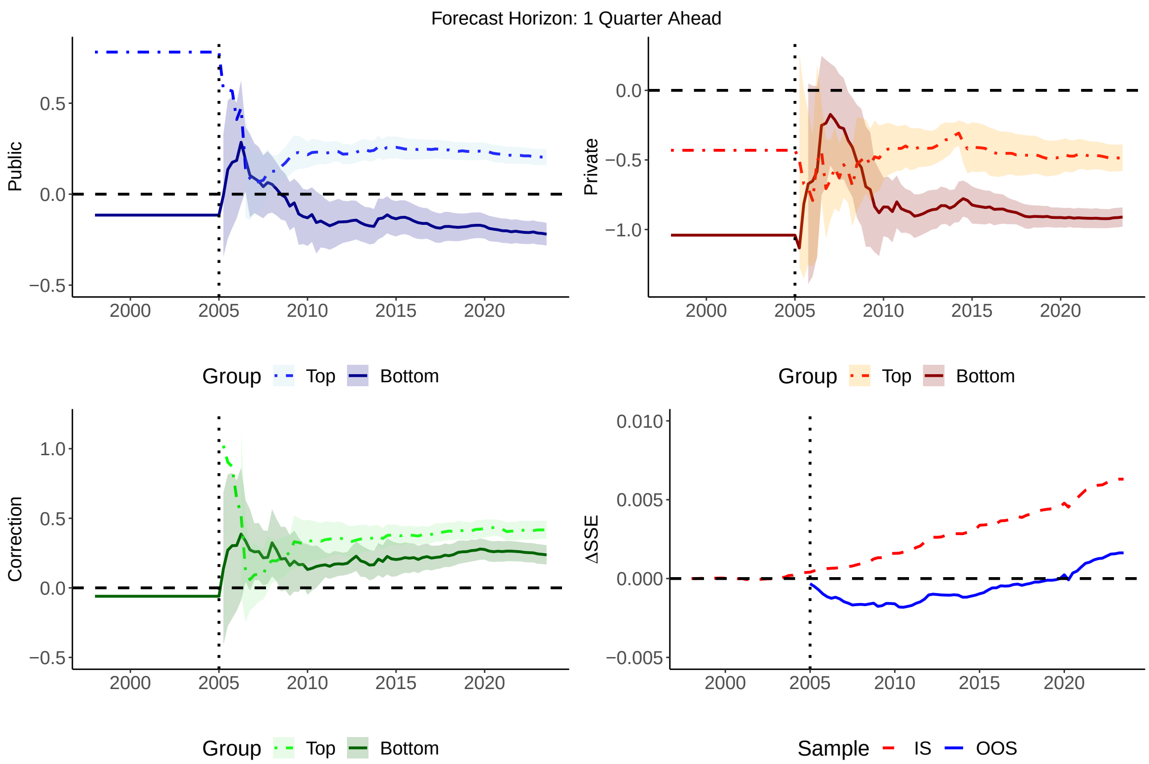

Figure 9: OOS performance of the decomposed model with cross-analyst heterogeneity

Notes: The first three plots show the time-varying estimated coefficients of the three factors of the decomposed model (public signals, private signals, correction coefficient) on two groups of analysts (Tob and Bottom quintiles). Every calendar quarter q in the testing period uses a coefficient that is estimated with a rolling forward methodology (see main text) on revenues that were released up until quarter $q-1$. The shaded areas are 95% confidence intervals using double-clustered standard errors at the announcement period (calendar quarter) and analyst. The lower right plot shows the In-Sample (IS) and Out-of-Sample (OOS) performance of a ΔSSE statistic that tests for the presence of cross-analyst heterogeneity (see main text). The vertical dotted lines show the end of the training period of the rolling forward methodology and the beginning of the testing period.

The findings collectively suggest that analysts' expectations are distorted by psychological factors (e.g., cognitive bias, overoptimism, myopia) and may be influenced by inherent personality traits (e.g., risk aversion), potentially introducing bias into their forecasts. Notably, when accounting for heterogeneous intrinsic characteristics, predictions do not exhibit significant improvement over the decomposed factors' accuracy (as shown in Section 5).

Therefore, a model that decomposes analysts' revisions into responses to different types of signals seems sufficient to identify bias in their reports while an extension of this model that allows for cross-analyst heterogeneity can only marginally explain forecast errors that are not explained by these asymmetric responses to signals.

7 Robustness Tests

I test the robustness of the results in a series of aspects and report them in the Online Appendix. First, I address the concern that the *rational* consensus might have not been properly identified. I propose two alternative modifications of the Kalman Filter process and test whether the results from the decomposition model change significantly. One alternative consists of parameterizing aggregate bias in the Kalman Filter (instead of parameterizing the variance of a measurement error related to this bias, as in the baseline model). The second alternative imposes a penalty parameter in the Kalman gain whenever the average forecast of day t in the observation equation is computed by a small pool of analysts–to account for potential distorted beliefs arising from insufficient diversification of the idiosyncratic measurement error. As an additional robustness check, I proxy for the *rational* consensus by substituting the Kalman Filter process with a simple average forecast as formed over the past 7 days. The results are robust to all three alternatives.

Second, I investigate whether public signals can predict forecast errors made by the *rational* consensus in an OOS evaluation. Detecting predictability in forecast errors would raise concerns about the accuracy of identification of the *rational* consensus (and hence, the public signals), as bias may not be effectively filtered out. However, the results reveal no predictive ability. This, together with the first series of robustness tests, indicate a proper decomposition of forecast updates.

The main results in this paper underline a combination of noisy information and behavioral factors that jointly drive forecast errors. Could these results be explained by the violation of the *Full Information* but not the violation of the *Rational Expectations* hypothesis? If these asymmetric responses to signals were driven by sticky information and not non-rational reactions, one would expect that the *rational* consensus still exhibits predictable forecast errors as the *rational* consensus has only eliminated non-rational bias. The aforementioned robustness test shows that the *rational* consensus's revisions to public signals fail to explain future forecast errors. This alleviates concerns regarding the nature of the bias identified in this study.

For the majority of forecast announcement days there has only been one analyst that reports a forecast for the same-firm same-quarter revenues. This raises concerns whether the *rational* consensus accumulates–and not eliminates–individuals' private information. To mitigate such concerns, I run regressions on observations that meet a minimum threshold on the number of forecasts reported on the same day. This way, I confirm that by restricting observations on revisions stemming from important public announcements, the observed asymmetry in responses to public and private signals is not the result of noise on individuals' reports.

Any evidence of behavioral bias in this study is inferred from forecasts that analysts revised sometime between 1 day to 4 weeks. To test whether the bias is maintained on updates that occur in lower frequency, I expand the revision window such that analysts' most recent historical forecast precedes their current forecast by more than 4 weeks; particularly, their historical forecast falls between 4-8 weeks, or between 8-12 weeks. The behavioral bias identified in the main regressions is confirmed in non-overlapping data with lower-frequency revisions.

A natural question that arises is if the behavioral bias identified in this study is driven by the majority of analysts or a minority of them whose forecast errors are sufficiently large to influence the results. Put it differently, the *behavioral* model that accounts for bias in the three factors of the decomposed model could outperform analysts' expectations OOS due to the presence of a minority of analysts whose sum of squared errors is sufficiently large while the rest of them announce accurate forecasts. I address this concern by proposing an OOS evaluation where the *behavioral* model competes with every analyst separately. The *behavioral* model outperforms the vast majority of analysts every period in the testing sample.

As I show in the Online Appendix, the decomposed model consistently improves the forecasts of the majority of analysts across all testing periods and horizons. Specifically, the decomposed model improves forecasts for an average of 56.89% of analysts (with a standard deviation of 3.30%) over time. This evidence is robust even when the decomposition process substitutes the Kalman Filter process with a 7-day average forecast to proxy for the *rational* consensus.

A caveat is weather the bias is explained by strategic incentives given that the forecasts are prepared by professional analysts (Gemmi and Valchev, 2025, Ottaviani and Sørensen, 2006, Ehrbeck and Waldmann, 1996). I assume that if analysts engaged in strategic activities, they would have an incentive to inflate or depress their forecasts at their initial reports. This would lead to bias in the forecast correction factor, but would not influence their response to news shocks. The fact that they overreact to both positive and negative private signals strengthens the argument that their bias has behavioral roots.[27]

[27]Bordalo et al. (2023) claim that distortions due to agency are stable and do not influence the time series

Finally, one might question that the decomposed model does not incorporate feedback mechanisms. For example, analysts might exhibit adaptive learning and overreaction to private signals could simply indicate that analysts are becoming more alert when responding to shocks after learning about past forecast errors they individually made. In the Online Appendix I show that when past forecast errors are included as a fourth factor in the regressions, the adjusted R-squared increases while the coefficients of the three factors do not change much with the exception of public signals that lose in terms of significance. That said, past forecast errors are related to public signals as firms announce their realized revenues, and therefore, the coefficient of reaction to public signals is expected to change with the inclusion of this variable.

8 Conclusion

The violation of the *Rational Expectations* hypothesis holds significant consequences for macroeconomics and finance, especially when the dynamics of behavioral bias differ from those of information frictions. A model challenging this fundamental assumption should demonstrate robustness when subjected to out-of-sample (OOS) evaluation. This raises an essential question to econometricians: '*Can you generate more accurate predictions than analysts conditional on their real-time information?*'. To explore this, I analyze data from professional analysts' forecasts of US firm revenues. The mixed frequency of analysts' revisions in this data, together with a decomposition methodology I propose, allows for identification of analysts' responses to public and private signals. This in turn allows me to study cross-signal asymmetries and cross-analyst heterogeneity in responses to news.

Following a rolling-forward methodology for OOS testing formally proposed by Eva and Winkler (2023), I show that a *behavioral* model that corrects for the bias identified from equity analysts' reactions to news fails to outperform their reported forecasts when the news is inferred from their *overall* revisions as in the work of Bordalo et al. (2020b). While this result is seemingly consistent with the *Rational Expectations* hypothesis, I show that the weak performance of the *behavioral* model masks simultaneous overreactions to private signals, underreactions to public signals and insufficient correction revisions of past reports due to overconfidence. Once cross-signal asymmetry of responses is taken into account, the *behavioral* model outperforms analysts' reported forecasts confirming bias that is robust OOS. While public signals typically represent the largest factor of analysts' revised expectations,

variation in forecasts. In my work, this need not be the case. Even if agency-related distortions are time-varying, correction revisions would arguably control for them. The reaction coefficients related to public and private signals are therefore attributed to behavioral bias.

the magnitude of private signals is also economically significant. Together, these results speak to the joint rejection of the *Full Information* and *Rational Expectations* hypotheses.

One explanation behind the rejection of FIRE might be related to strategic incentives (Gemmi and Valchev, 2025). However, if professional analysts' predictable forecast errors were solely driven by strategic incentives, one should expect strong under-correction of their past forecast reports without exhibiting asymmetric non-rational response to news. The data suggest that controlling for the correction factor, underreaction to public signals and overreaction to private signals remains statistically significant. A more plausible explanation that is consistent with my findings is that equity analysts assign a relatively higher weight on information they perceive as '*exclusive*', while they undervalue the importance of information that is publicly communicated. Due to this asymmetry of response, their *overall* forecast revisions are often appeared as weak predictors of forecast errors which explains the contradictory empirical results in the literature.

The findings unveil several questions for future research. First, is this asymmetry of response to signals observed to different agents in the economy (eg, households and firm managers)? Second, why are private signals perceived as more valuable than public signals? Third, what are the dynamics of this asymmetry? For example, McClure et al. (2024) find that US managers' inflation expectations can drive their price- and wage-setting decisions. In the same spirit, could asymmetric responses to signals lead to heterogeneous stock price distortions? In line with this question is a growing literature on the diagnostic expectations and its effects on stock prices (Bordalo et al., 2019, 2020a, 2023). I hope this work incentivizes the consideration of additional features in such models to enhance their predictive ability.

References

Afrouzi, Hassan, Laura Veldkamp et al., "Biased inflation forecasts," in "2019 Meeting Papers, no" 2019.

_ , **Spencer Y Kwon, Augustin Landier, Yueran Ma, and David Thesmar**, "Overreaction in expectations: Evidence and theory," *The Quarterly Journal of Economics*, 2023, p. qjad009.

Andolfatto, David, Scott Hendry, and Kevin Moran, "Are inflation expectations rational?," *Journal of Monetary Economics*, 2008, *55* (2), 406–422.

Angeletos, George-Marios and Zhen Huo, "Myopia and anchoring," *American Economic Review*, 2021, *111* (4), 1166–1200.

_ , _ , **and Karthik A Sastry**, "Imperfect macroeconomic expectations: Evidence and theory," *NBER Macroeconomics Annual*, 2021, *35* (1), 1–86.

Bianchi, Francesco, Cosmin L Ilut, and Hikaru Saijo, "Implications of diagnostic expectations: Theory and applications," 2021.

_ , **Sydney C Ludvigson, and Sai Ma**, "Belief distortions and macroeconomic fluctuations," *American Economic Review*, 2022, *112* (7), 2269–2315.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, "Diagnostic expectations and credit cycles," *The Journal of Finance*, 2018, *73* (1), 199–227.

_ , _ , **and** _ , "Overreaction and diagnostic expectations in macroeconomics," *Journal of Economic Perspectives*, 2022, *36* (3), 223–244.

_ , _ , **Rafael La Porta, and Andrei Shleifer**, "Diagnostic expectations and stock returns," *The Journal of Finance*, 2019, *74* (6), 2839–2874.

_ , _ , **Rafael La Porta, and Andrei Shleifer**, *Expectations of fundamentals and stock market puzzles*, National Bureau of Economic Research, 2020.

_ , _ , _ , **Matthew O'Brien, and Andrei Shleifer**, "Long term expectations and aggregate fluctuations," Technical Report, National Bureau of Economic Research 2023.

_ , _ , **Yueran Ma, and Andrei Shleifer**, "Overreaction in macroeconomic expectations," *American Economic Review*, 2020, *110* (9), 2748–82.

Broer, Tobias and Alexandre N Kohlhas, "Forecaster (mis-) behavior," *Review of Economics and Statistics*, 2022, pp. 1–45.

Chen, Hui, Winston Wei Dou, and Leonid Kogan, "Measuring "dark matter" in asset pricing models," Technical Report, National Bureau of Economic Research 2019.

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

_ , _ , **and Michael Weber**, "Monetary policy communications and their effects on household inflation expectations," *Journal of Political Economy*, 2022, *130* (6), 1537–1584.

_ , _ , **and Saten Kumar**, "How do firms form their expectations? new survey evidence," *American Economic Review*, 2018, *108* (9), 2671–2713.

Ehrbeck, Tilman and Robert Waldmann, "Why are professional forecasters biased? Agency versus behavioral explanations," *The Quarterly Journal of Economics*, 1996, *111* (1), 21–40.

Eva, Kenneth and Fabian Winkler, "A Comprehensive Empirical Evaluation of Biases in Expectation Formation," 2023.

Gabaix, Xavier and David Laibson, "Myopia and discounting," Technical Report, National bureau of economic research 2017.

Gemmi, Luca and Rosen Valchev, "Biased surveys," *Journal of Monetary Economics*, 2025, p. 103868.

Gennaioli, Nicola, Yueran Ma, and Andrei Shleifer, "Expectations and investment," *NBER Macroeconomics Annual*, 2016, *30* (1), 379–431.

Karnaukh, Nina and Petra Vokata, "Growth forecasts and news about monetary policy," *Journal of financial economics*, 2022, *146* (1), 55–70.

Kohlhas, Alexandre N and Ansgar Walther, "Asymmetric attention," *American Economic Review*, 2021, *111* (9), 2879–2925.

Kučinskas, Simas and Florian S Peters, "Measuring under-and overreaction in expectation formation," *Review of Economics and Statistics*, 2024, *106* (6), 1620–1637.

McClure, Ethan ML, Vitaliia Yaremko, Olivier Coibion, and Yuriy Gorodnichenko, "The macroeconomic expectations of US managers," *Journal of Money, Credit and Banking*, 2024.

Nickell, Stephen, "Biases in dynamic models with fixed effects," *Econometrica: Journal of the econometric society*, 1981, pp. 1417–1426.

Ottaviani, Marco and Peter Norman Sørensen, "The strategy of professional forecasting," *Journal of Financial Economics*, 2006, *81* (2), 441–466.

Rozsypal, Filip and Kathrin Schlafmann, "Overpersistence bias in individual income expectations and its aggregate implications," *American Economic Journal: Macroeconomics*, 2023, *15* (4), 331–371.

Silveira, Rava Azeredo Da, Yeji Sung, and Michael Woodford, "Optimally imprecise memory and biased forecasts," Technical Report, National Bureau of Economic Research 2020.

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

Welch, Ivo and Amit Goyal, "A comprehensive look at the empirical performance of equity premium prediction," *The Review of Financial Studies*, 2008, *21* (4), 1455–1508.

Woodford, Michael, "Imperfect Common Knowledge and the Effects of Monetary Policy," Working Paper 8673, National Bureau of Economic Research December 2001.

Cited Data:

Wharton Research Data Services. "WRDS" wrds.wharton.upenn.edu, accessed 2023-11-22.

Chicago Board Options Exchange, CBOE Volatility Index: VIX [VIXCLS], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/VIXCLS, December 19, 2023.

Appendix A Identification of *Rational* Consensus

The methodology of forecast revision decomposition as described in Section 2.2 requires identification of the *rational* consensus, $\bar{\mathbb{E}}^j_{t+h|t}$, as formed any day t when at least one analyst reported their forecast. The *rational* consensus can be thought of as a representative individual whose information set consists of public signals and her expectations are rational. A consequence of this definition is that the *rational* consensus effectively diversifies any idiosyncratic component that arises from either private information, individual sentiments, or strategic behavior and is not driven by aggregate behavioral bias. In this Appendix, I show how to estimate the *rational* consensus with the Kalman Filter. In the Online Appendix I propose two modifications of this algorithm to test if the identification of the *rational* consensus is robust.

A.1 *Rational* Consensus Inference in High-Frequency Data

The literature conventionally takes the average or median forecast of a predefined window to estimate consensus. There are two disadvantages associated with this practice. First, if the window is wide enough, say 1 month, forecasts reported close to the beginning of the period might be outdated if news shocks received afterward have changed analysts' expectation formation. To deal with this issue one can drop out all but the most recent vintage for every analyst before averaging but there is no warrant that analysts will report a new forecast when exposed to news shocks. On the other hand, if the window is narrow enough to ensure that analysts have reported their most updated beliefs the pool of analysts will be small enough, and the consensus formation will be distorted with idiosyncratic components that do not represent consensus but individual analysts' beliefs that were not effectively diversified.[28]

The second problem associated with this approach is that there might be periods of optimism or pessimism influencing the reaction coefficient for the majority of analysts. As a result, aggregate bias might be present in the consensus estimate and therefore $\bar{\mathbb{E}}^j_{t+h|t}$ might no longer represent *rational* beliefs given all publicly available information.

To demonstrate the main results of this paper, I use Kalman Filter to estimate $\bar{\mathbb{E}}^j_{t+h|t}$ and support the robustness of the results with the use of alternative measures, e.g., the use of 7-Day average forecast. Let $\mathbb{F}_t = \frac{1}{N_t}\sum_{i \in I} F^{i,j}_{q|t}$ be the cross-analyst average forecast of

[28]In addition, one could use data from the summary file of IBES to infer consensus; this file summarizes forecasts in a window of one month and is reported on the third Thursday of the month. One could use the most recent of these estimates as part of the information set of individual analyst i. In that case, the analysis is not only liable to the first concern, but also, analysts who submitted their forecasts farther from the third Thursday of a calendar month will be assigned an information set that has not incorporated recent public signals (if any).

day t on firm j's end-of-quarter q revenues as formed by the N_t individuals who reported a forecast the same day. Consequently, for every firm j and any fiscal quarter q, $\mathbb{F}_t$ describes a time series of forecasts that are observable.

