Heterogeneous Overreaction in Expectation Formation: Evidence and Theory*

HENG CHEN
University of Hong Kong
GUANGYU PEI
Chinese University of Hong Kong
QIAN XIN
Harbin Institute of Technology
XU LI
University of Hong Kong

June 3, 2023

Abstract. Using firm-level earnings forecasts and managerial guidance data, we construct guidance surprises for analysts, i.e., differences between managerial guidance and analysts' initial forecasts. We document new evidence on expectation formation: (i) analysts overreact to guidance surprises; (ii) the overreaction is state-dependent, i.e., it is stronger for negative guidance surprises but weaker for larger surprises; and (iii) forecast revisions are neither symmetric in guidance surprises nor monotonic. We organize these facts with a model where analysts are uncertain about the quality of managerial guidance. Structural estimation reveals that a reasonable degree of ambiguity aversion is necessary to account for the documented heterogeneous overreaction.

Keywords. overreaction, expectation formation, managerial guidance, forecast revision, asymmetry, nonmonotone, ambiguity aversion

JEL Classification. C53, D83, D84

*Heng Chen: Department of Economics, The Faculty of Business and Economics, University of Hong Kong; Guangyu Pei: Department of Economics, Chinese University of Hong Kong; Qian Xin: Area of Accounting, School of Economics and Management, Harbin Institute of Technology, Shenzhen; and Xu Li: Area of Accounting, The Faculty of Business and Economics, University of Hong Kong. This research project is partly funded by the Seed Fund for Basic Research of Hong Kong University (Project No. 202011159014).

1. Introduction

The mechanisms underlying expectation formation are crucial for understanding economic decisions. While it is documented that individuals in general overreact to information (Bordalo, Gennaioli, Ma, and Shleifer 2020), there has been growing interest in the circumstances under which the overreaction is stronger or weaker and under which an underreaction can arise. In this paper, we provide new evidence that the degree of overreaction can be heterogeneous across individual forecasters, even when they receive the same information. To organize the facts, we propose a forecasting model where agents make forecasts based on noisy information (e.g., economic or financial data) and are uncertain about information quality.

To test how agents form expectations in general and how they react to new information in particular, it would be ideal to have a testing ground in which (i) the new information acquired by agents is observable and measurable, and (ii) agents' forecasts before and after receiving the new information are available. We consider an environment that is fairly close to this: financial analysts forecast the earnings of firms, firms release managerial guidance for earnings, and then analysts update their earnings forecasts. Forecast revisions are then defined to be the differences between analysts' updated forecasts after receiving managerial guidance and their initial forecasts before receiving it. That is, forecast revisions are constructed to reflect the impact of the guidance on earnings.

Using earnings forecasts data (individual analysts' EPS forecasts from the I/B/E/S Estimates) and managerial guidance data (the I/B/E/S Guidance data) from 1994 to 2017, we provide a number of findings. First, analysts' forecasts overreact to information that arrives during the time window that is constructed to encompass managerial guidance. We show that forecast revisions are negatively correlated with forecast errors, which are defined to be the differences between realized earnings and analysts' updated forecasts. This suggests that upward (downward) revisions can predict negative (positive) forecast errors, i.e., there is too much revision relative to the rational benchmark. This result is consistent with the existing findings of Bordalo, Gennaioli, Ma, and Shleifer (2020) using macroeconomic survey data.

Second, our new finding in this paper is that the overreaction is heterogeneous across analysts. We define *guidance surprises* to be the differences between the managerial guidance and analysts' initial forecasts. We construct surprises at the firm-quarter-analyst level, rank those surprises from the most negative to the most positive and then group them into deciles. Estimating the degree of overreaction in each decile subsample, we find that the overreaction is stronger when the surprises are negative; the overreaction tends to be weaker when the surprises are larger in size.