Let $\mathbb{E}^R_{i,t}$ be analyst i's expectations regarding the realization of firm j's revenues in quarter q given her information set at day t and assuming her expectations are not driven by behavioral bias. Similarly, $\mathbb{E}^R_t$ is the consensus's expectations given all publicly available information by day t and the absence of any behavioral bias. While analyst i's expectations of day t are not observed, her announced forecast is observable.[29] Each individual's rational component of their expectation given their information set is characterized according to:

$$\mathbb{E}^R_{i,t} = \mathbb{E}^R_{i,t-1} + s_{i,t} + \bar{s}_t$$

where, $s_{i,t}$ is the change in the forecast due to a private signal that individual i observes at time t, and $\bar{s}_t$ is the change in the forecast due to a public signal. While ideally, one would like to estimate $\mathbb{E}^R_{i,t}$ for every individual, that is a toilsome process that demands a lot of observations per individual and will require strong assumptions. To deal with this issue, I reduce the dimensionality of this problem by estimating a series of rational expectations given the publicly available information. That is, I estimate the consensus's expectations as if the consensus exhibited a rational reaction to the news. The *state transition equation* is:

$$\mathbb{E}^R_t = \mathbb{E}^R_{t-1} + \bar{s}_t \tag{A.1}$$

Assuming that rational updates follow a white noise process, the state transition equation is a random walk. Here, $\bar{s}_t \sim N(0, \sigma^2_{\bar{s}})$. Notice that the absence of private signals, $s_{i,t}$, follows the assumption that the consensus effectively diversifies individual updates following these signals. Importantly, $\mathbb{E}^R_t$ is a hidden variable. The importance of estimating this variable using recursive methods, like the Kalman Filter, is aligned with the absence of an appropriate direct measure of estimating the consensus in high-frequency data since the signals are not observable but are inferred from forecast updates. The *observation equation* is:

$$\mathbb{F}_t = \mathbb{E}^R_t + \varepsilon_t + \upsilon_t \tag{A.2}$$

where, $\upsilon_t \equiv \frac{1}{N_t}\sum_i \upsilon_{i,t}$. The measurement errors $\epsilon_t \sim N(0, \sigma^2_{\epsilon,t})$ and $\upsilon_t \sim N(0, \frac{\sigma^2_{i,t}}{N_t})$ capture aggregate bias and an idiosyncratic dispersion from the *rational* consensus respectively. The former induces all analysts who submit their forecasts at t to depart from $\mathbb{E}^R_t$ due to prevail-

[29]For simplicity, I assume that on days when no analyst announced a (new) forecast, there is no public signal, and both the state and the state covariance are not updated in that case.

ing pessimism/optimism. The latter is a diversifiable component that captures analyst i's dispersion from the *norm*, either because of private signals that add up to her information set, or because of sentiments that are not aggregated.

I assume that whenever $N_t = 0$, then $\varepsilon_t = \upsilon_t = \bar{s}_t = 0$. Each day, t, reflects an iteration in the Kalman methodology and the estimated value of the hidden variable follows the state update equation:[30]

$$\hat{\mathbb{E}}^R_{t|t} = \hat{\mathbb{E}}^R_{t|t-1} + K_t \left(\mathbb{F}_t - \hat{\mathbb{E}}^R_{t|t-1} \right) \tag{A.3}$$

where, K_t is the Kalman gain:

$$K_t = \frac{\hat{P}_{t|t-1}}{\hat{P}_{t|t-1} + \sigma^2_{\epsilon,t} + \sigma^2_{\upsilon,t}} \tag{A.4}$$

with $\hat{P}_{t|t-1}$ being the predicted state estimate variance. The estimated state variance is updated according to:

$$\hat{P}_{t|t} = (1 - K_t)\hat{P}_{t|t-1} \tag{A.5}$$

Given that every day, t, represents an iteration of the methodology, the initialization step of the algorithm will crucially determine the efficiency of convergence to the consensus estimate, assuming that the measurement variances describe well the system. For every time series of firm and fiscal-quarter revenues, I choose the initial daily average forecast observed in the data as the initial guess. That is, I set $\hat{\mathbb{E}}^R_{0|0} \equiv \mathbb{F}_0$. Of course, the degree of confidence of the initial guess varies depending on how many individuals reported a forecast on day $t = 0$. For this reason, I parameterize the initial guess of the state variance as:

$$\hat{P}_{0|0} \equiv \left(\frac{N - N_0}{N} \right) (\tilde{x}^j_q - \mathbb{E}_{0|0})^2$$

where N denotes the total number of analysts who have ever reported (at least once) their forecast on firm j's revenues of quarter q. These revenues are denoted by $\tilde{x}^j_q$. The mean squared error, while incorporating realized revenues that are not known by time $t = 0$, serves as a good proxy of confidence of the initial guess. Notice that as the number of forecasters who form the initial guess increases and approaches the total number of forecasters, the

[30] To avoid any confusion regarding the notation used here, in this section $\hat{\mathbb{E}}^R_{t|t}$ is an updated estimate of the *rational* consensus from the Kalman Filter's t-th iteration/day regarding firm j's revenues that are announced on day $t + h$. This series is used as an estimation of $\bar{\mathbb{E}}_{t+h|t}$ as presented in the decomposition process of Section 2.2.

initial estimated state variance converges to zero reflecting augmented confidence that the initial guess represents the consensus.

In this recursive method, the presence of the measurement error ϵ_t of Equation (A.2) determines the characterization of $\mathbb{E}^R$ as a *rational* consensus. The absence of ϵ_t would suggest this method identified the consensus that may not necessarily exhibit *rational* reaction to the news. A consequence of the addition of this term is that the proxy of its variance selected will effectively determine the success of this characterization.

To proxy for the variance of aggregate bias, I work as following. I estimate the day-by-day growth rate of the volatility index (VIX) as $log(VIX_t/VIX_{t-1})$ and regress forecast errors on this growth rate:

$$log(x_q^j/\mathbb{F}_t) = \alpha_v + \beta_v log(VIX_t/VIX_{t-1}) + \varepsilon_{v,t}$$

I then proxy for the time-varying variance of aggregate bias as $\sigma^2_{\bar{\epsilon},t} \approx (\hat{\beta}_v)^2 \sigma^2_{VIX,t}$ where σ^2_{VIX} denotes the moving-variance of the daily growth rate of VIX index over the past 30 working-days (including day t). Intuitively, if the β_v coefficient is different than zero, then forecast errors are partially predicted by the growing uncertainty that was prevailing by the time analysts reported their expectations. This results in a larger variance of aggregate bias in the observation equation as β_v further departs from zero.

The variance of the diversifiable component is proxied by the variance of day-t dispersion from the mean forecast:

$$\sigma^2_{v,t} = \frac{1}{N_t^2} \sum_i \sigma^2_{i,t} \approx \frac{var(log(\mathbb{F}_{i,t}/\mathbb{F}_t))}{N_t}$$

The higher the number of analysts who reported a forecast on day t, N_t, the smaller the idiosyncratic variance. A concern with the use of this proxy is that whenever a daily mean forecast is represented by only one analyst, the variance becomes zero. To deal with this issue, whenever $N_t = 1$, $\sigma^2_{i,t}$ is proxied by $(log(\mathbb{F}_{i,t}/\mathbb{F}_{t-1}))^2$. Lastly, the variance of the public signal, $\sigma^2_{\bar{s}}$, is parameterized as the variance of $log(\mathbb{F}_t/\mathbb{F}_{t-1})$. The algorithmic procedure of Kalman Filter is given below. The updated state, $\mathbb{E}^R_{t|t}$, from the iterative process is the estimate of $\bar{\mathbb{E}}^j_{t+h|t}$ in the decomposition method.

Final Remark. — This process must be followed for every firm, and every horizon separately. The econometrician's problem is to infer the consensus's rational expectation of day t given the real-time publicly available information, for all t when at least one analyst announced a forecast. Having the entire time series of the rational consensus estimated, the public signal received from $t-k$ to t is the first-order change of the *rational* consensus estimate.

Online Appendix

Heterogeneous Responses to Signals and the Predictability of Forecast Errors

Symeon Taipliadis

This online appendix supports the main text of the paper as follows: In Section O.1 I present two modifications of the Kalman Filter to estimate the time series of the *rational* consensus's expectations for every firm and period-end date of firm revenues. In Section O.2 I present some robustness tests briefly mentioned in Section 7 of the main text. In Section O.3 I present additional tables and figures that complement the results discussed in the main text.

O.1 Alternative Proxies for the *Rational* Consensus

Here, I present two alternative methodologies to proxy for the *rational* consensus each of them representing extensions of the Kalman Filter presented in the main paper. These methodologies are used in the robustness tests to ensure that the *rational* consensus is well identified and effectively decomposes *overall* revisions into the three factors presented in the main text. Of course, a naive statistic, for example, a simple average of individual analysts' forecasts submitted over the past 7 days may also be used to proxy for the *rational* consensus. I use this average to run robustness checks and show that the Kalman Filter methodology is a valid tool to identify the public signals and the *rational* consensus.

O.1.1 Steps of the (Baseline) Kalman Filter Algorithm

Before presenting the extensions of the algorithm, I present the steps followed in the baseline algorithm to facilitate comparison:

1. Series selection:
    - Select firm j and quarter-q revenues.
    - Identify the number of analysts N ever announced a forecast for firm j's revenues of quarter q.
    - Find the number of analysts who reported a forecast on day t, $\mathbb{F}_t$.
    - Compute the daily-mean forecast of day t, $\mathbb{F}_t$, for every day t there was at least one forecast announcement.

2. Initialization:

   - Identify the initial day, $t = 0$, when at least one forecast was reported.
   - Set $\hat{\mathbb{E}}_{0|0} \equiv \mathbb{F}_0$.
   - Set $\hat{P}_{0|0} \equiv \left(\frac{N-N_0}{N}\right)(\tilde{X} - \mathbb{E}_{0|0})^2$

3. Prediction:

   - Set $\hat{\mathbb{E}}_{1|0} = \hat{\mathbb{E}}_{0|0}$.
   - Set $\hat{P}_{1|0} = \hat{P}_{0|0} + \sigma^2_{\bar{s}}$

4. Iteration Process (for $t \in (1, T)$):

   - Identify N_t and $\mathbb{F}_t$ from the data.
   - Estimate measurement variance as $\sigma^2_{\epsilon,t} + \sigma^2_{v,t}$
   - Estimate the Kalman Gain: $K_t = \frac{\hat{P}_{t|t-1}}{\hat{P}_{t|t-1}+\sigma^2_{\epsilon,t}+\sigma^2_{v,t}}$.
   - Estimate current state: $\hat{\mathbb{E}}_{t|t} = \hat{\mathbb{E}}_{t|t-1} + K_t\left(\mathbb{F}_t - \hat{\mathbb{E}}_{t|t-1}\right)$.
   - Update current state variance: $\hat{P}_{t|t} = (1 - K_t)\hat{P}_{t|t-1}$.
   - Predict next period's consensus estimate given the available information of day t: $\hat{\mathbb{E}}_{t+1,t} = \hat{\mathbb{E}}_{t,t}$.
   - Predict variance of next period's state estimate given the available information of day t: $\hat{P}_{t+1|t} = \hat{P}_{t|t} + \sigma^2_{\bar{s}}$.
   - Update day t to $t + 1$.

5. Drop observations of day $t = 0$ from the sample.

   - The initial observation is dropped as it was used to proxy for the initial guess.
   - This rule does not impose further restrictions in this study, as by definition, there is no forecast revision available associated with the initial observation.

O.1.2 Extension #1: Inclusion of a Penalty Parameter

The Kalman Filter algorithm used in this study (see Appendix A.1) updates the state equation by comparing the measurement variance with the variance of the public signal. When the measurement variance gets smaller, a larger weight is given to the new observation (daily mean), and the rational consensus estimate is updated aggresively. If the measurement variance is not proxied properly, the weight given to forecasts reported by a small number of analysts (e.g., 1 analyst) might cause misidentification of signals that are *de facto* private as public signals.

To deal with this concern, I modify the Kalman Filter algorithm by including a parameter on the idiosyncratic component of the observation equation (A.2) that aims to impose a penalty when the number of daily reported forecasts is relatively small. Specifically, the hidden idiosyncratic component $\upsilon_{i,t}$ of (A.2) is now defined as:

$$\upsilon_{i,t} = \sqrt{\frac{2N - N_t}{N_t}}\ \gamma_{i,t} \tag{O.1}$$

where, as before, N is the number of forecasters who ever reported a forecast and N_t is the number of forecasters who reported a forecast on day t. Notice that whenever N_t is relatively small, $\upsilon_{i,t}$ now becomes larger. The intervention of the penalty parameter can be better seen on variances:

$$\sigma_t = var(\upsilon_t) = \left(\frac{2N - N_t}{N_t}\right)\frac{\sigma_{i,t}}{N_t} \tag{O.2}$$

By proxying $\sigma_{i,t}$ as in the main algorithm, now the variance that results from the mean-dispersion σ_t becomes larger when N exceeds N_t. On the upper limit, when the number of reports of the day t equals the number of all analysts that form the consensus, the penalty parameter gets the value of 1 due to the increased confidence that the daily forecast well represents the consensus. The penalty parameter is expected to provide more conservative updates whenever the observed daily mean forecast is not well represented by a large body of individuals. The steps of the Kalman Filter algorithm are modified as follows:

1. Series selection:
   - Same as with the Baseline algorithm.
2. Initialization:
   - Same as with the Baseline algorithm.
3. Prediction:
   - Same as with the Baseline algorithm.
4. Iteration Process (for $t \in (1, T)$):
   - Identify N_t and $\mathbb{F}_t$ from the data.
   - Estimate measurement variance as $\sigma^2_{\mu,t} \equiv \sigma^2_{\epsilon,t} + \left(\frac{2N-N_t}{N}\right)\sigma^2_{\upsilon,t}$

- Estimate the Kalman Gain: $K_t = \frac{\hat{P}_{t|t-1}}{\hat{P}_{t|t-1}+\sigma^2_{\mu,t}}$.
- Estimate current state: $\hat{\mathbb{E}}_{t|t} = \hat{\mathbb{E}}_{t|t-1} + K_t\left(\mathbb{F}_t - \hat{\mathbb{E}}_{t|t-1}\right)$.
- Update current state variance: $\hat{P}_{t|t} = (1 - K_t)\hat{P}_{t|t-1}$.
- Predict next period's consensus estimate given the available information of day t: $\hat{\mathbb{E}}_{t+1,t} = \hat{\mathbb{E}}_{t,t}$.
- Predict variance of next period's state estimate given the available information of day t: $\hat{P}_{t+1|t} = \hat{P}_{t|t} + \sigma^2_{\bar{s}}$.
- Update day t to $t + 1$.

5. Drop observations of day $t = 0$ from the sample.

O.1.3 Extension #2: Inclusion of an Aggregate Bias Parameter

Another concern regarding the Kalman Filter algorithm used in the main results (see Appendix A) is that the observation equation restricts the parameter of the hidden variable to equal 1. A more generalized observation equation can be specified by modifying Equation (A.2) as:

$$\mathbb{F}_t = H_t\mathbb{E}^R_t + \upsilon_t \tag{O.3}$$

Notice that in this specification, the hidden measurement error, ε_t, that aims to capture aggregate behavioral bias is no longer entering the equation, and instead, a time-varying parameter, H_t, intends to capture this bias. This parameter is greater than one when the average daily reported forecast is driven by prevailing optimism and below one when it is driven by prevailing pessimism. Whenever there is no aggregate bias, any departure from the rational consensus is due to idiosyncratic errors that have not been well-diversified. To parameterize H_t I estimate the mean forecast error from all daily mean forecasts reported on (calendar) month τ regardless of the firm and horizon as:

$$MFE_\tau = \frac{1}{N_\tau}\sum_j\sum_h log(\mathbb{F}^j_{\tau+h|\tau}/\tilde{x}^j_{\tau+h}) \ \ \forall j, h$$

where N_τ is the total number of consensus daily forecasts reported on month τ. I then scale all monthly observations of MFE_τ to ensure the series is a zero-mean:

$$SMFE_\tau = MFE_\tau - MFE$$

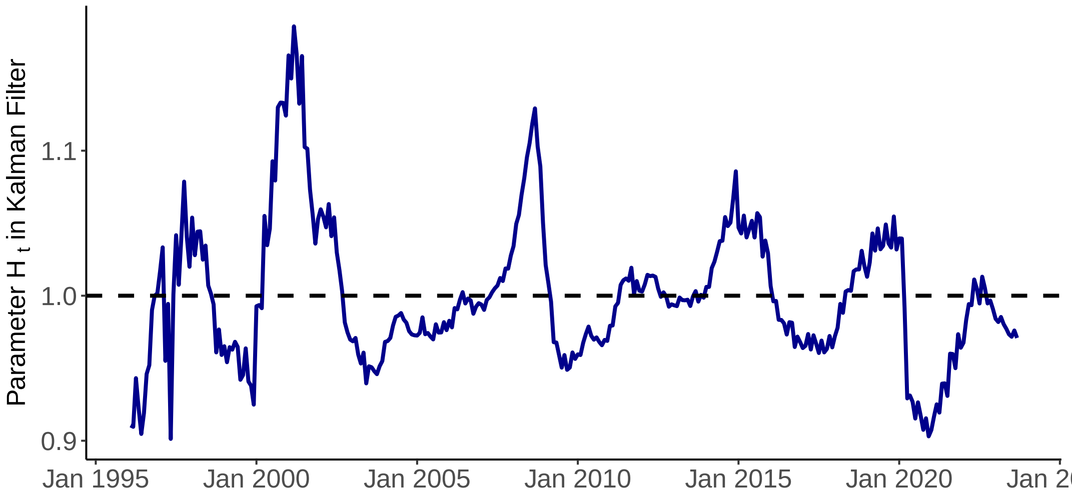

Figure O1: Proxy for aggregate bias

where MFE is the cross-month average mean forecast error. The aggregate bias is proxied as:

$$H_t = 1 + SMFE_\tau \quad \text{with,} \;\; t \in \tau \tag{O.4}$$

The intuition is that the mean forecast error within a calendar month should approximate zero due to the large volume of forecasts announced for a sufficient number of firms and forecast horizons. Since the forecast horizon is not fixed, if $SMFE_\tau$ deviates significantly from zero, that is an indication of prevailing optimism in the economy ($SMFE_\tau > 0$) or pessimism ($SMFE_\tau < 0$). We should then expect that during periods of optimism (pessimism) the parameter H_t is greater (smaller) than one as analysts tend to inflate (depress) their reports.

Figure O1 shows the fluctuation of H_t around its mean value over time. While all forecasts reported are up to 6 quarters ahead and no long-term predictions are reported, the H_t parameter seems to identify periods of optimism, for example before the emerge of the Global Financial Crisis of 2007-09, and pessimism, for example early 2020. One could argue that H_t might not be part of analysts' information set, and hence, its inclusion to the Kalman Filter might feed the algorithm with excess information. Since these forecasts are reported by professional analysts it is fair to assume that they have a fair understanding when the markets are characterized by optimism or pessimism. In any case, the inclusion of the parameterized bias in the Kalman Filter aims to test if the results from the decomposed model are robust to alternative specifications of the *rational* consensus. The steps of the algorithm are modified as follows:

1. Series selection:
   - Same as with the Baseline algorithm.
2. Initialization:
   - Same as with the Baseline algorithm.
3. Prediction:
   - Same as with the Baseline algorithm.
4. Iteration Process (for $t \in (1, T)$):
   - Identify N_t, $\mathbb{F}_t$ and H_t from the data.
   - Estimate measurement variance as $\sigma^2_{\mu,t} \equiv \sigma^2_{\epsilon,t} + \sigma^2_{v,t}$
   - Estimate the Kalman Gain: $K_t = \frac{H_t \hat{P}_{t|t-1}}{H_t^2 \hat{P}_{t|t-1} + \sigma^2_{\mu,t}}$.
   - Estimate current state: $\hat{\mathbb{E}}_{t|t} = \hat{\mathbb{E}}_{t|t-1} + K_t \left(\mathbb{F}_t - H_t \hat{\mathbb{E}}_{t|t-1} \right)$.
   - Update current state variance: $\hat{P}_{t|t} = (1 - H_t K_t) \hat{P}_{t|t-1}$.
   - Predict next period's consensus estimate given the available information of day t: $\hat{\mathbb{E}}_{t+1,t} = \hat{\mathbb{E}}_{t,t}$.
   - Predict variance of next period's state estimate given the available information of day t: $\hat{P}_{t+1|t} = \hat{P}_{t|t} + \sigma^2_{\bar{s}}$.
   - Update day t to $t + 1$.
5. Drop observations of day $t = 0$ from the sample.

O.2 Robustness Tests

O.2.1 Proxying the *Rational* Consensus

The Kalman Filter algorithm (as shown in the Appendix A.1) provides estimations of the *rational* consensus that converge to the real expectations as the number of iterations increases. Every iteration essentially represents a different forecast announcement day. One concern is that the initial iterations can give estimates that are far from the *rational* consensus, especially if the initial guess is not defined properly. Indeed, for many series, the initial guess is formed by a single forecast as during the initial day only one analyst reported their forecast over a specific firm and period-end date. The results however seem robust to alternative proxies of the *rational* consensus estimate. A straightforward alternative for $\bar{\mathbb{E}}_{t+h|t}$ without using the Kalman Filter is to calculate the mean forecast based on all forecasts announced in the past 7 days, including day t. This proxy can be expressed as:

$$\bar{\mathbb{E}}^{j}_{t+h|t} = \frac{1}{7}\sum_{\rho=0}^{6} F^{i,j}_{t+h|t-\rho}$$

Figure O2 supplements Figure 4 by decomposing forecast updates with the use of a 7-day mean forecast instead of the Kalman Filter's state equation estimates. The results are almost identical. Moreover, Figure O3 shows the OOS evaluation of the decomposed model when this alternative is used (series: 7-Day Mean). As can be seen, while the ΔSSE statistic fall slightly below the decomposed model that uses Kalman Filter, a substantial outperformance of the decomposed model is confirmed with this alternative.

Another concern is that the *rational* consensus estimates are heavily updated from forecasts that include substantial idiosyncratic measurement error. To deal with this concern, in Appendix O.1.2 I modify the Kalman Filter algorithm by adding a penalty parameter that aims to provide more conservative updates of the state equation whenever the average daily forecast is formed by a relatively small number of analysts.

The penalty parameter is added in the variance of the idiosyncratic measurement error of the observation equation. For all series (i.e., firms and period-end dates) the total number of analysts who have ever reported a forecast is tracked. On days when the number of analysts submitted a forecast equals the total number of analysts in this series, the penalty parameter disappears as there is increased confidence that during that day the average daily forecast represents the consesus's beliefs. On the other hand, when the number of analysts reported on day t falls below the total number of analysts, the penalty parameter effectively increases the idiosyncratic measurement variance and the Kalman gain shrinks. The observed average forecast of day t will therefore been given a relatively smaller weight on the state update of

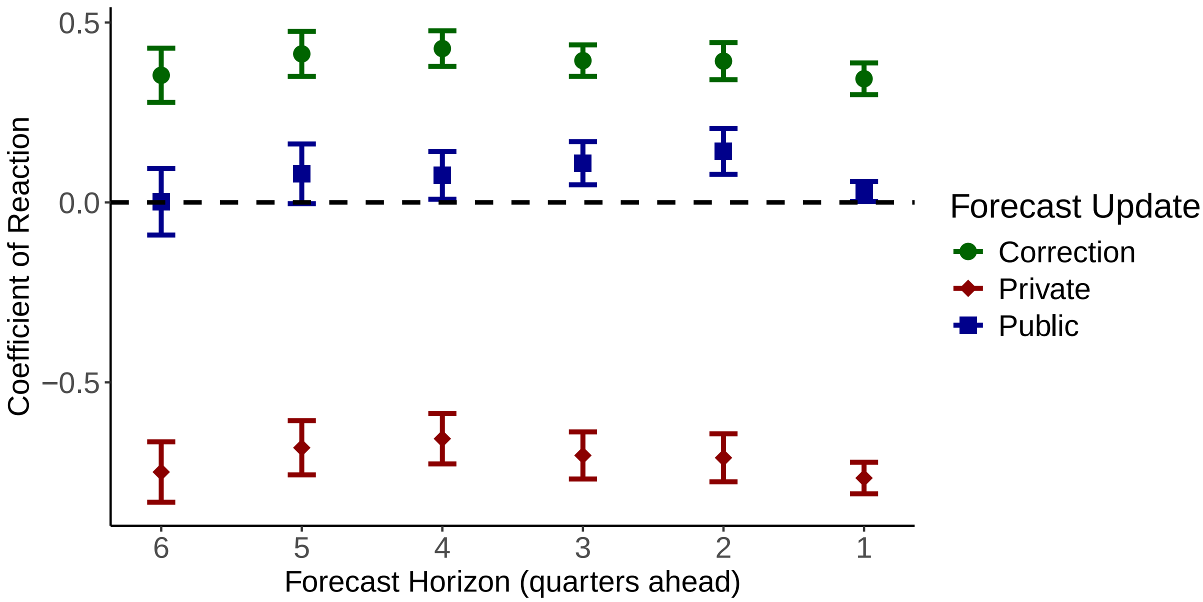

Figure O2: Estimated coefficients from the Decomposed model when the rational consensus is proxied by a 7-day mean forecast

the *rational* consensus in that case (compared to the baseline Kalman Filter algorithm of A.1). Figure O3 shows that the results from the OOS evaluation of the decomposed model remain robust when the penalty parameter is imposed to the Kalman Filter (series: Kalman (w. penalty)). The ΔSSE under this alternative consistently follows the statistic of the main results.

A third concern is that the Kalman Filter algorithm of Appendix A.1 assumes that the aggregate component of the measurement error of the observation equation is independent of the hidden variable of the *rational* consensus. The observed daily average forecast, however, could depend linearly on the *rational* consensus. In other words, there could be periods of time that all analysts are either too optimistic or too pessimistic and their aggregated reported forecasts are more or less than one times the *rational* consensus. This can be solved by allowing for a flexible time-varying parameter in the observation equation. I extend the baseline algorithm by parameterizing this aggregate bias in the observation equation. This modification is presented analytically in Appendix O.1.3. Robustness of the results are shown in Figure O3 (series: Kalman (param. bias)). There are two comments regarding this alternative proxy for the *rational* consensus. First, for all horizons, the decomposed model outperforms analysts' forecasts when the aggregate bias is parameterized in the Kalman Filter. Second, for all horizons greater than 1, the addition of the aggregate bias parameter results in significantly higher accuracy of *ex ante* expectations in the OOS evaluation.

The consistent conclusion drawn from all alternative proxies for the *rational* consensus indicates that the failure of the BGMS coefficient to identify and correct the behavioral bias in reported forecasts is not indicative of evidence supporting rational expectations. Instead, the presence of heterogeneous responses to different sources of signals obscures the predictability of forecast errors from *overall* forecast revisions.

O.2.2 Credibility of the *Rational* Consensus Estimate

Do estimates of the *rational* consensus effectively eliminate the behavioral bias? I address this question by running an OOS rolling-forward methodology at the consensus level where the sum of squared errors that arise from the *rational* consensus estimates are compared with the sum of squared errors from a *behavioral* model that aims to improve these estimates by taking into account the consensus's revision from the public signal. In essence, for every quarter $q \in [Q1\ \ 2005, Q3\ \ 2023]$ the first stage involves estimating the β_q coefficient from:

$$\widehat{FE}^{j}_{t+h} = \hat{\alpha}_q + \hat{\beta}_q\ \ \Delta\hat{E}^{j}_{t+h} + \epsilon^{i,j}_{t,h} \quad \text{with } \ t \in [Q1\ \ 1998, q-1] \tag{O.5}$$

where $\widehat{FE}^{j}_{t+h} \equiv \tilde{x}^{j}_{t+h} - \hat{E}^{j}_{t+h|t}$ and $\Delta\hat{E}^{j}_{t+h} \equiv \hat{E}^{j}_{t+h|t} - \hat{E}^{j}_{t+h|t-k}$.

Once again, I drop all forecasts for which the firm announced the realized revenues any time beyond calendar quarter q from these rolling-forward regressions to avoid feeding the OOS predictions with information that the analysts could not observe. The $\hat{\beta}_q$ is then used in the second stage to infer the *ex ante* expectations as formed by a *behavioral* consensus model that aims to correct for any estimated bias that has not been effectively eliminated in the Kalman Filter methodology. Here, the *ex ante* expectations of the *behavioral* consensus model are:

$$\hat{x}^{j}_{t+h|t} \equiv \hat{E}^{j}_{t+h|t} + \hat{\beta}_q\ \ \Delta\hat{E}^{j}_{t+h} \quad \text{with } \ t \in q \tag{O.6}$$

The sum of squared errors of the *rational* consensus and the *behavioral* consensus are:

$$\overline{SSE}^{R}_{q} \ = \ \sum_t \sum_h \sum_j (x^{j}_{t+h} - \hat{E}^{j}_{t+h|t})^2 \ \ \forall t \in [Q1\ \ 2005, q] \tag{O.7}$$

$$\overline{SSE}^{B}_{q} \ = \ \sum_t \sum_h \sum_j (x^{j}_{t+h} - \hat{x}^{j}_{t+h|t})^2 \ \ \forall t \in [Q1\ \ 2005, q] \tag{O.8}$$

At any quarter q, if there is any bias inherent in the estimations of $\hat{E}^{j}_{t+h|t}$ that should be

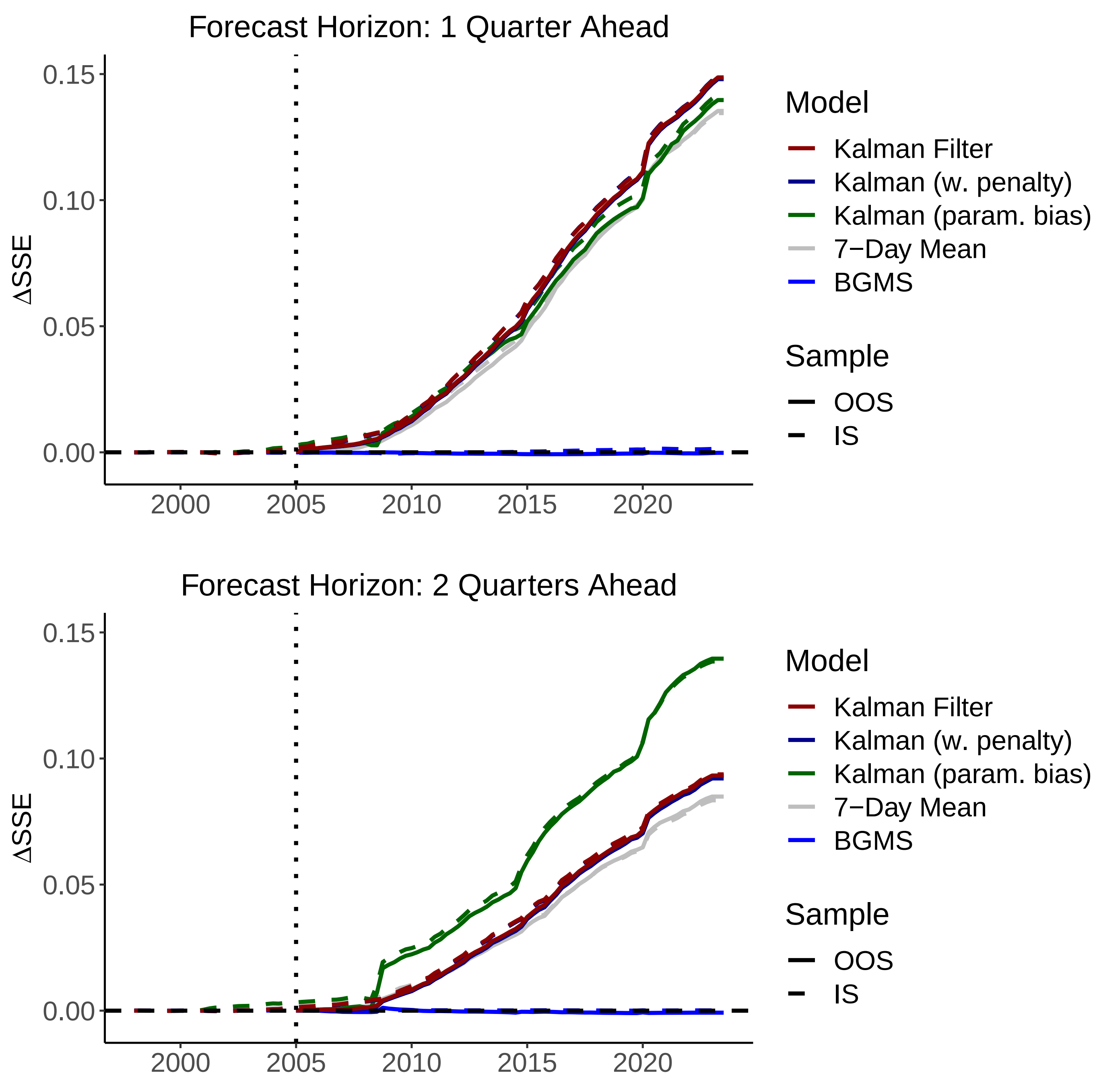

Figure O3: Robustness tests OOS

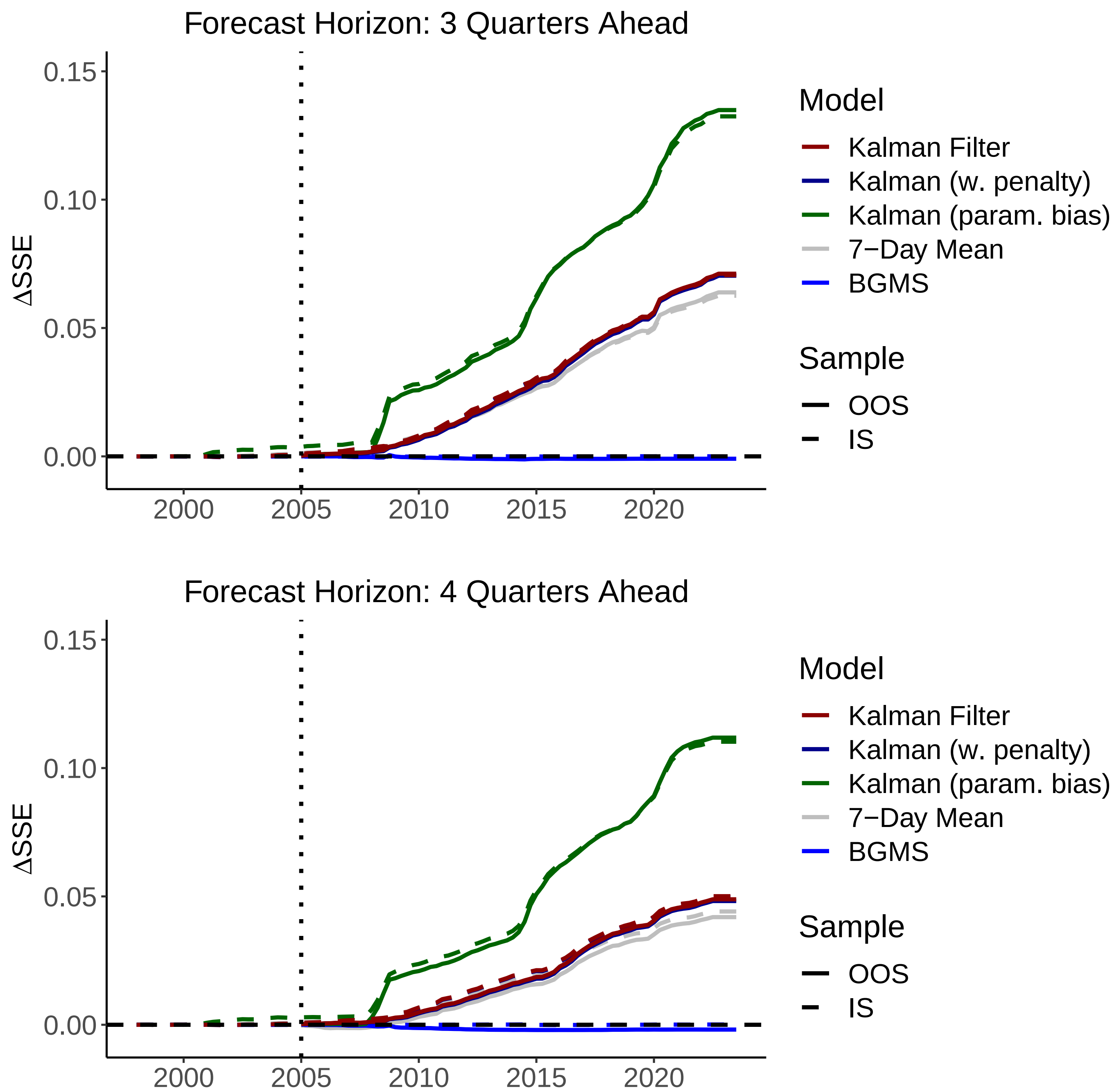

Figure O3 (cont.): Robustness tests OOS

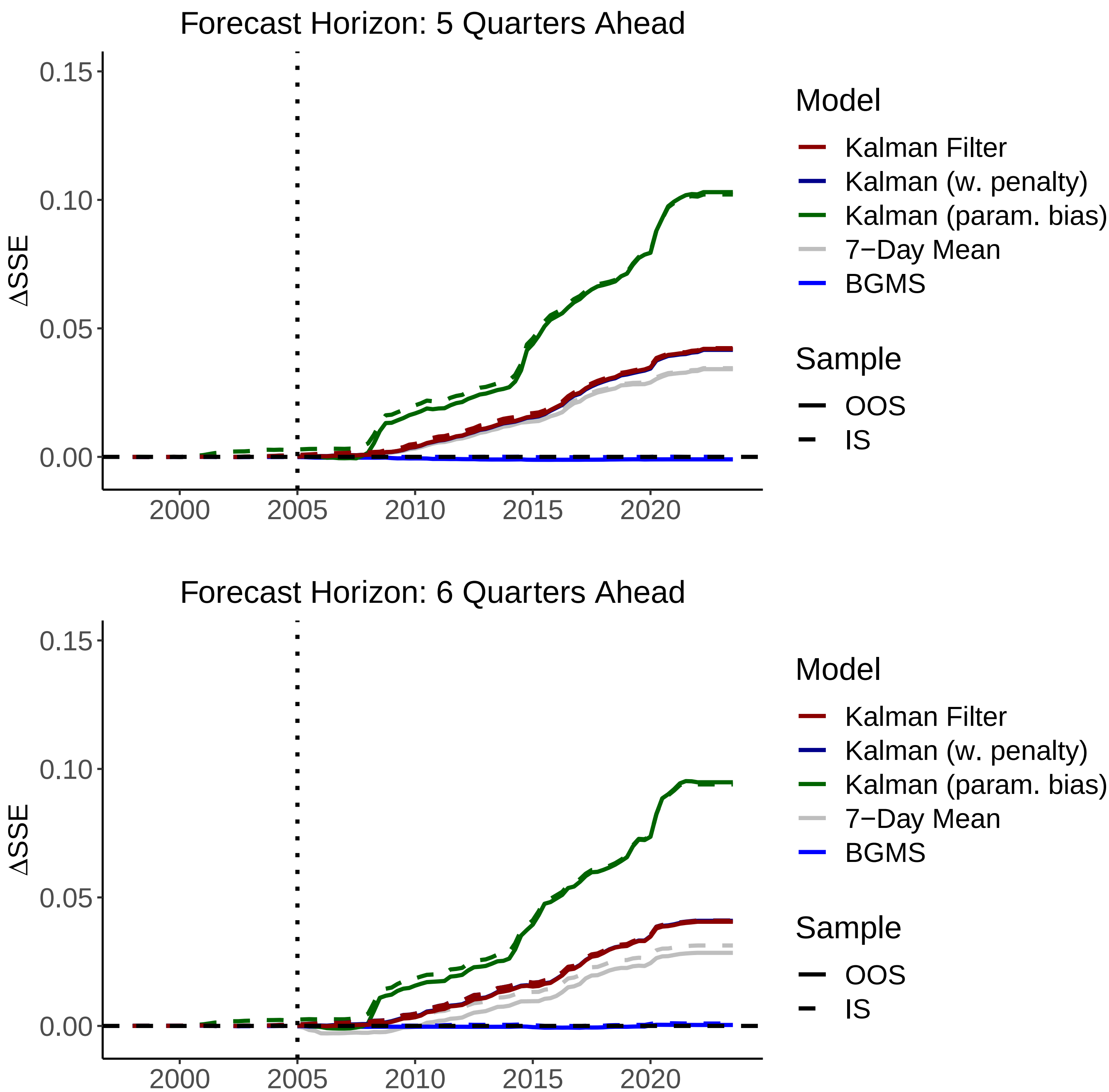

Figure O3 (cont.): Robustness tests OOS

identified by an upward trend in the following statistic:

$$\Delta SSE_q = \frac{\overline{SSE}_q^R - \overline{SSE}_q^B}{\overline{SSE}_T^R} \quad \text{where } T = \{Q2\ \ 2023\} \tag{O.9}$$