Third, we further directly explore how forecast revisions respond to guidance surprises with nonparametric estimations. We find that forecast revisions are asymmetric in surprises: forecast revisions are stronger when the surprises are negative than those when the surprises are of the same magnitude but positive. Furthermore, forecast revisions are not monotonically increasing in surprises either: when the surprises are large enough, forecast revisions decrease in surprises. Thus, the estimated relationship between forecast revisions and surprises displays a pattern of asymmetry and non-monotonicity. It is worth pointing out that the two new facts, i.e., heterogeneous overreaction and asymmetric and non-monotonic relation between forecast revisions and surprises, are consistent with each other.[1]

The new evidence on the documented heterogeneous overreaction pattern calls for a new theory, in which optimal forecasts have to be *state-dependent*. In general, forecasting models with state-independent responses would predict that forecast revisions are linear in surprises and are therefore inconsistent with the new facts documented. In particular, the extent to which forecast revisions respond to new information should vary, depending on the size and direction of surprises contained in the new information, which is a necessary condition for the overreaction pattern to be heterogeneous.

In section 3, we propose one such state-dependent forecasting model where analysts have access to both private information about the earnings of a firm (unobservable to the econometrician) and managerial guidance for earnings from the firm (observable to the econometrician). The key departures from standard forecasting models are (a) that analysts are ambiguous about the quality of the managerial guidance and (b) that they are ambiguity averse.

Given assumption (a), analysts update their beliefs about the quality of guidance based on the guidance itself and update their beliefs about earnings for any possible quality. On the one hand, the forecast revision should be large when a surprise is large. On the other hand, when a surprise is large, a Bayesian analyst would believe that its quality is likely low. When surprises are large enough, the latter force can dominate the former, which explains why forecast revisions could decrease in surprises.

We incorporate analysts' aversion to ambiguity (i.e., assumption (b)) with the smooth model of ambiguity as proposed in Klibanoff, Marinacci, and Mukerji (2005), where the degree of ambiguity aversion is finite. Given ambiguity-averse preferences, analysts wish to act in a robust fashion. In general, analysts behave as if, in their posterior beliefs, they optimally overweigh the states of the world where their expected utility is low and underweight the states where their expected utility is high. Suppose

[1]If forecast revisions are linear in surprises, then the extent of overreaction to new information cannot be heterogeneous; and if overreaction is heterogeneous in size and direction of surprises, then forecast revisions cannot be linear in surprises. This connection will be characterized in Section 6.1.

specifically that analysts consider high earnings realizations to be favorable. Then, they would subjectively "discount" the quality of favorable news, because it improves analysts' expected utility. In contrast, analysts would subjectively "overcount" the quality of unfavorable news, because it reduces analysts' expected utility. Therefore, ambiguity-averse analysts are less responsive to favorable than to unfavorable surprises, which explains the asymmetry of forecast revisions.

In section 4, we demonstrate that it is crucial to allow agents to possess a *finite* degree of ambiguity aversion to simultaneously capture both nonmonotonicity and asymmetry in the relationship between forecast revisions and surprises. Without ambiguity aversion, analysts' forecast revisions are symmetric, despite the sign of surprises. With extreme ambiguity aversion (i.e., the Wald (1950) maxmin criterion), analysts' forecast revisions are monotonic in surprises, despite the uncertainty in information quality.

In this model, to what extent analysts overreact or underreact to information, when revising their forecasts, depends critically on how analysts perceive the quality of the managerial guidance. As predicted in our model, when surprises are negative, analysts tend to infer the quality of guidance to be relatively high, which leads to a larger overreaction. When surprises are large enough, analysts tend to infer the quality of guidance to be relatively low, which leads to a milder overreaction (or even underreaction). Both predictions are qualitatively consistent with the pattern of heterogeneous overreaction found in the data.

In section 5, we estimate the model with the simulated method of moments (SMM) and quantitatively evaluate the impact of ambiguity aversion. Our estimated model can successfully predict a cross-sectional pattern of overreaction that is consistent with the data, even though it is not targeted in our estimations. Our theory underlies the role of uncertain information quality in organizing the new facts regarding expectation formation.

We stress that uncertainty in information quality is one of the keys to rationalize the observed pattern. The flip side of our theory says that once the uncertainty is very low, analysts' forecast revisions should be almost linear in guidance surprises. We show in section 5.4 that this auxiliary prediction is empirically supported.