This is equivalent to saying that the model from Coibion and Gorodnichenko (2015) (CG model) is evaluated OOS when the consensus is approximated by the *rational* consensus estimate. Figure O4 shows the results from the rolling-forward methodology on forecasts with horizon of 1 to 6 quarters ahead. While in all regressions with horizon from 2 to 6 quarters ahead, the end-point of ΔSSE falls below zero or is at most located at zero, the same is not true on forecasts with horizon of 1 quarter ahead. However, in the latter case, the improvement in the sum of squared errors OOS is not economically significant as it would only improve the *rational* consensus estimates by a little less than 0.62% by the end of the testing period. Moreover, while ΔSSE slopes upward after 2012 in the OOS evaluation, the improvement in the sum of squared errors is trivial.

The key insight from these out-of-sample results is that the Kalman Filter methodology employed in this paper effectively captures rational expectations based on all publicly available information. This alleviates concerns about significant bias in the estimates, strengthening the credibility of the results.

O.2.3 Imposing a minimum threshold on the number of forecasts reported daily

The prevalence of cases where only one individual has reported their forecast on a given forecast announcement day, t, for a specific firm and quarter might raise concerns about the differentiation between private and public signals. It is possible that revisions from private signals are heavily influenced by measurement errors and not by deviations from the *rational* consensus due to private signals. To address this potential issue, I introduce a minimum threshold on the number of forecasts reported daily regarding revenues of the same firm and quarter, ensuring that these observations represent 'busy' days with important news releases.

Figures O5 and O6 demonstrate the robustness of the results by conducting regressions on forecasts announced on days with at least 2 and 4 analysts reporting their forecasts for the same firm's revenues, respectively. Despite the smaller number of observations that meet these minimum thresholds, this filter reveals a more pronounced differentiation between reactions to public and private signals. Specifically, the overreaction to private signals remains consistent with estimates from the main results, while a more pronounced underreaction to public signals is observed compared to the main findings.

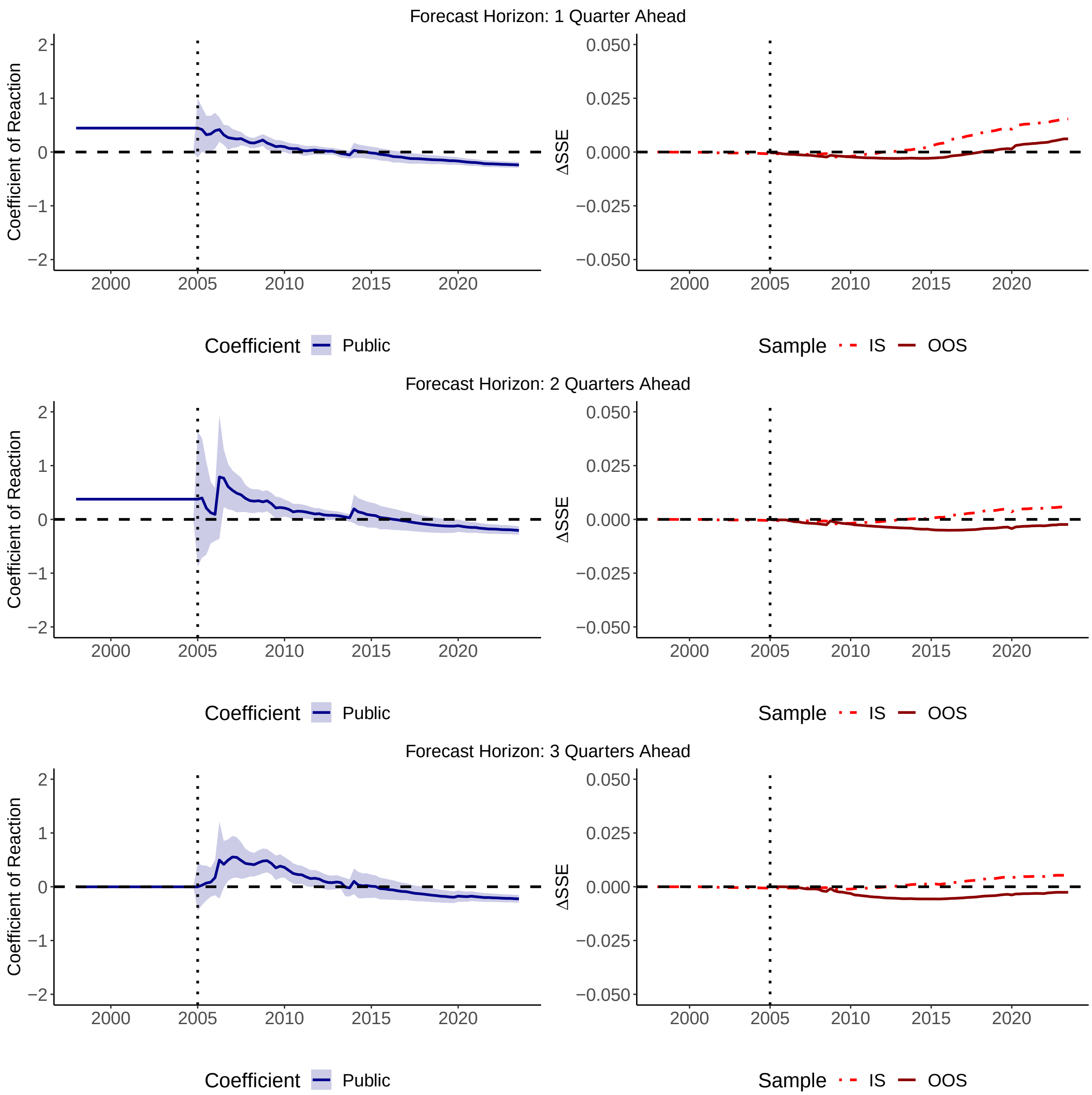

Figure O4: OOS Credibility Test of the Rational Consensus Estimate

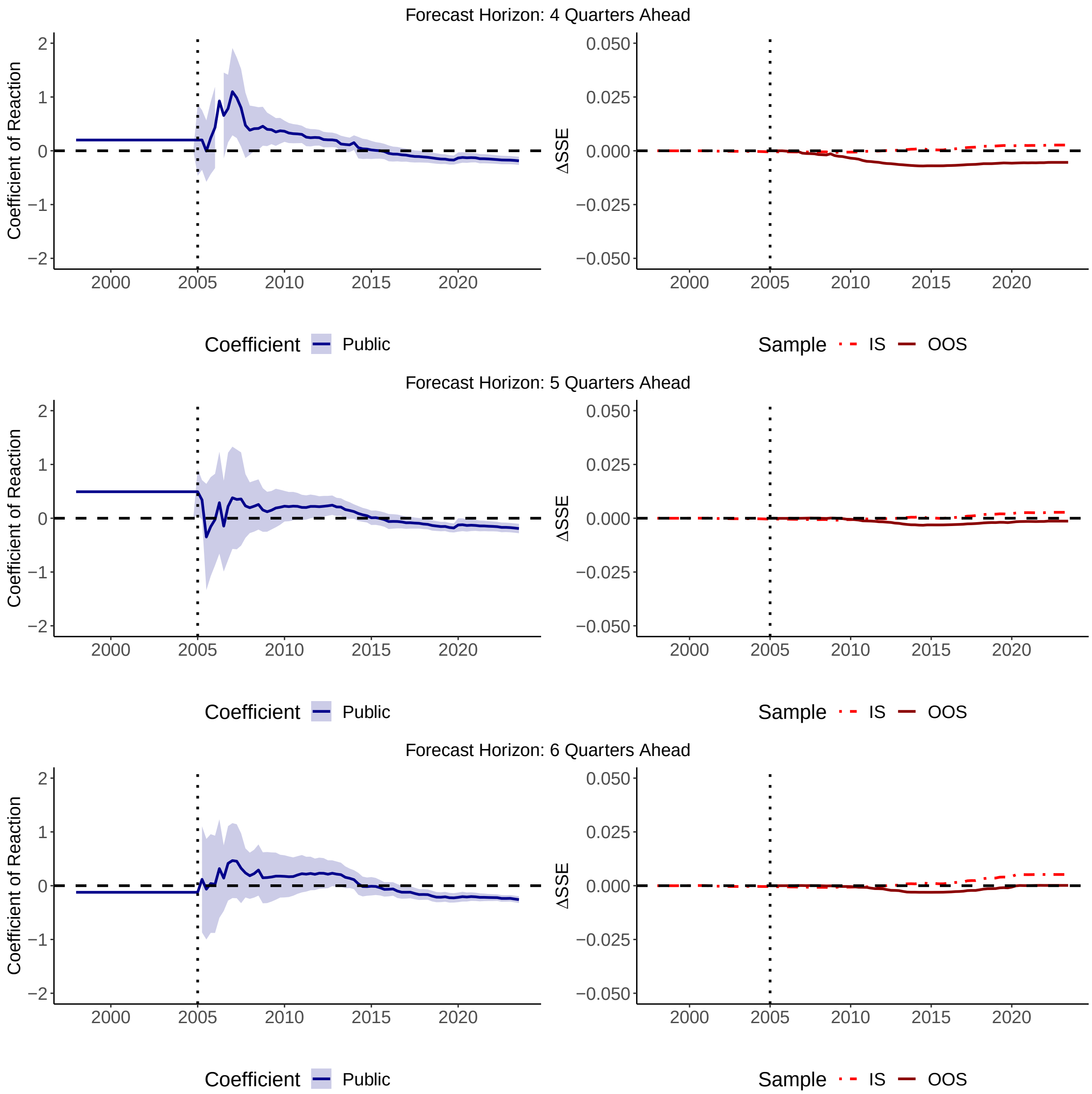

Figure O4 (cont.): OOS Credibility Test of the Rational Consensus Estimate

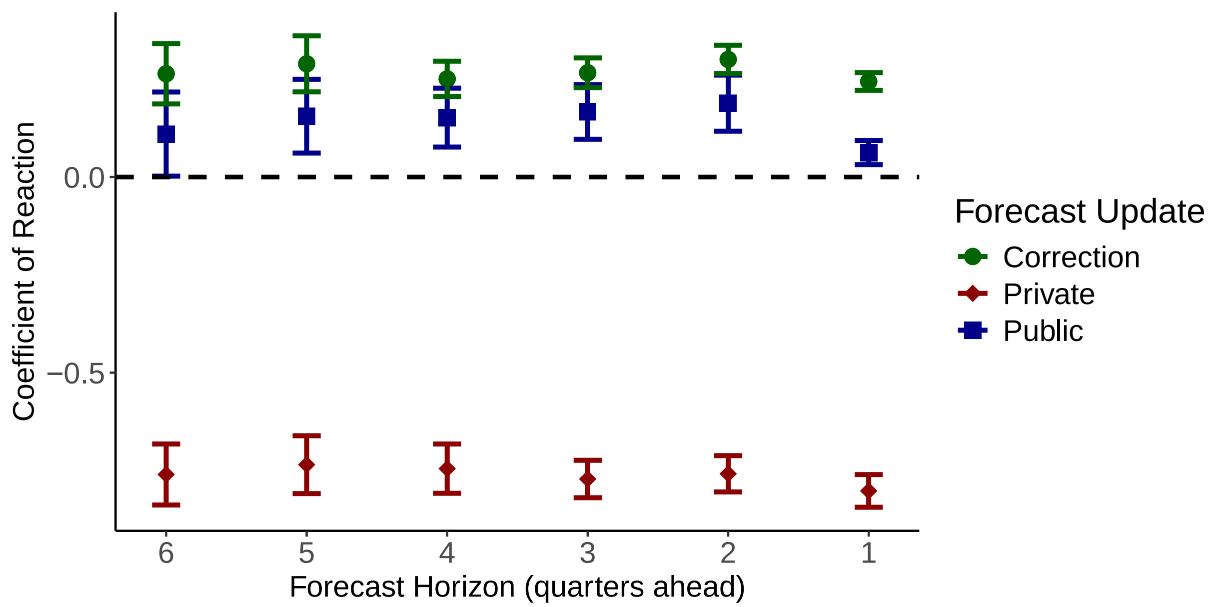

Figure O5: Estimated coefficients from the Decomposed model when at least 2 analysts reported on day t.

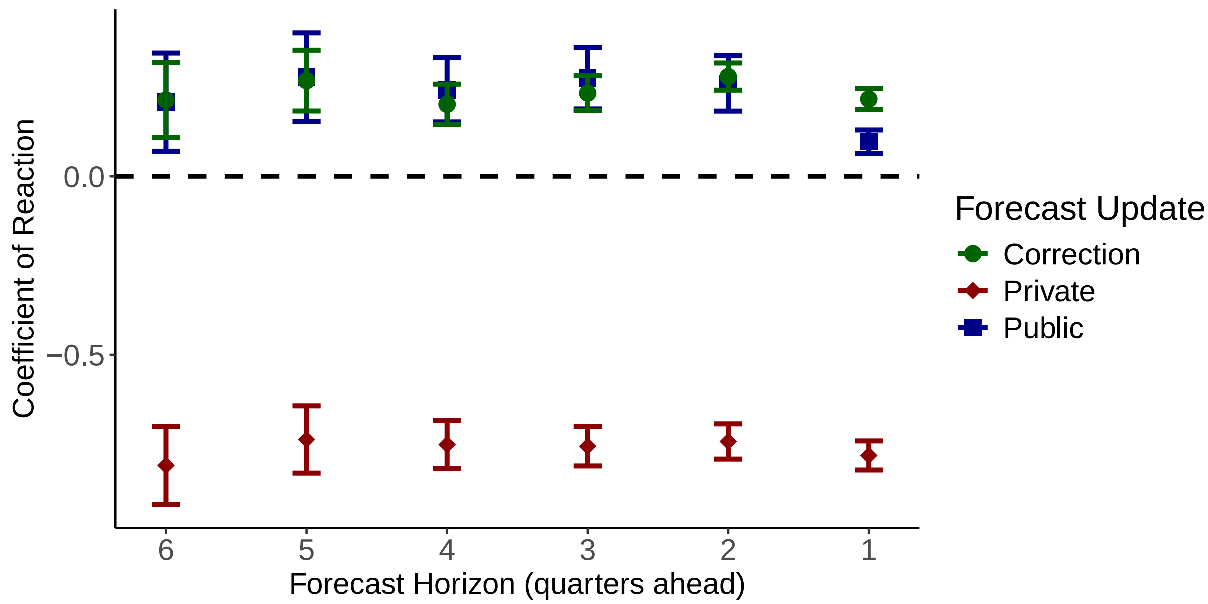

Figure O6: Estimated coefficients from the Decomposed model when at least 4 analysts reported on day t.

O.2.4 Lowering the Frequency of Revision

The OOS tests shown in Figure 6 suggest significant improvement of expectations formation once analysts' reported forecasts are corrected for the presence of bias identified in any of the three factors of the decomposed model. But are these estimations robust to lower-frequency data? To test this, I regress forecast errors on forecast revisions, focusing on forecasts that revised past reports within a window spanning 8 to 12 weeks. In essence, I extend the revision horizon k in (1) to capture updates over a larger time frame. Recall these forecast updates are defined as:

$$\Delta_k F^{i,j}_{t+h} \equiv F^{i,j}_{t+h|t} - F^{i,j}_{t+h|t-k}$$

Figure O7 suggests that the results are robust to this specification.

O.2.5 More Insights on the Success Rate of the Behavioral Model

Does the *behavioral* model exhibit superior performance to the *rational* model as a consequence of a minority of analysts reporting *biased* forecasts? To examine this, I expand the rolling-forward methodology and assess the following OOS statistic:

$$RSSE_{i,q} = \frac{SSE^B_{i,q}}{SSE^R_{i,q}} \tag{O.10}$$

where,

$$SSE^R_{i,q} = \sum_t \sum_h \sum_j (x^j_{t+h} - F^{i,j}_{t+h|t})^2 \quad \forall t \in q \tag{O.11}$$

$$SSE^B_{i,q} = \sum_t \sum_h \sum_j (x^j_{t+h} - \hat{x}^{i,j}_{t+h|t})^2 \quad \forall t \in q \tag{O.12}$$

There are two key observations in this context. Firstly, the sum of squared errors is computed individually for each analyst in every quarter throughout the testing period. Secondly, the sum of squared errors does not accumulate past periods' errors as the testing sample progresses. The $RSSE_{i,q}$ statistic enables the *behavioral* model to compete with each analyst individually. Any value of $RSSE_{i,q}$ below one indicates that, in quarter q, the *behavioral* model produced more accurate expectations than analyst i's reported forecasts when both were evaluated based on their performance over the same set of forecasts.[31] The $RSSE_{i,q}$ statistic is insightful, as each quarter allows an assessment of whether the *behavioral* model outperforms the median analyst. In other words:

$$MRSSE_q = \texttt{median}(RSSE_{i,q}) \tag{O.13}$$

If $MRSSE_q$ is less than 1, it signifies that in testing quarter q, the *behavioral* model out-

[31] It's important to note that the *behavioral* model doesn't possess superior information that analyst i wasn't privy to, as $\hat{x}^{i,j}_{t+h|t}$ is estimated with coefficient estimates from a training sample preceding the testing sample (refer to Equation (17)).

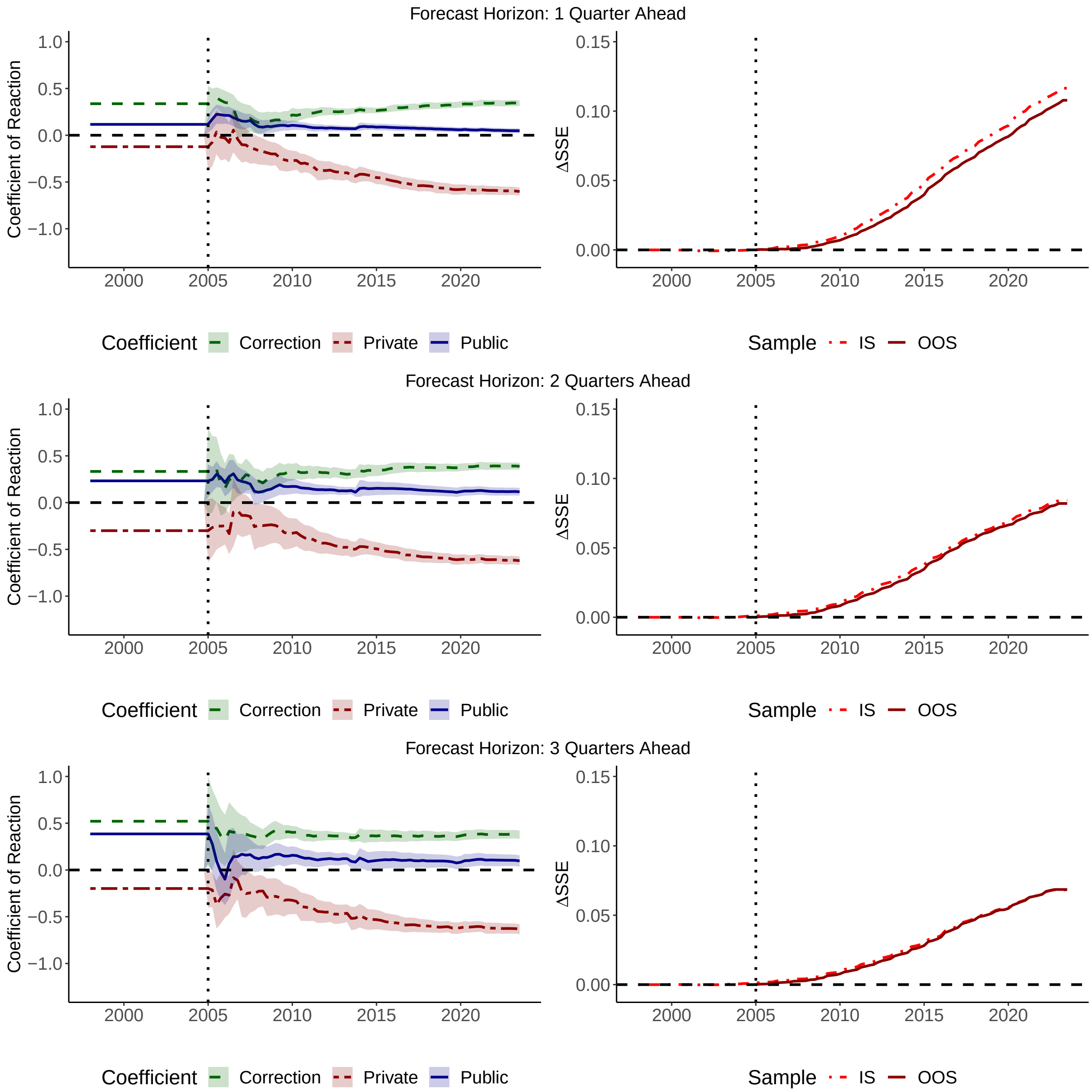

Figure O7: OOS Predictability of Forecast Errors from the Decomposed Model in Low-Frequency Data

Notes: This figure tests the robustness of the results of Figure 6 in forecast updates of lower-frequency. The revision horizon k in forecast revisions is modified such that the current forecast succeeds the initial forecast in a window between 8 to 12 weeks. The plots on the LHS show the three coefficients of the decomposed model from a regression of Forecast Errors on Forecast Revisions due to public signals, private signals, and the correction factor, at the individual (analyst) level using high-frequency data. Every calendar quarter q in the testing period uses coefficients that are estimated with a rolling forward methodology (see main text) on revenues that were released up until quarter $q-1$. The shaded areas are 95% confidence intervals using double-clustered standard errors at the announcement period (calendar quarter) and analyst. The plots on the RHS show the In-Sample (IS) and Out-of-Sample (OOS) performance of the ΔSSE statistic (see main text). The vertical dotted lines show the end of the training period of the rolling forward methodology and the beginning of the testing period.

performs the majority of analysts by minimizing their forecast errors after correcting for the behavioral bias. Additionally, the success rate in any quarter q is defined as the percentage of analysts who reported less accurate forecasts than the behavioral model's expectations. In other words, the success rate identifies the percentage of individuals with $RSSE_{i,q} \leq 1$.

Figure O8 presents the two statistics for forecasts with horizons of 1, 2, and 3 quarters. The plots on the LHS depict the $MRSSE_q$ for three alternative *behavioral* models. The first corrects bias identified by the BGMS coefficient, the second corrects bias identified in all three factors of the decomposed model when the *rational* consensus is estimated with the Kalman Filter algorithm, and the third corrects bias in all three factors of the decomposed model when a simple 7-day average forecast proxies for the *rational* consensus. Across all horizons, the BGMS coefficient exhibits poor performance, fluctuating around the value of 1 (i.e., the cutoff point determining whether the median analyst's forecasts can be improved.) In contrast, the decomposed model consistently outperforms the median forecaster, and the results remain robust to different specifications of the *rational* consensus in the decomposition process.

The RHS plots in Figure O8 demonstrate that the decomposed model consistently improves the forecasts of the majority of analysts across all testing periods and horizons. Specifically, the baseline decomposed model (utilizing Kalman Filter in the decomposition process) improves forecasts for an average of 56.89% of analysts (with a standard deviation of 3.30%) over time, as indicated by the line plots. This evidence is robust even when the decomposition process uses a 7-day average forecast to proxy for the *rational* consensus. In contrast, the BGMS model exhibits a different pattern. A *behavioral* model aiming to correct bias with the BGMS coefficient shows a more limited ability to improve forecasts for the majority of analysts across most tested periods. On average, 50.32% of analysts (with a standard deviation of 5.49%) could have improved their reported forecasts by considering the bias identified from the BGMS coefficient.

The results underscore three crucial observations. First, any evidence suggesting that analysts do not report forecasts consistent with the rational expectations hypothesis in this study is not attributed to a minority of analysts reporting significantly worse expectations than the *behavioral* model, but rather to the majority of analysts tested. Second, despite the BGMS model's failure to consistently outperform analysts out-of-sample, a non-trivial percentage of analysts could have at least marginally improved their forecasts using this model. Third, the supporting evidence in favor of heterogeneous responses to signals, as demonstrated in the main results, is further reinforced by the decomposed model's ability to effectively improve forecasts for a greater number of analysts compared to the BGMS model.

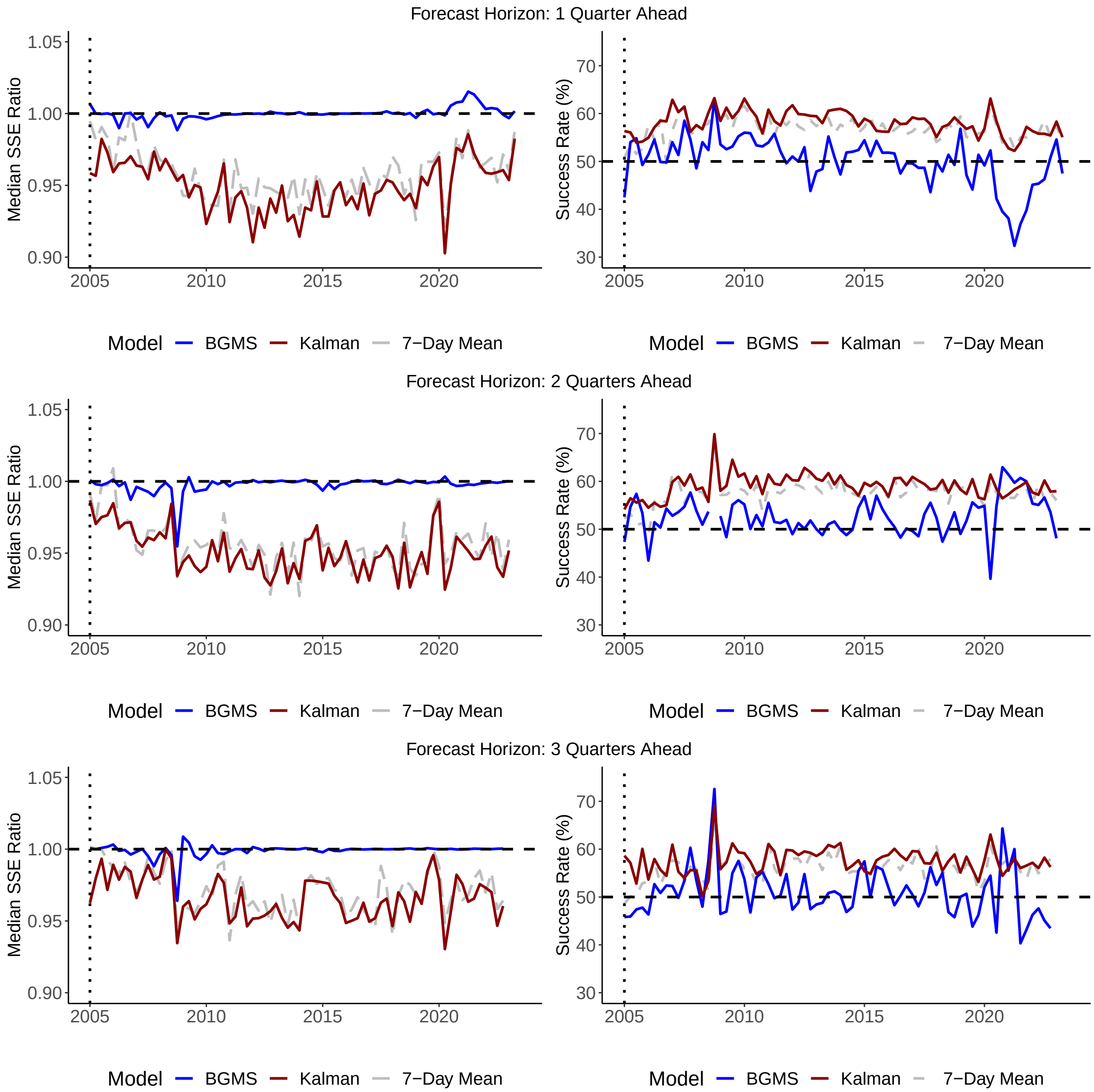

Figure O8: Does the behavioral model effectively correct the bias for a majority or minority of analysts?

Notes: This figure tests the robustness of the results by estimating the performance of the behavioral model relative to the median forecaster and the percentage of forecasters who could have improved their reported forecasts by taking into account the behavioral bias identified in any of the three models (BGMS model; Decomposed model that uses the Kalman Filter to proxy for the *rational* consensus; and Decomposed model that uses a 7-Day mean forecast to proxy for the *rational* consensus). The plots on the LHS show the median sum-squared-error ratio (see equation O.13). The plots on the RHS report the success rate defined as the percentage of analysts whose forecasts were worse than or at most as successful as the behavioral model's expectations.

O.2.6 Controlling for Adaptive Learning

The decomposed model proposed in the main paper identifies three factors from analysts' *overall* revisions as candidate factors to explain forecast errors. Analysts might learn from past forecast errors and adapt their response to news signals. This in turn might suggest that the observed overreaction to private signals coincides with analysts becoming more responsive to shocks after learning about forecast errors they made on their past forecasts. To control for this feedback mechanism, I include analysts' past forecast errors as a fourth factor in the regressions. For most observations, past forecast errors are available on forecasts announced within 60 days prior to their contemporaneous forecast.

Figure O9 shows the coefficient estimates from the IS regressions. According to the results, the adjusted R-squared increases while the coefficients of the three factors do not change much with the exception of public signals that lose in terms of significance. That said, past forecast errors are related to public signals as firms announce their realized revenues, and therefore, the coefficient of reaction to public signals is expected to change with the inclusion of this variable. Analytical results are shown in Table O9.

Figure O17 shows that the results are robust out-of-sample, and in fact, ΔSSE statistic improves compared to the decomposed three-factor model of the main results.

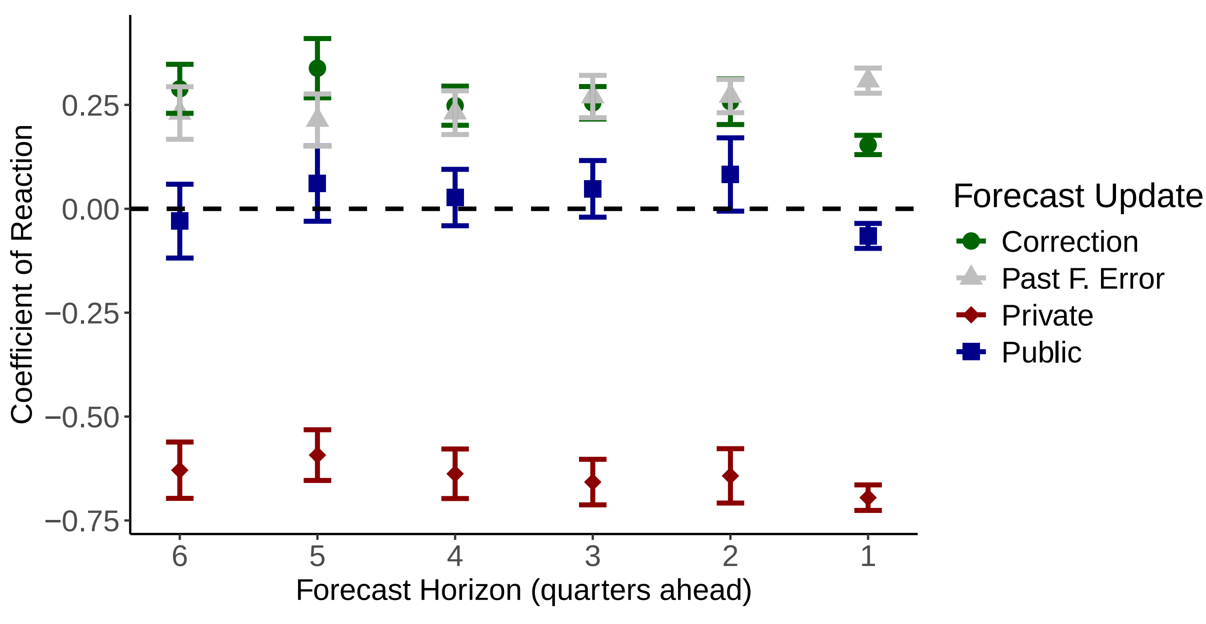

Figure O9: Controlling for Past Forecast Errors

O.3 Supporting Tables and Figures

Table O1: Summary Statistics

Panel A [a]						
Forecast Horizon (Quarters Ahead)	1	2	3	4	5	6
Observations	421,785	368,671	314,946	247,911	169,056	113,884
Companies	7,798	7,594	7,304	6,954	6,267	5,634
Analysts	4,465	4,465	4,462	4,386	3,759	3,364
Brokerage Houses affiliated with Analysts	448	438	428	399	358	323
Earliest Forecast Announcement	Apr 1996	Jan 1998	Apr 1996	Oct 1997	Sep 1997	Sep 1997
Latest Forecast Announcement	Jun 2023	Mar 2023	Dec 2022	Sep 2022	Jun 2022	Mar 2022
Panel B						
	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Observations by analyst [b]	366.462	590.762	16	66	418	9,172
Companies covered by analyst	19.827	14.179	1	9	27	112
Companies covered per month by analyst	2.565	2.602	1	1	3	55
Forecast announcement days by analyst	90.165	111.146	2	20	116	1,307
Forecasts reported for each firm by analyst [c]	18.483	30.063	1	4	20	789
Times revised by analyst [d]	2.032	1.723	1	1	2	40
Analysts per outcome [e]	3.901	3.790	1	1	5	43
Panel C						
	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Forecast Error	−0.011	0.184	−1.039	−0.063	0.057	0.754
Overall Forecast Revision	−0.004	0.077	−1.781	−0.019	0.015	1.566
Revision from Public Signals	−0.005	0.074	−1.583	−0.020	0.015	1.479
Revision from Private Signals	−0.002	0.053	−1.545	−0.011	0.011	0.961
Correction Revision	0.003	0.065	−1.152	−0.007	0.009	1.488
Panel D [a]						
	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Forecasts per outcome	35.714	39.175	3	10	46	546
Analysts per outcome	8.118	6.421	2	4	11	61

Notes: This table reports summary statistics of forecasts with horizon between (incl.) 1 and 6 quarters. Forecast horizon measures how many quarters the forecast was announced ahead of the firm-revenue announcement date. All forecast errors and forecast revisions are computed by taking realized revenues and forecasts in logs.

Footnotes:

[a] Panels A, B, and C report statistics on data of which analysts revised their initial forecasts and revision was reported within 4 weeks. These observations were used in main regressions. Panel D reports statistics on all forecasts (including forecasts where analysts did not revise their initial forecasts or revised beyond the 4-week window); these observations were used in the Kalman Filter to estimate the rational consensus. A minimum of 2 analysts per outcome and 3 forecasts per outcome was a requirement in the Kalman Filter process.

[b] A minimum of 30 observations per analyst was required in the forecast-decomposition and group-formation process. The minimum 16 shown here is due to some observations falling off the window of 1-6 quarters-ahead horizon.

[c] Number of forecasts reported by an analyst regardless of the period-end date of the realization of revenues.

[d] Number of times an analyst revised her initial forecast on revenues about a specific firm and period-end date.

[e] Outcome denotes revenues of a specific firm and period-end date (fiscal quarter).

Table O2: IS Cross-Horizon Coefficient estimates from the BGMS and CG Models.

Horizon (Quarters Ahead):	1	2	3	4	5	6
Panel A			BGMS Model			
Forecast revision	-0.0521***	0.0148	-0.0179	-0.0218	-0.0264	-0.1025**
	(0.0140)	(0.0301)	(0.0277)	(0.0309)	(0.0387)	(0.0416)
Constant	0.0056**	-0.0082	-0.0166**	-0.0215***	-0.0234**	-0.0283***
	(0.0025)	(0.0054)	(0.0067)	(0.0075)	(0.0095)	(0.0108)
Observations	421,785	368,671	314,946	247,911	169,056	113,884
Adjusted R^2	0.0011	0.0000	0.0000	0.0001	0.0001	0.0009
Panel B			CG Model			
Forecast revision	0.1292***	0.2276***	0.2437***	0.2865***	0.9745***	-0.0964
	(0.0141)	(0.0256)	(0.0283)	(0.0465)	(0.3348)	(0.3974)
Constant	-0.0062***	-0.0216***	-0.0306***	-0.0351***	-0.1142***	-0.1060***
	(0.0021)	(0.0039)	(0.0051)	(0.0060)	(0.0206)	(0.0307)
Observations	713,772	645,154	560,990	154,645	1,177	584
Adjusted R^2	0.0032	0.0052	0.0040	0.0024	0.0224	-0.0015

Notes: The table shows the estimated coefficients of the BGMS model (Panel A) and the CG model (Panel B) from a regression of forecast errors on (overall) forecast revisions. The regression of the BGMS model is at the individual (analyst) level while the regression of the CG model is at the consensus level. The BGMS model uses observations from the IBES Detail File while the CG model uses observations from the IBES Summary File. Separate regressions are run for different forecast horizons (from 1 to 6 quarters ahead of firm-revenue announcement.) Standard errors are double-clustered (quarter of forecast announcement-individual analyst) in the BGMS model, and clustered at the quarter of forecast announcement in the CG model. Significance levels: 10% (*), 5% (**) and 1% (***).