While this paper is the first that discovers and rationalizes this set of cross-sectional patterns in the literature of expectation formation, we acknowledge that there could be other mechanisms that simultaneously contribute to the observed patterns. To highlight our theoretical contributions to the literature, we examine a number of existing theories in section 6, such as diagnostic beliefs, over-confidence, loss aversion, dynamic models, and agency theory.

Contributions While overreaction to information has been well documented and studied, we discover and rationalize heterogeneous overreactions to shocks of different properties, which is one step further from the existing literature.

We also add a new theory to the literature by explicitly studying information with uncertain quality. It is commonplace that forecasters receive noisy economic or financial data to make forecasts but are uncertain about data quality. However, the way they react to noisy data of uncertain quality has this far received little attention in the expectation formation literature. Our paper fills this gap and scrutinizes how uncertainty of this sort plays a role in expectation formation.

On the empirical front, we construct an empirical setting, which has unique advantages for uncovering how the properties of surprises (i.e., new information) affect the degree of overreaction. First, managerial guidance is observable and likely the most important part of the information flow in the time window of our intentional construction. Second, analysts have dispersed information before receiving the guidance, summarized by their initial forecasts. The two features combined imply that the same managerial guidance delivers different surprises to analysts with different initial forecasts. The variations in surprises at the analyst level enable us to explore the cross-sectional features of overreaction. Third, in contrast to studies using the Survey of Professional Forecasters, this setting is not dynamic: we utilize within-quarter variations in surprises among analysts to uncover how analysts update their forecasts. Therefore, it is cleaner for exploring cross-sectional variations in expectation formation. Finally, it has been well documented that analysts cast doubt on the quality of managerial guidance.[2] Therefore, this setting offers a natural environment to study expectation formation when the information quality is uncertain.

Literature Review Both the facts documented and the mechanisms characterized in this paper are relevant for the expectation formation literature in general and studies concerning overreactions to information in particular. The empirical part of this paper builds on a new literature that empirically explores information frictions and expectation formation. Coibion and Gorodnichenko (2015) provide a new empirical methodology in which they regress forecast errors, i.e., the difference between the realized random variable and the revised forecast of the forecaster, on forecast revisions, i.e., the difference between the revised and the initial forecast. Under the full information rational expectation (FIRE) assumption, the coefficient will be zero. Thus, a statistically significant coefficient suggests a departure from FIRE. Coibion

[2]Prior studies suggest that firm managers have various incentives to bias their forecasts either upwards or downwards, which renders the guidance doubtful to analysts, such as litigation concerns (Skinner 1994; Rogers and Stocken 2005), deterring entry (Newman and Sansing 1993; Rogers and Stocken 2005) and signaling their ability to survive and recover from financial distress (Frost 1997; Rogers and Stocken 2005).

and Gorodnichenko (2015) find that consensus forecasts of macroeconomic variables tend to underreact relative to FIRE. Applying the same approach to individual forecasts, Bordalo, Gennaioli, Ma, and Shleifer (2020) find that analysts overreact to information in general. This pattern is also discovered by Broer and Kohlhas (2022) with macroeconomic survey data. The same approach is applied to firm earnings forecast data: Bordalo, Gennaioli, Porta, and Shleifer (2019) document that an overreaction of individual analysts' forecasts is present in forecast data on firms' long-term earnings growth, and Bouchaud, Krueger, Landier, and Thesmar (2019) discover that underreaction is present in the case of short-term earnings growth.[3]

Part of our work relies on the aforementioned "FE-on-FR" approach. However, instead of focusing only on the average behavior of the whole sample, we consider differences across groups: analysts who are positively surprised vs. those who are negatively surprised and analysts who are more surprised vs. those who are less surprised. Furthermore, we provide a complementary empirical approach of "FR-on-Surprise," that directly explores the relationship between forecast revisions and observable new information and can be a useful tool for the literature. It is worthwhile to highlight that we construct a novel and nontrivial environment to study expectation formation, which is useful for other related research in this area.[4]

In an experimental setting, Afrouzi, Kwon, Landier, Ma, and Thesmar (2022) establish that the overreaction is stronger for a less persistent data generation process and stronger for longer forecast horizons. They account for the facts by allowing recent observations to have a larger influence on expectations. We focus on cross-sectional variations in overreaction and explore how the characteristics of surprises affect forecast revisions. These two works are complementary for understanding the determinants of overreaction to information.

The building blocks of our model have precedents in the literature on uncertain information quality. Both Gentzkow and Shapiro (2006) and Chen, Lu, and Suen (2016) show that Bayesian agents who are uncertain about the quality of news would rationally discount its quality when the news received is far from their priors. In those models, the direction of surprises does not matter, and there is no asymmetry. Both Ep-

[3]Other recent studies also provide evidence on the forecasts of financial market participants, such as Amromin and Sharpe (2014), Barrero (2022), Ma, Ropele, Sraer, and Thesmar (2020) and Greenwood and Shleifer (2014).

[4]We use managerial guidance to facilitate the exploration because this is among the very few kinds of information that are observable, measurable and systematically accessible to econometricians. Management earnings guidance is one of the most significant events that releases new information to the market during a quarter. For instance, Beyer, Cohen, Lys, and Walther (2010) show that the release of management earnings forecasts accounts for more than 50% of the variations in returns during a quarter, indicating that market participants pay close attention to it. It is thus of first-order importance to understand how sell-side financial analysts, as an important information intermediary, revise their beliefs on earnings projections upon managerial guidance.

stein and Schneider (2008) and Baqaee (2020) characterize the process of expectation formation when agents have an extreme ambiguity-averse preference (i.e., multiple priors) and show that belief updating is asymmetric in the contexts of asset pricing and business cycles, respectively.[5] Our work allows for a finite degree of aversion in the smooth ambiguity model by following Klibanoff, Marinacci, and Mukerji (2005) and Cerreia-Vioglio, Maccheroni, and Marinacci (2022). However, our model differs in that agents do not know the second moments of the data generating process, i.e., uncertainty in information quality.[6] Our results differ qualitatively from the two aforementioned polar cases in the literature. The empirical and quantitative exercises in this paper show that such a theoretical deviation is relevant and necessary.[7]

The theory part of this paper adds to a growing literature on expectation formation that deviates from the rational expectation benchmark. To rationalize such a deviation in the theoretical literature, one approach pursued is to relax the full-information assumption. Prominent examples include rational inattention (Sims 2003), sticky information (Mankiw and Reis 2002), higher-order uncertainty (Morris and Shin 2002; Woodford 2003; Angeletos and Lian 2016) and asymmetric attention (Mackowiak and Wiederholt 2009; Kohlhas and Walther 2021). Another approach is to introduce behavioral features. Prominent examples include diagnostic expectations (Bordalo, Gennaioli, and Shleifer 2018, Bordalo, Gennaioli, Ma, and Shleifer 2020, Bianchi, Ilut, and Saijo 2022), overconfidence (Broer and Kohlhas 2022), cognitive discounting (Gabaix 2020), level-K thinking (García-Schmidt and Woodford 2019, Farhi and Werning 2019), narrow thinking (Lian 2020), autocorrelation averaging (Wang 2020) and loss aversion (Capistrán and Timmermann 2009).[8] The most recent studies combine both, such as overextrapolation with dispersed information (Angeletos, Huo, and Sastry 2020).[9] A common feature of the aforementioned theories is that, in a linear Gaussian environments, forecast revisions are monotonically increasing in surprises and the direction of surprises does not matter. Our model differs in both aspects.

[5]In addition, Baqaee (2020) provides evidence that households' inflation expectations are more responsive to inflationary news than to disinflationary news and that the downward nominal wage rigidity can be driven by this asymmetric response to inflationary and disinflationary news.

[6]In the existing literature, it is often assumed that agents are ambiguous about the first moments and have multiple priors preferences, such as Ilut (2012), Ilut and Schneider (2014), Ilut and Saijo (2021).

[7]Models that feature constant ambiguity aversion have become common in the recent literature, and Baliga, Hanany, and Klibanoff (2013) is one such example.