Table O3: IS Cross-Horizon Coefficient estimates from the Decomposed Model.

	Forecast Horizon (Quarters Ahead)					
	1	2	3	4	5	6
Public	0.0179	0.1329***	0.0949***	0.0836**	0.0878*	0.0077
	(0.0160)	(0.0415)	(0.0356)	(0.0364)	(0.0452)	(0.0487)
Private	-0.7293***	-0.6529***	-0.6486***	-0.6117***	-0.5806***	-0.6179***
	(0.0233)	(0.0333)	(0.0290)	(0.0301)	(0.0307)	(0.0341)
Correction	0.3218***	0.3946***	0.3840***	0.3696***	0.4162***	0.3849***
	(0.0165)	(0.0233)	(0.0192)	(0.0224)	(0.0283)	(0.0266)
Constant	0.0040	-0.0099*	-0.0184***	-0.0229***	-0.0249***	-0.0300***
	(0.0025)	(0.0053)	(0.0067)	(0.0075)	(0.0094)	(0.0107)
Observations	421,785	368,671	314,946	247,911	169,056	113,884
Adjusted R^2	0.1481	0.0950	0.0729	0.0520	0.0440	0.0429

Notes: The table shows the estimated coefficients of the decomposed model from a regression of forecast errors on three factors, namely, revisions from public signals, revisions from private signals, and forecast correction (past disagreement with consensus) as shown in equation 9. Separate regressions are run for different forecast horizons (from 1 to 6 quarters ahead of firm-revenue announcement.) Double-clustered standard errors are used (quarter of forecast announcement-individual analyst). Significance levels: 10% (*), 5% (**) and 1% (***).

Table O4: Do analysts exhibit more or less profound underreaction to negative public signals?

	Forecast Horizon (Quarters Ahead)					
	1	2	3	4	5	6
Public	0.0059	0.0272	-0.0500	-0.0526*	-0.0292	-0.1246***
	(0.0238)	(0.0248)	(0.0305)	(0.0284)	(0.0377)	(0.0438)
Public (Negative Revision)	0.0184	0.1722***	0.2543***	0.2603***	0.2194***	0.2652***
	(0.0304)	(0.0498)	(0.0473)	(0.0518)	(0.0683)	(0.0834)
Private	-0.7295***	-0.6543***	-0.6509***	-0.6101***	-0.5784***	-0.6164***
	(0.0233)	(0.0327)	(0.0282)	(0.0298)	(0.0308)	(0.0347)
Correction	0.3220***	0.3967***	0.3875***	0.3779***	0.4222***	0.3927***
	(0.0164)	(0.0230)	(0.0189)	(0.0233)	(0.0290)	(0.0279)
Constant	0.0043**	-0.0068	-0.0140**	-0.0178**	-0.0205**	-0.0248**
	(0.0022)	(0.0048)	(0.0063)	(0.0071)	(0.0089)	(0.0102)
Observations	421,785	368,671	314,946	247,911	169,056	113,884
Adjusted R^2	0.1481	0.0960	0.0746	0.0535	0.0449	0.0440

Notes: The table repeats the results from Table O3 by including an interaction term of revisions from public signals with a dummy variable that is 1 when the public signal is negative and zero otherwise, as a fourth factor in the decomposed model. A statistically significant coefficient of the interaction term 'Public (Negative Revision)' is evidence that analysts' reaction to public signals changes significantly depending on if news have a positive or negative impact on expectations formation. Separate regressions are run for different forecast horizons (from 1 to 6 quarters ahead of firm-revenue announcement.) Double-clustered standard errors are used (quarter of forecast announcement-individual analyst). Significance levels: 10% (*), 5% (**) and 1% (***).

Table O5: Do analysts exhibit more or less profound overreaction to negative private signals?

	Forecast Horizon (Quarters Ahead)					
	1	2	3	4	5	6
Public	0.0152	0.1301***	0.0931***	0.0824**	0.0868*	0.0081
	(0.0168)	(0.0407)	(0.0353)	(0.0365)	(0.0456)	(0.0500)
Private	-0.8234***	-0.7637***	-0.7709***	-0.7595***	-0.7450***	-0.7654***
	(0.0483)	(0.0420)	(0.0431)	(0.0402)	(0.0479)	(0.0592)
Private (Negative Revision)	0.1539*	0.1836**	0.2076**	0.2667***	0.2941***	0.2635**
	(0.0804)	(0.0787)	(0.0806)	(0.0934)	(0.1056)	(0.1252)
Correction	0.3234***	0.3958***	0.3860***	0.3752***	0.4203***	0.3890***
	(0.0162)	(0.0230)	(0.0191)	(0.0234)	(0.0289)	(0.0275)
Constant	0.0056***	-0.0076	-0.0157***	-0.0193***	-0.0209**	-0.0263***
	(0.0022)	(0.0046)	(0.0060)	(0.0068)	(0.0086)	(0.0099)
Observations	421,785	368,671	314,946	247,911	169,056	113,884
Adjusted R^2	0.1487	0.0956	0.0735	0.0528	0.0449	0.0436

Notes: The table repeats the results from Table O3 by including an interaction term of revisions from private signals with a dummy variable that is 1 when the private signal is negative and zero otherwise, as a fourth factor in the decomposed model. A statistically significant coefficient of the interaction term 'Private (Negative Revision)' is evidence that analysts' reaction to private signals changes significantly depending on if news have a positive or negative impact on expectations formation. Separate regressions are run for different forecast horizons (from 1 to 6 quarters ahead of firm-revenue announcement.) Double-clustered standard errors are used (quarter of forecast announcement-individual analyst). Significance levels: 10% (*), 5% (**) and 1% (***).

Table O6: Do analysts exhibit persistency in overconfidence or underconfidence?

	Forecast Horizon (Quarters Ahead)					
	1	2	3	4	5	6
Public	0.0163	0.1317***	0.0938***	0.0827**	0.0872*	0.0082
	(0.0165)	(0.0415)	(0.0355)	(0.0364)	(0.0454)	(0.0491)
Private	-0.7309***	-0.6540***	-0.6512***	-0.6142***	-0.5829***	-0.6213***
	(0.0232)	(0.0333)	(0.0286)	(0.0294)	(0.0302)	(0.0333)
Correction	0.2959***	0.3619***	0.3256***	0.3028***	0.3342***	0.3034***
	(0.0270)	(0.0284)	(0.0244)	(0.0290)	(0.0419)	(0.0483)
Correction (Negative Disagreement)	0.0659	0.0825**	0.1421***	0.1605***	0.1921***	0.1866**
	(0.0461)	(0.0411)	(0.0441)	(0.0530)	(0.0659)	(0.0839)
Constant	0.0048**	-0.0088*	-0.0164***	-0.0206***	-0.0221**	-0.0272***
	(0.0022)	(0.0051)	(0.0063)	(0.0071)	(0.0090)	(0.0102)
Observations	421,785	368,671	314,946	247,911	169,056	113,884
Adjusted R^2	0.1483	0.0952	0.0734	0.0525	0.0446	0.0434

Notes: The table repeats the results from Table O3 by including an interaction term of the Correction revision with a dummy variable that is 1 when the Correction revision is negative and zero otherwise, as a fourth factor in the decomposed model. A statistically significant coefficient of the fourth factor 'Correction (Negative Disagreement)' is evidence that the coefficient of the 'Correction' revision changes significantly depending on if analysts were more confident or less confident than the *rational* consensus during period $t-1$. Separate regressions are run for different forecast horizons (from 1 to 6 quarters ahead of firm-revenue announcement.) Double-clustered standard errors are used (quarter of forecast announcement-individual analyst). Significance levels: 10% (*), 5% (**) and 1% (***).

Table O7: IS Inference of Cross-analyst heterogeneity in the BGMS model.

	Forecast Horizon (Quarters Ahead)					
	1	2	3	4	5	6
Bottom	-0.2518***	-0.2434***	-0.3329***	-0.3742***	-0.4573***	-0.5534***
	(0.0233)	(0.0380)	(0.0326)	(0.0290)	(0.0375)	(0.0475)
Low-Mid	-0.1250***	-0.0662*	-0.1162***	-0.1348***	-0.1813***	-0.2385***
	(0.0155)	(0.0364)	(0.0326)	(0.0397)	(0.0455)	(0.0475)
Mid	-0.0320*	0.0222	0.0141	0.0351	0.0358	-0.0646
	(0.0184)	(0.0340)	(0.0332)	(0.0393)	(0.0505)	(0.0520)
High-Mid	0.0603***	0.1695***	0.1606***	0.1566***	0.1998***	0.1857***
	(0.0171)	(0.0305)	(0.0261)	(0.0282)	(0.0255)	(0.0327)
Top	0.1751***	0.3618***	0.3976***	0.4414***	0.5707***	0.5728***
	(0.0170)	(0.0329)	(0.0362)	(0.0391)	(0.0473)	(0.0509)
Wald (Top - Bottom)	219.16	144.64	224.71	281.03	290.14	261.78
Wald (High-Mid - Low-Mid)	64.61	24.62	43.99	35.72	53.38	54.15
Observations	421,785	368,671	314,946	247,911	169,056	113,884
Adjusted R^2	0.0093	0.0084	0.0135	0.0177	0.0196	0.0235

Notes: The table reports the cross-group BGMS coefficients from an OLS regression of equation 21. All analysts are assigned to one of the five quintile groups (Bottom, Low-Mid, Mid, High-Mid, Top) based on their individual β^i coefficient that is estimated from an OLS regression of (20). Columns 1-6 report the results for forecasts with different horizon (from 1 to 6 quarters ahead of firm-revenue announcement.) Double-clustered standard errors are used (quarter of forecast announcement-individual analyst). Significance levels: 10% (*), 5% (**) and 1% (***). Wald (Top - Bottom) is a one-sided Wald statistic that tests whether the coefficient of the Top group is greater than the coefficient of the Bottom group. Similarly, Wald (High-Mid - Low-Mid) tests if the coefficient from the High-Mid group is greater than the coefficient of the Low-Mid group.

Table O8: IS Inference of Cross-analyst heterogeneity in the Decomposed model.

	Forecast Horizon (Quarters Ahead)					
	1	2	3	4	5	6
	Public Signal					
Bottom	−0.1961***	−0.1643***	−0.2685***	−0.3071***	−0.3729***	−0.5049***
	(0.0299)	(0.0484)	(0.0348)	(0.0336)	(0.0511)	(0.0593)
Low-Mid	−0.0458**	0.0571	−0.0040	−0.0326	−0.0798	−0.1308**
	(0.0211)	(0.0520)	(0.0416)	(0.0462)	(0.0530)	(0.0550)
Mid	0.0413*	0.1562***	0.1416***	0.1649***	0.1572***	0.0594
	(0.0216)	(0.0491)	(0.0453)	(0.0492)	(0.0585)	(0.0647)
High-Mid	0.1182***	0.2826***	0.2784***	0.2527***	0.3170***	0.2995***
	(0.0188)	(0.0396)	(0.0354)	(0.0352)	(0.0352)	(0.0454)
Top	0.2098***	0.4443***	0.4956***	0.5240***	0.6548***	0.6628***
	(0.0210)	(0.0400)	(0.0436)	(0.0451)	(0.0528)	(0.0623)
	Private Signal					
Bottom	−0.8859***	−0.8551***	−0.8960***	−0.9134***	−0.8843***	−0.9443***
	(0.0374)	(0.0553)	(0.0491)	(0.0461)	(0.0319)	(0.0425)
Low-Mid	−0.8048***	−0.7096***	−0.7238***	−0.7052***	−0.7056***	−0.7091***
	(0.0390)	(0.0474)	(0.0399)	(0.0449)	(0.0483)	(0.0587)
Mid	−0.6791***	−0.6413***	−0.5948***	−0.5443***	−0.5066***	−0.6018***
	(0.0354)	(0.0295)	(0.0344)	(0.0390)	(0.0458)	(0.0403)
High-Mid	−0.6033***	−0.5248***	−0.5162***	−0.4368***	−0.3911***	−0.3468***
	(0.0301)	(0.0323)	(0.0364)	(0.0330)	(0.0411)	(0.0504)
Top	−0.4691***	−0.3188***	−0.3055***	−0.2226***	−0.0965	−0.0688
	(0.0381)	(0.0518)	(0.0523)	(0.0675)	(0.0782)	(0.0906)

Table O8: [cont.]

	Forecast Horizon (Quarters Ahead)					
	1	2	3	4	5	6
	Correction Revision					
Bottom	0.2512***	0.2789***	0.1936***	0.0820**	0.0848*	0.0450
	(0.0338)	(0.0450)	(0.0314)	(0.0373)	(0.0453)	(0.0601)
Low-Mid	0.3066***	0.3660***	0.3529***	0.3292***	0.3244***	0.3382***
	(0.0286)	(0.0424)	(0.0356)	(0.0348)	(0.0398)	(0.0352)
Mid	0.2991***	0.3602***	0.3856***	0.4128***	0.4880***	0.4117***
	(0.0237)	(0.0264)	(0.0267)	(0.0370)	(0.0473)	(0.0442)
High-Mid	0.3523***	0.4492***	0.4324***	0.4414***	0.4947***	0.4943***
	(0.0233)	(0.0254)	(0.0231)	(0.0317)	(0.0367)	(0.0415)
Top	0.4078***	0.6332***	0.6928***	0.7382***	0.9065***	0.9401***
	(0.0252)	(0.0336)	(0.0400)	(0.0397)	(0.0545)	(0.0628)
Wald (Public)	123.34	93.97	187.56	218.87	195.88	184.36
Wald (Private)	61.08	50.15	67.75	71.42	87.01	76.54
Wald (Correction)	13.77	39.82	96.16	145.25	134.45	106.16
Observations	421,785	368,671	314,946	247,911	169,056	113,884
Adjusted R^2	0.1558	0.1032	0.0860	0.0691	0.0627	0.0649

Notes: The table reports the cross-group coefficients of the three factors from the decomposed model (namely, revisions from public signals, private signals, and correction revisions) from an OLS regression of (22). All analysts are assigned to one of the five quintile groups (Bottom, Low-Mid, Mid, High-Mid, Top) based on their individual β^i coefficient that is estimated from an OLS regression of (20). Columns 1-6 report the results for forecasts with different horizons (from 1 to 6 quarters ahead of firm-revenue announcement.) Double-clustered standard errors are used (quarter of forecast announcement-individual analyst). Significance levels: 10% (*), 5% (**) and 1% (***). For every factor, a Wald one-sided statistic is reported that tests whether the factor's coefficient of the Top group is greater than the coefficient of the Bottom group.

Table O9: Inclussion of Past Forecast Errors in IS Regressions of the Decomposed Model

	Dependent Variable: Forecast Errors					
Forecast Horizon:	(1)	(2)	(3)	(4)	(5)	(6)
Public	-0.0653***	0.0825*	0.0480	0.0270	0.0606	-0.0298
	(0.0153)	(0.0450)	(0.0348)	(0.0346)	(0.0462)	(0.0453)
Private	-0.6951***	-0.6427***	-0.6577***	-0.6377***	-0.5928***	-0.6291***
	(0.0157)	(0.0334)	(0.0280)	(0.0304)	(0.0311)	(0.0346)
Correction	0.1536***	0.2577***	0.2550***	0.2481***	0.3383***	0.2887***
	(0.0118)	(0.0280)	(0.0200)	(0.0240)	(0.0364)	(0.0301)
Past Forecast Error	0.3083***	0.2713***	0.2701***	0.2313***	0.2136***	0.2305***
	(0.0154)	(0.0205)	(0.0260)	(0.0268)	(0.0320)	(0.0322)
Constant	0.0019	-0.0120**	-0.0205***	-0.0242***	-0.0255***	-0.0307***
	(0.0021)	(0.0050)	(0.0064)	(0.0073)	(0.0090)	(0.0103)
Observations	349,409	304,673	258,829	204,709	140,733	95,264
Adjusted R^2	0.2202	0.1257	0.0954	0.0644	0.0535	0.0515

Notes: The table shows the estimated coefficients from a regression of forecast errors on the three decomposed factors, namely, revisions from public signals, revisions from private signals, and forecast correction (past disagreement with consensus) as shown in Equation (9) and a control variable, namely, past forecast error. Past Forecast Error represents the individual analyst's forecast error on her most recent historical forecast reported (prior to day t) about realized firm revenues that become publicly available. Separate regressions are run for different forecast horizons (from 1 to 6 quarters ahead of firm-revenue announcement.) Standard errors are double clustered at the quarter of forecast announcement and (individual) analyst level. Significance levels: 10% (*), 5% (**) and 1% (***).

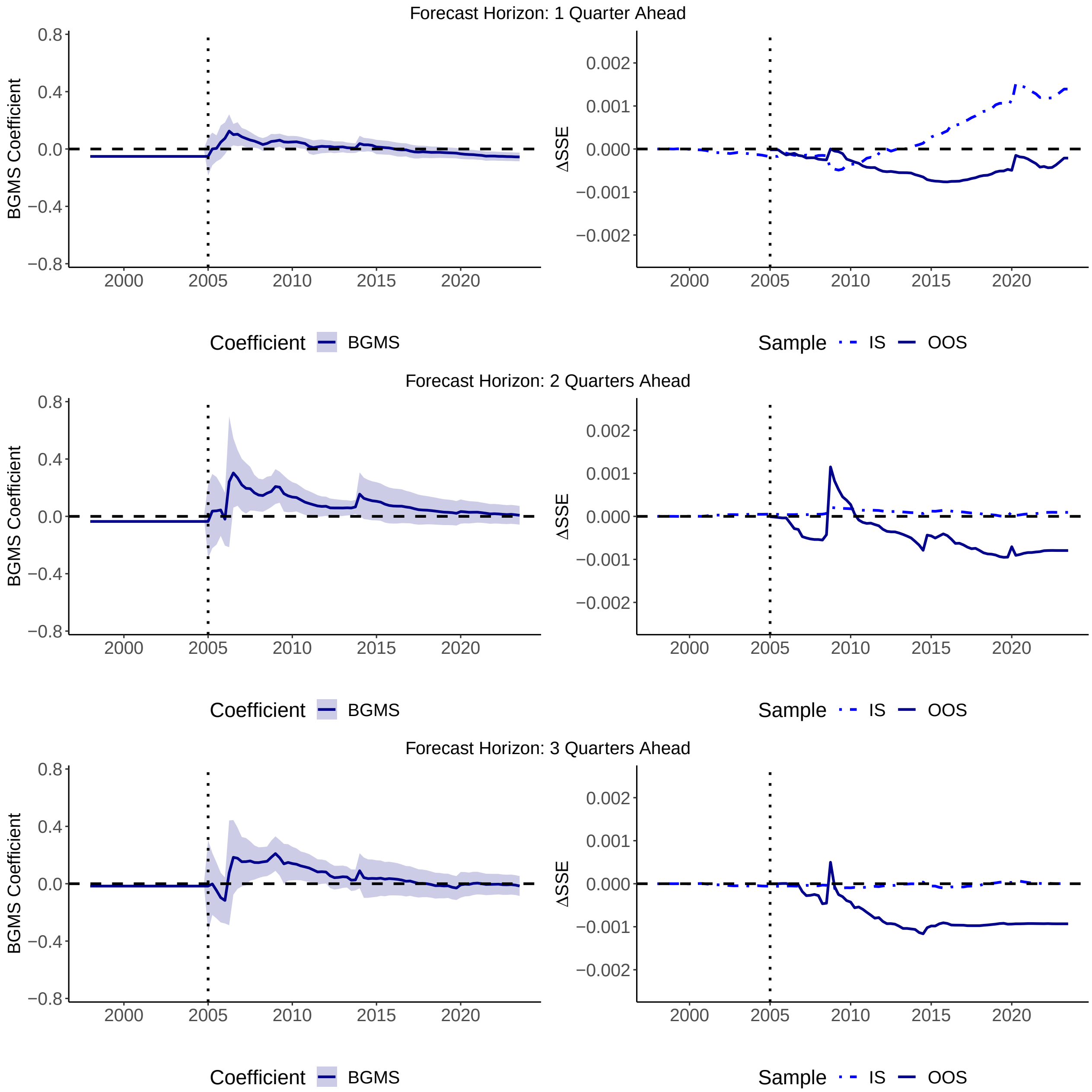

Figure O10: OOS Predictability of Forecast Errors from the BGMS Model

Notes: This figure supplements Figure 5 in the main text by showing all results for forecasts with a horizon of 1-6 quarters ahead.