[8]In section 6.2, we contrast our model with models featuring loss aversion as in Elliott and Timmermann (2008) and Elliott, Komunjer, and Timmermann (2008). We then demonstrate that under the flexible setup of loss aversion, though FR-on-Surprise relation is still monotonically increasing.

[9]Farmer, Nakamura, and Steinsson (2021) maintain the assumption that forecasters are Bayesian and show that in a dynamic environment, slow learning over the unit root long-run trend can rationalize a set of forecasting anomalies at the consensus level.

2. Evidence

2.1. Data, Sample and Timing

In this section, we explore how analysts revise their earnings forecasts upon newly received information. Our goal is to construct a scenario where the information flow is observable, measurable and accessible to the econometrician.

Toward this end, we focus on managerial guidance, which is among the very few information sources that satisfy such criteria. In financial markets, the management teams of publicly listed firms issue guidance for the earnings of the current quarter between the last quarter's and current quarter's earnings announcements. That is a crucial opportunity for firms to provide information about earnings to market participants, such as financial analysts. Because of its importance, earnings guidance often triggers analysts' forecast updates: analysts likely revise their forecasts a few days after receiving earnings guidance, i.e., on average 4 days in our sample (constructed in this section). Furthermore, it is common that firms continue to provide earnings guidance for an extensive period of time, and the discontinuation in earnings guidance is typically perceived unfavorably by the market (Chen, Matsumoto, and Rajgopal 2011). Earnings guidance includes various forms, such as point estimates and range estimates.

The Thomson Reuters I/B/E/S Guidance data provides quantitative managerial expectations, such as earnings per share, from press releases and transcripts of corporate events. The data cover managerial guidance from more than 6,000 companies in North America that can date back to as early as 1994. Furthermore, the I/B/E/S Guidance data are offered and presented on the same accounting basis as the I/B/E/S Estimates that provide individual analysts' forecast data. This makes it feasible to rigorously identify the timing of events and to compare managerial guidance and analysts' forecasts for the same firm in a certain period. Our sample construction based on the I/B/E/S Guidance and Estimates data is detailed as follows.

First, we retrieve all quarterly earnings guidance from the I/B/E/S Guidance Detail file issued for the current quarter by firm management from 1994 to 2017. The sample starts in 1994 as this is the first year when the I/B/E/S systematically collected information on managerial guidance.[10] We only keep closed-ended managerial

[10]The coverage bias in the management forecast data documented by Chuk, Matsumoto, and Miller (2013) is less of a concern in this particular setting. First, we obtain management forecast data from the I/B/E/S Guidance Detail file rather than the problematic First Call CIG database. Second, the focus of this paper is to understand how analysts update their beliefs given new information, i.e., management guidance in our setting. While the decision on the issuance of management guidance itself is also an important research question, it is not the focus of this paper. Third, the fact that we require at least one analyst issuing forecasts for a firm alleviates the concern that guidance data are more likely to be

guidance, including point and range forecasts, to quantify and compare them with analysts' forecasts. Consistent with the literature, the value of the guidance is set to equal the midpoint if it is a range forecast.

Second, given that our focus is on analysts' belief-updating process upon receiving new information from firm management, we exclude all managerial guidance bundled with earnings announcements.[11] We only consider unbundled guidance, partly because it is nearly impossible to distinguish whether a forecast revision reflects information gained from forward-looking managerial guidance or from the realized prior earnings when both of them occur simultaneously.

Third, for firm-quarters in which managers provide multiple rounds of earnings guidance (at different dates during the period from two days after the prior quarter earnings announcement date and the current quarter earnings announcement date), we only retain the latest guidance before the current quarter earnings announcement. However, our results are not sensitive to this specific choice and are qualitatively unchanged if we either keep the earliest guidance issued during a quarter or discard all firm-quarters with multiple guidance.