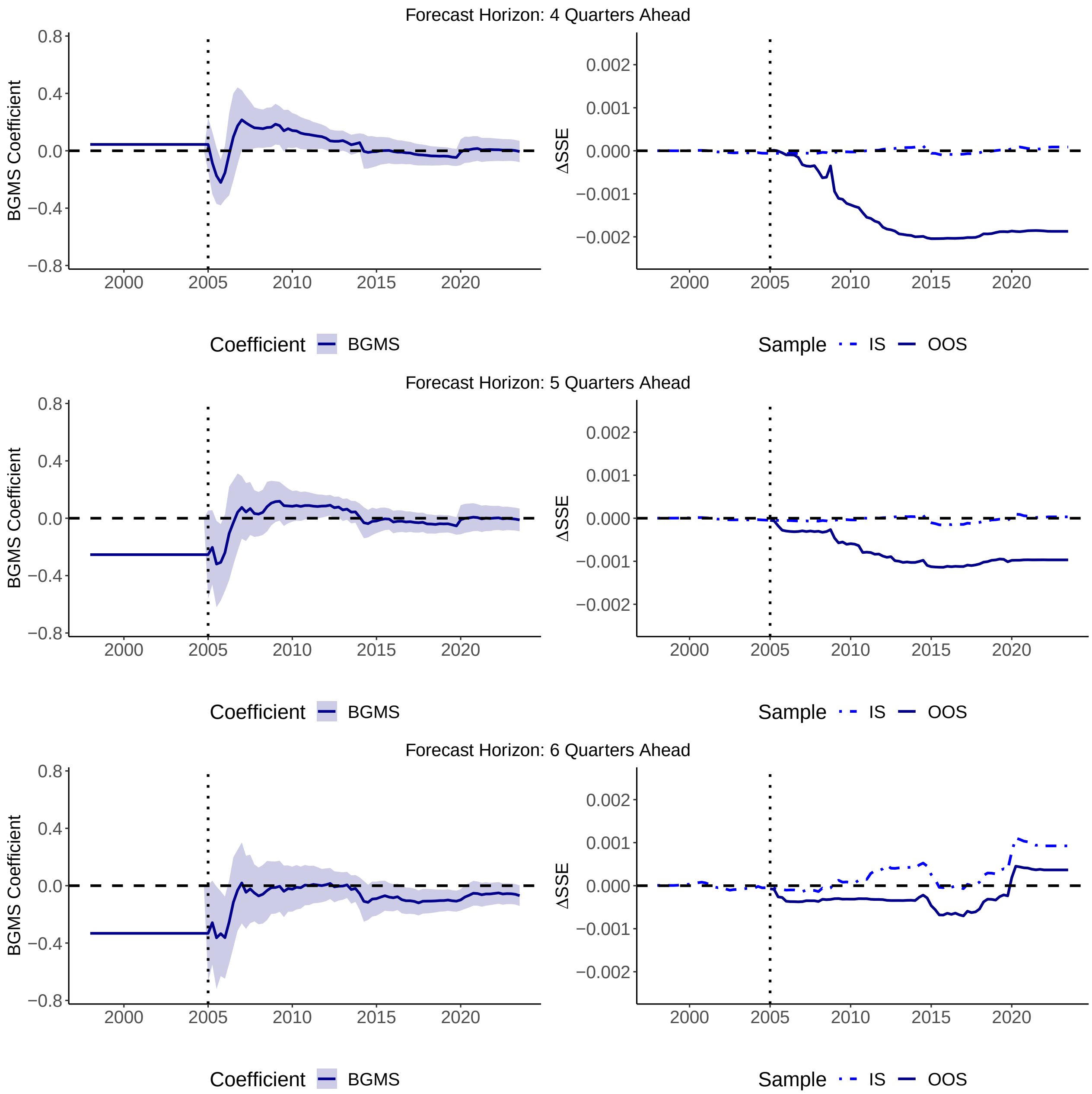

Figure O10 (cont.): OOS Predictability of Forecast Errors from the BGMS Model

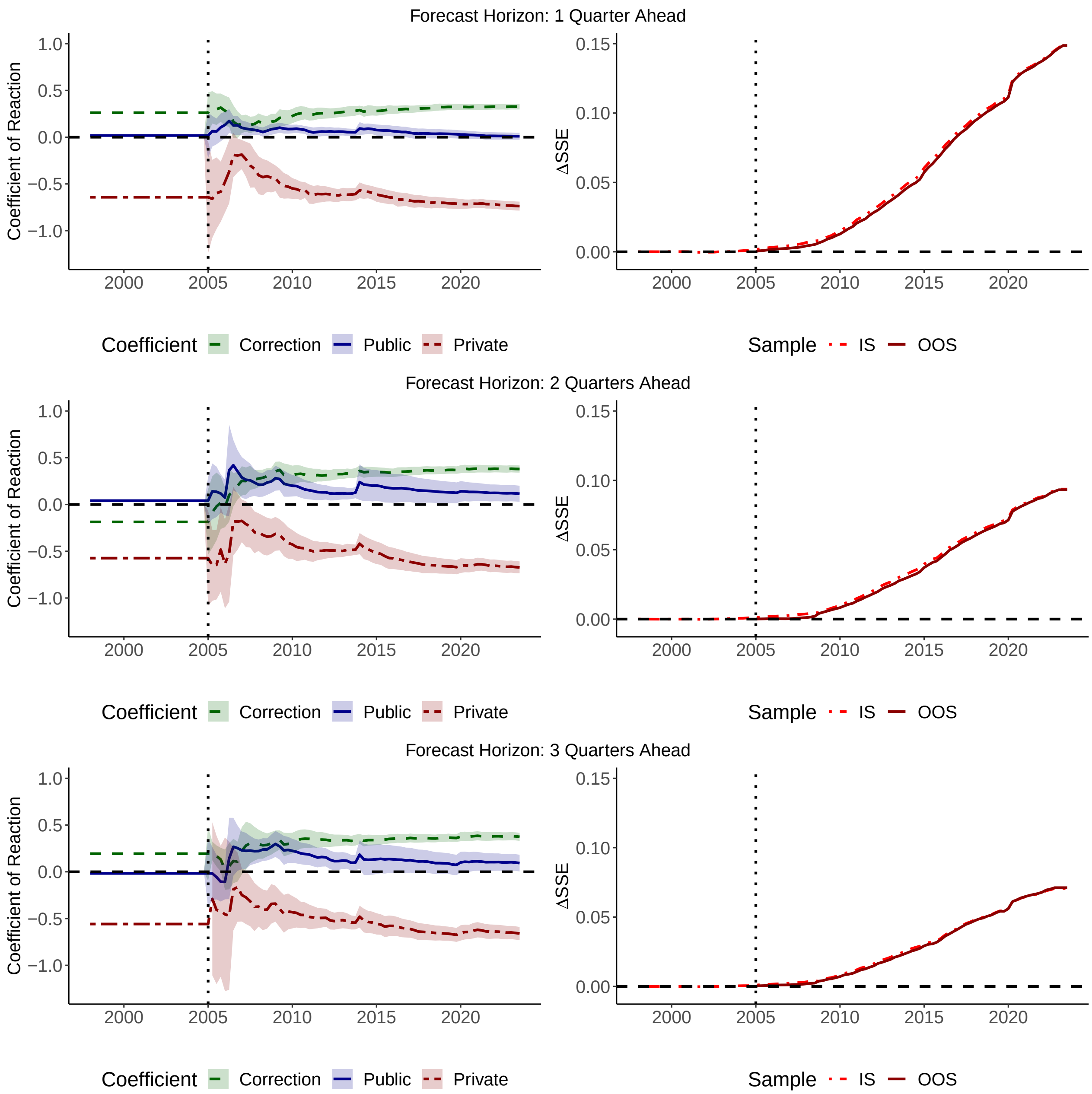

Figure O11: OOS Predictability of Forecast Errors from the Decomposed Model

Notes: This figure supplements Figure 6 by showing all results for forecasts with a horizon of 1-6 quarters ahead.

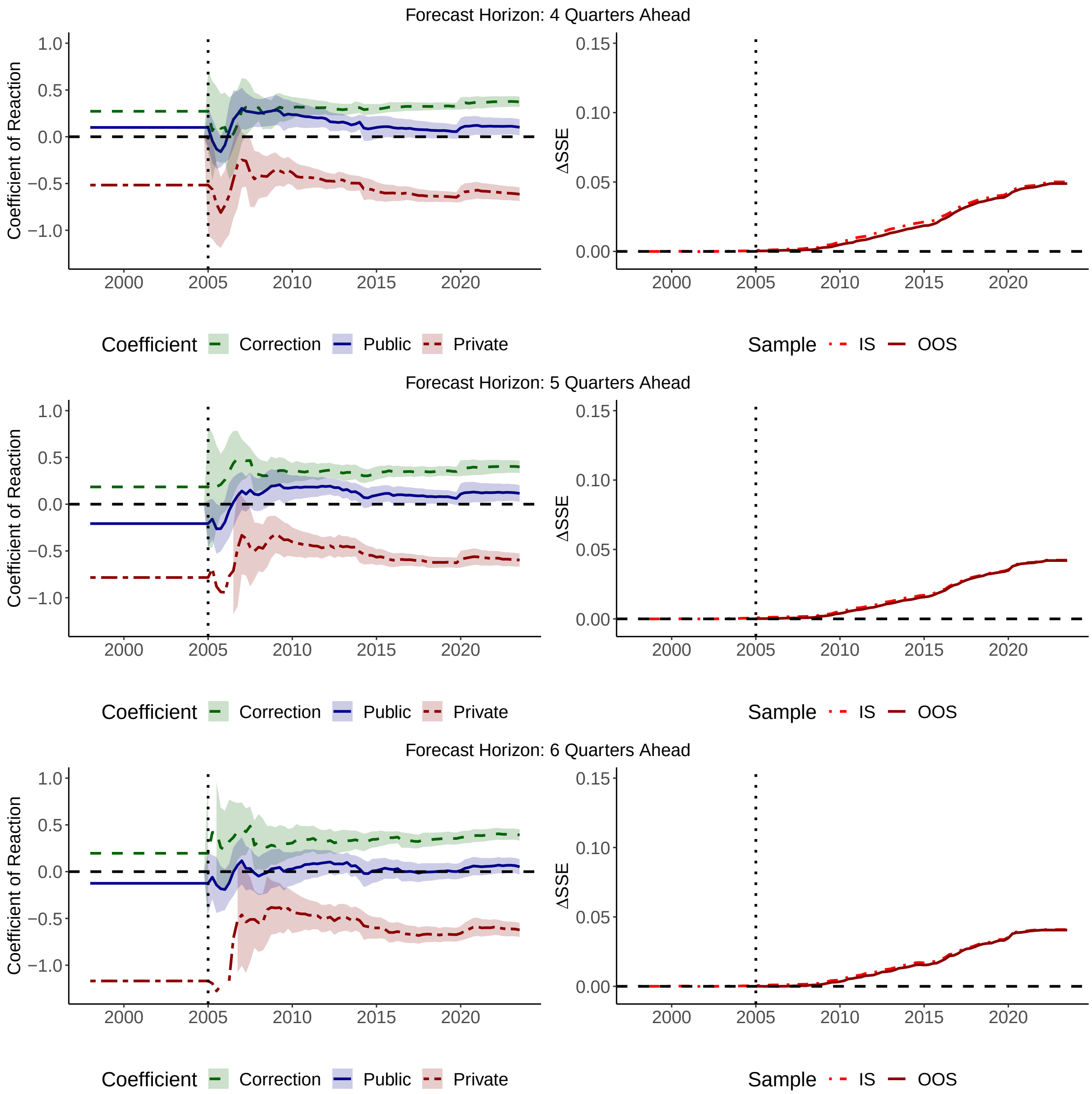

Figure O11 (cont.): OOS Predictability of Forecast Errors from the Decomposed Model

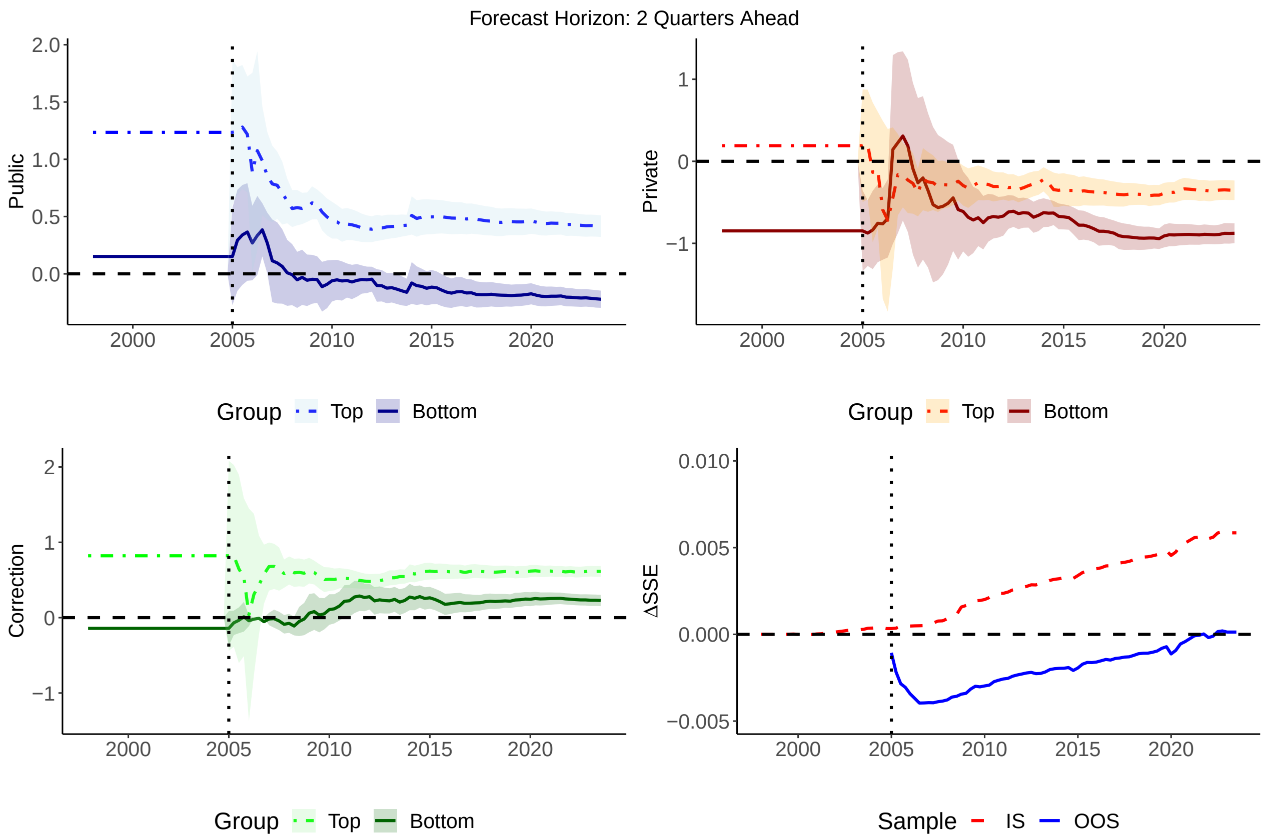

Figure O12: OOS performance of the decomposed model with cross-analyst heterogeneity

Notes: This figure supplements Figure 9 by showing the results from regressions of forecasts with horizon 2 quarters ahead.

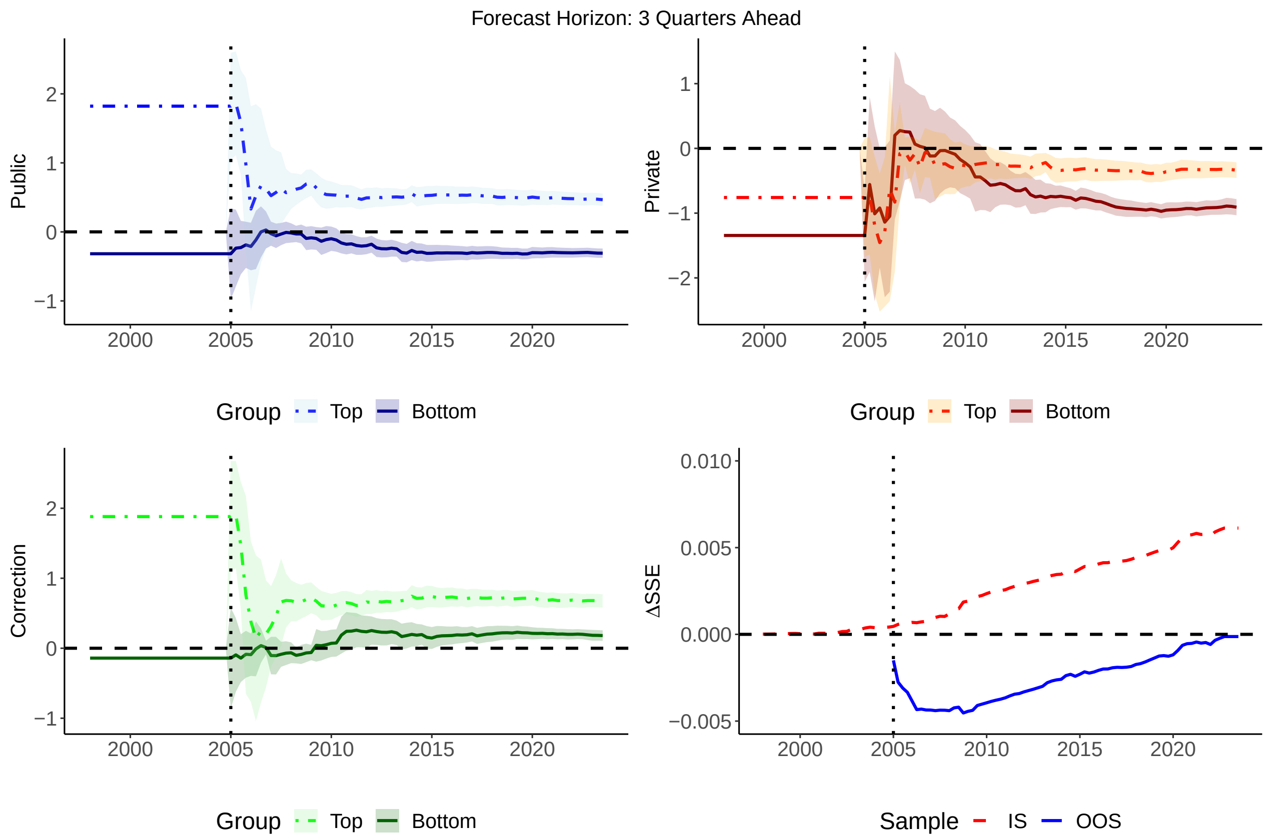

Figure O13: OOS performance of the decomposed model with cross-analyst heterogeneity

Notes: This figure supplements Figure 9 by showing the results from regressions of forecasts with horizon 3 quarters ahead.

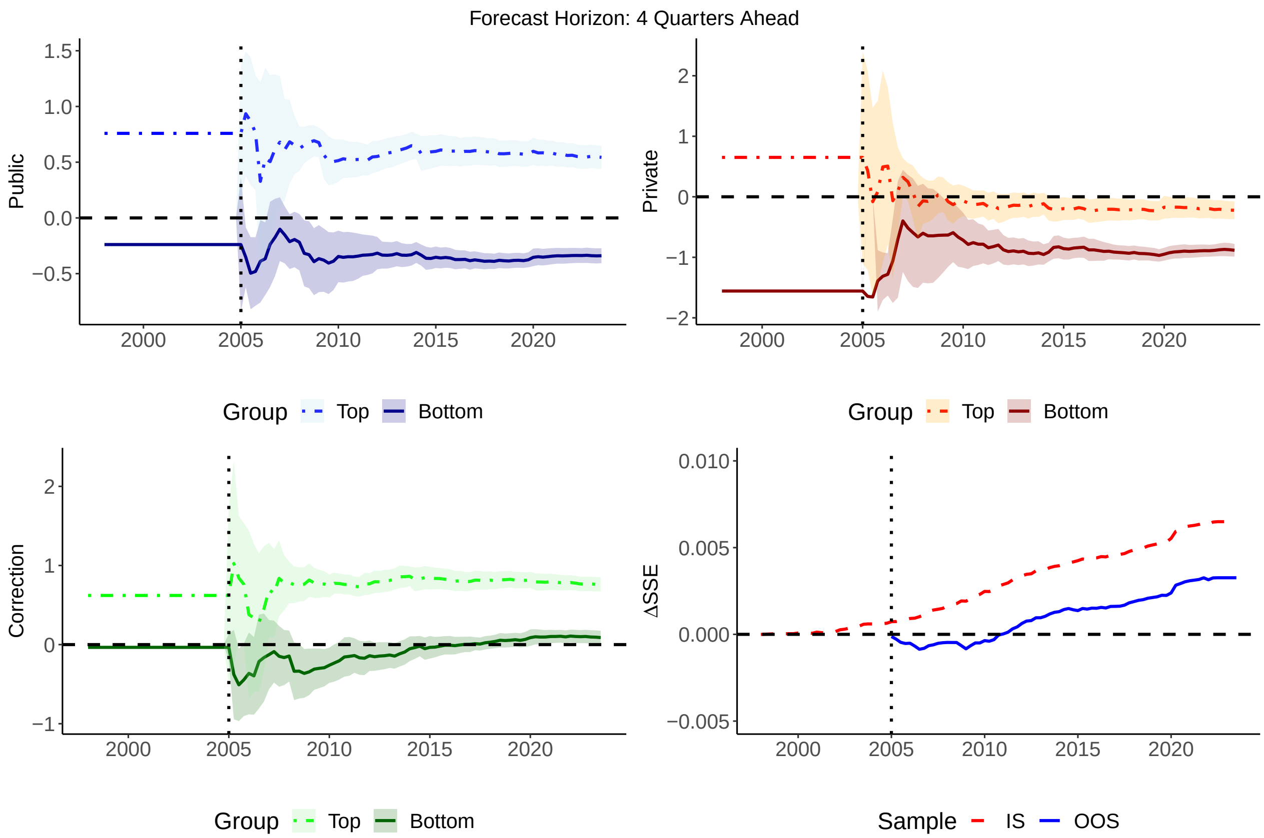

Figure O14: OOS performance of the decomposed model with cross-analyst heterogeneity

Notes: This figure supplements Figure 9 by showing the results from regressions of forecasts with horizon 4 quarters ahead.

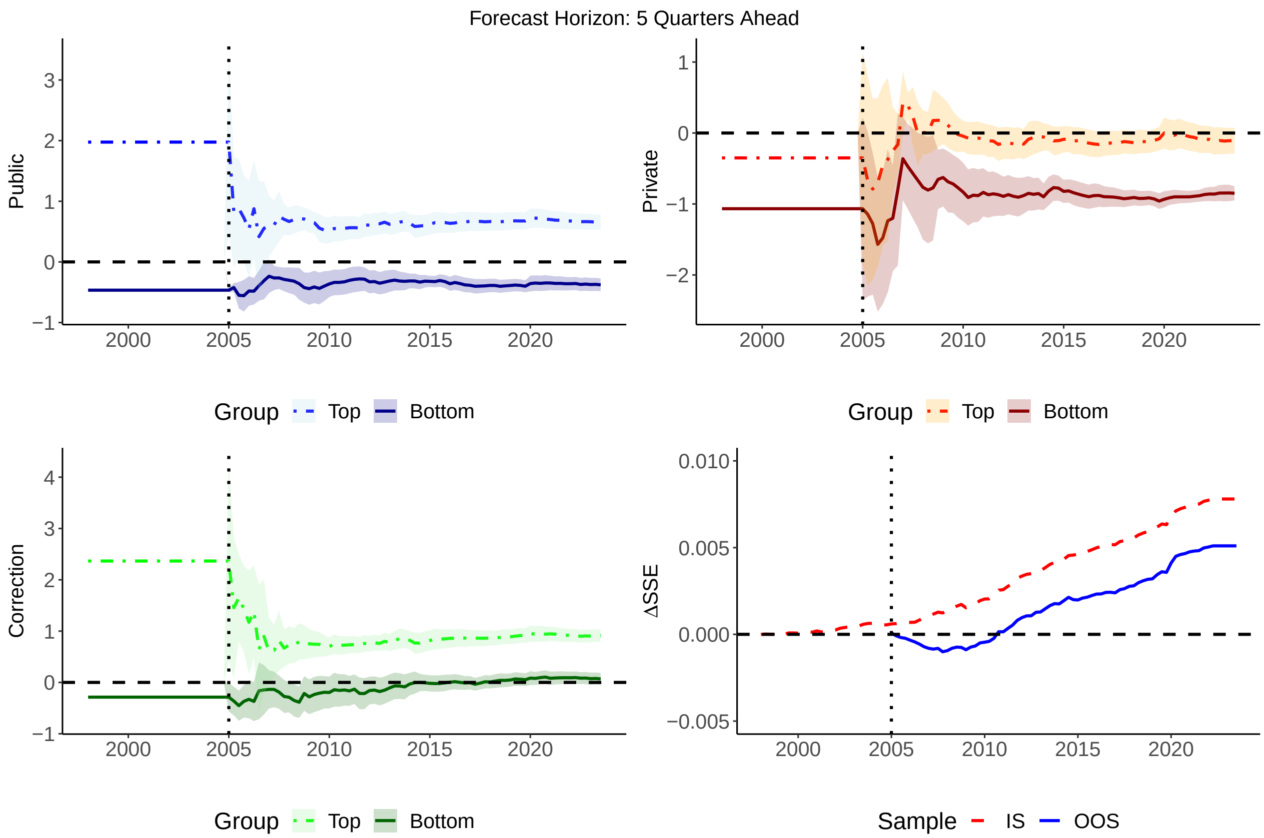

Figure O15: OOS performance of the decomposed model with cross-analyst heterogeneity

Notes: This figure supplements Figure 9 by showing the results from regressions of forecasts with horizon 5 quarters ahead.

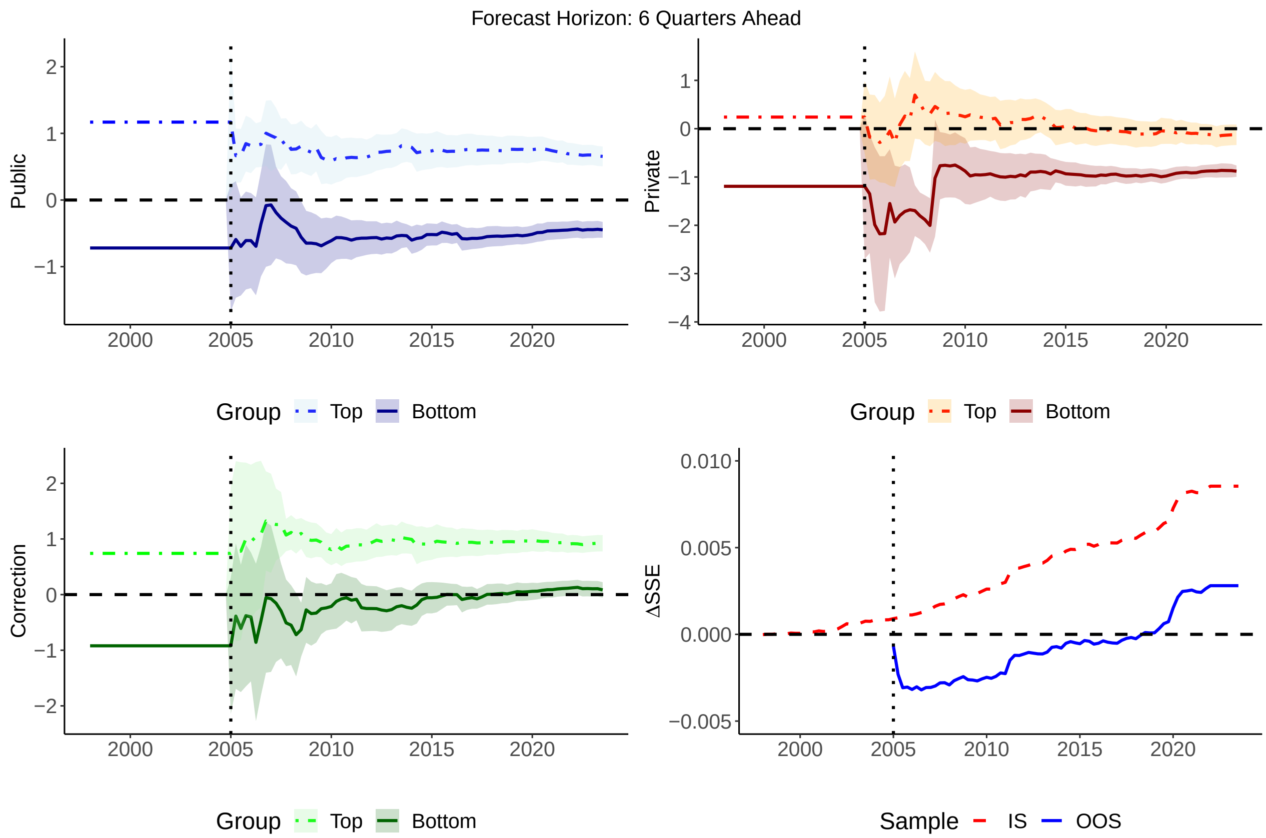

Figure O16: OOS performance of the decomposed model with cross-analyst heterogeneity

Notes: This figure supplements Figure 9 by showing the results from regressions of forecasts with horizon 6 quarters ahead.

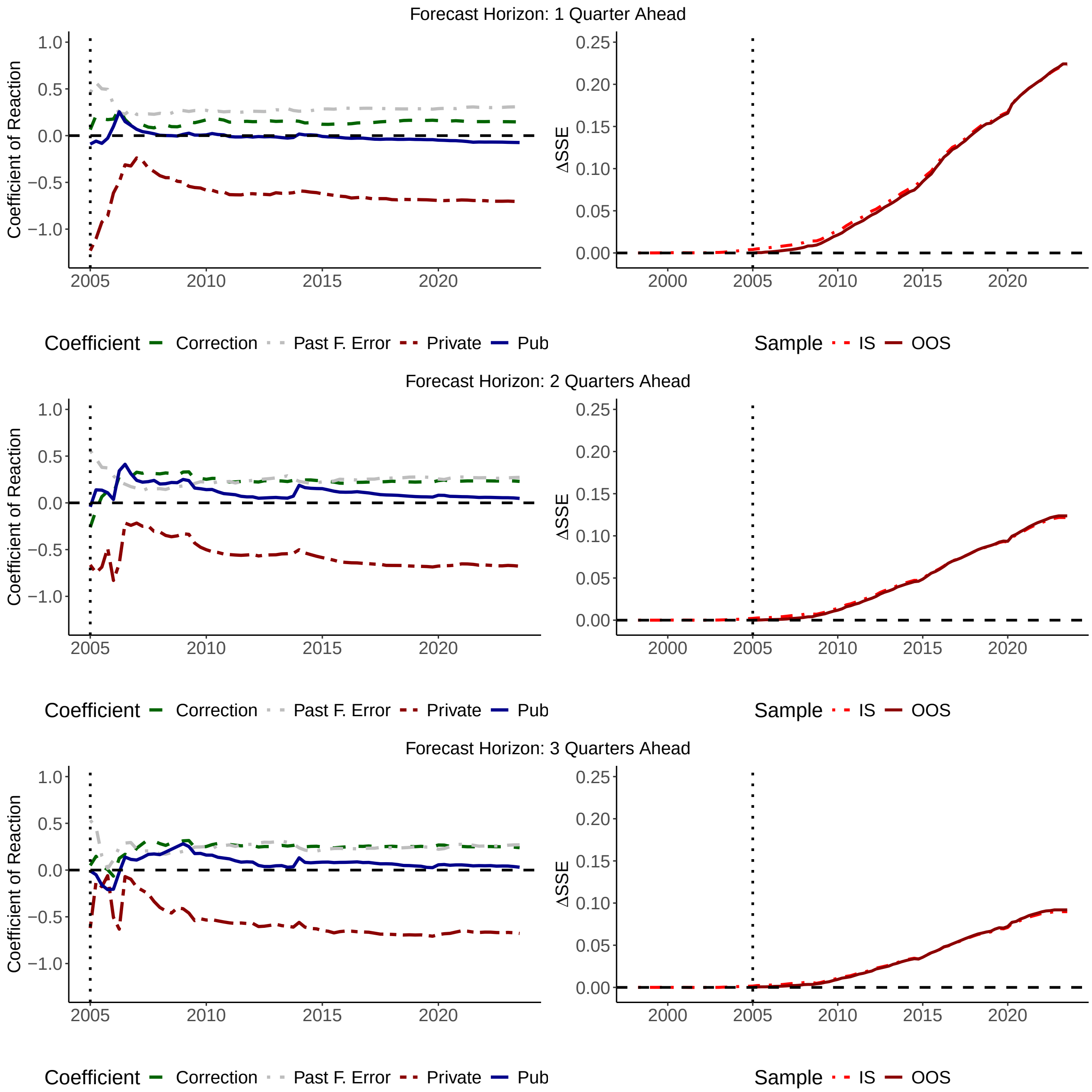

Figure O17: OOS Predictability of Forecast Errors from the Decomposed Model controlling for Past Forecast Errors

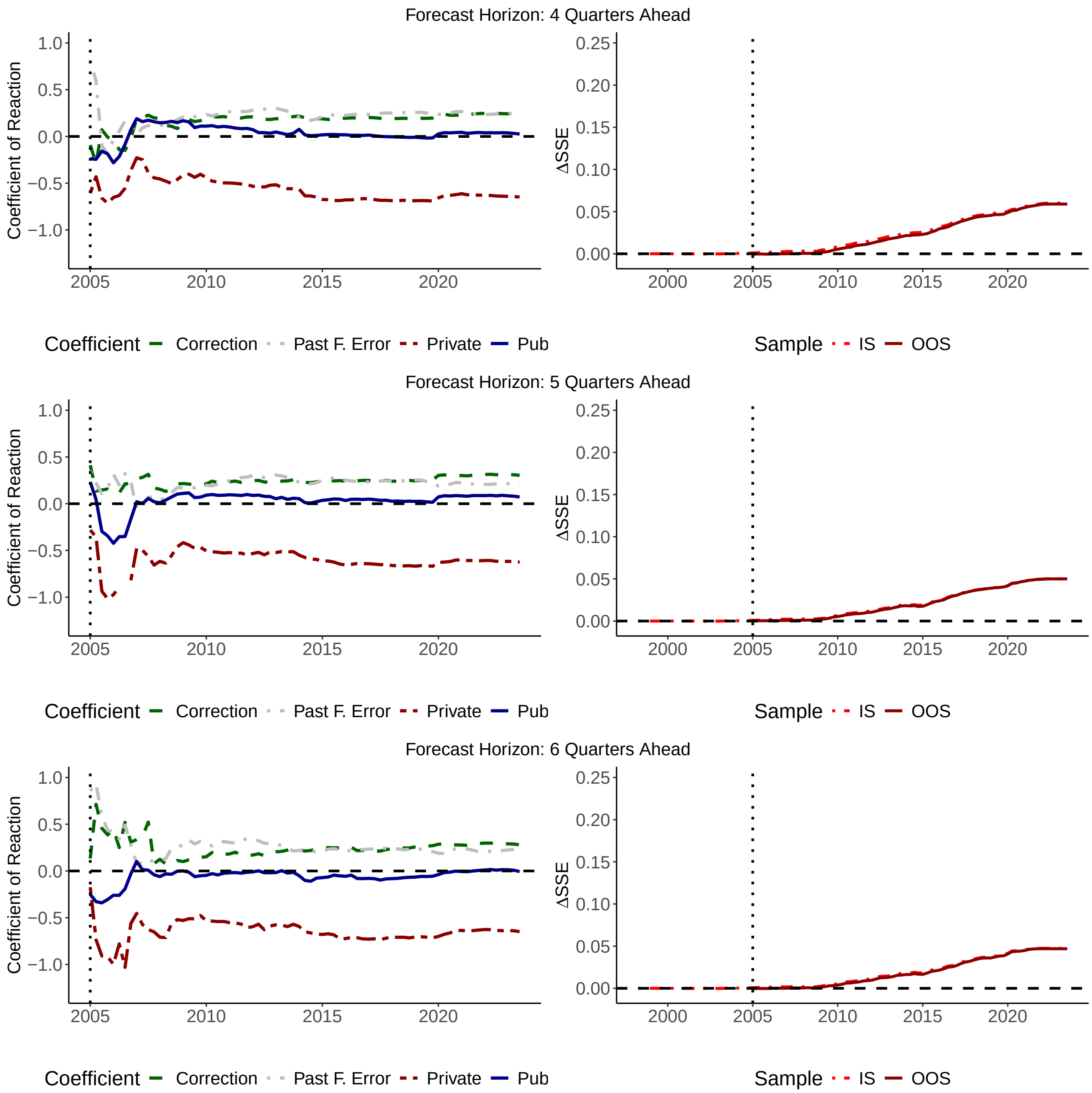

Figure O17 (cont.): OOS Predictability of Forecast Errors from the Decomposed Model controlling for Past Forecast Errors