Fourth, we then obtain individual analysts' EPS forecasts for a firm-quarter from the I/B/E/S Estimates (the Unadjusted Detail History file) and match them with the I/B/E/S Guidance data using the same firm identifier (I/B/E/S ticker). Because earnings projections in the I/B/E/S Guidance Detail file are provided on a split-adjusted basis, we manually split-adjust both individual analysts' forecasts and managerial projections so that they are comparable with the ultimate realized earnings announced for the forecasted quarter. The realized earnings data are also obtained from the I/B/E/S Estimates. Following a standard practice in the literature, we deflate the EPS estimates by the stock price at the beginning of the quarter using data retrieved from the CRSP.[12] To avoid the small price deflator problem that may distort the distribution, we exclude observations with a stock price of less than one dollar.

Finally, in these data, the initial analyst forecasts are defined and constructed by individual analyst forecasts that are issued after the prior quarter earnings announcement date and are the most updated forecasts before the earnings guidance. The revised analyst forecasts are defined as those issued by the same set of analysts on or immediately after the earnings guidance date. For analysts who initially offer forecasts but provide no forecast revisions until the earnings announcement, we assume

collected for firms with analyst coverage. Fourth, our results are robust to starting the sample period in 1998, after which the coverage bias has been shown to be relatively small.

[11]Bundled guidance is defined as the managerial forecasts issued within 2 days around the actual earnings announcement date (Rogers and Van Buskirk 2013).

[12]We provide a robustness check for our empirical results without deflating the EPS estimates with stock prices and show that this practice does not affect our findings.

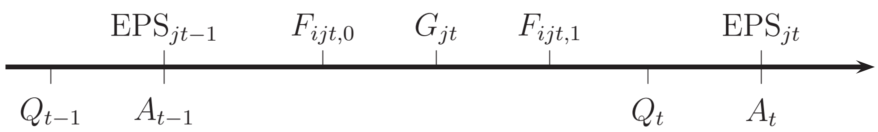


Figure 1. *Timeline. We consider managerial guidance G_t issued between A_{t-1} and A_t. If the guidance for EPS in quarter t is released on the date of A_{t-1} or within two days after A_{t-1}, then it is bundled. If the guidance is released between Q_t and A_t, it is a preannouncement. If more than one guidance is released between A_{t-1} and A_t, we choose the latest one.*

that their revised forecasts remain the same as their initial forecasts, a practice consistent with prior literature (Feng and McVay 2010; Maslar, Serfling, and Shaikh 2021).

We stress that we intentionally construct a time window where analysts' initial and updated forecasts encompass the earnings guidance of the current quarter. This construction allows us to analyze how forecasts are updated in response to information observable to the econometrician. The construction procedure can be better apprehended with the aid of Figure 1, which delineates the sequence of major events.[13] Analyst i learns firm j's EPS for quarter $t-1$ at the date of A_{t-1}, which is $\mathrm{EPS}_{j,t-1}$. Then he or she issues a forecast $F_{ijt,0}$ for firm j's EPS in quarter t. Firm j offers guidance G_{jt} for firm j's earnings in quarter t. Then, analyst i updates his or her forecast for firm j's EPS in quarter t (i.e., $F_{ijt,1}$). Quarter t ends at the date of Q_t, and firm j announces its EPS for quarter t at the date of A_t. In sum, in this setting, both initial and updated forecasts are made within the same period, after A_{t-1} and before A_t.

Our full sample consists of 110,895 pairs of individual analysts' forecasts (initial and updated forecasts) issued by 6,987 different analysts for 3,226 district firms over the period from 1994 to 2017. A summary of statistics is reported in Appendix I.A.

[13] Suppose that a typical fiscal quarter ends at Q_t, and its realized earnings are usually announced at A_t after the end of the quarter Q_t (The Securities and Exchange Commission requires public firms to file their financial statements within 45 days after the end of the fiscal quarter). Similarly, the earnings announcement date A_{t-1} for quarter $t-1$ would also happen after Q_{t-1}. In this paper, we retrieve earnings guidance that is issued by firm management on any date between A_{t-1} and A_t. Because an increasing number of firms bundle their earnings projections for quarter t with the announcement of the realized earnings for quarter $t-1$, we further require the guidance to be unbundled (as justified earlier). That is, we only consider guidances issued between two dates, i.e., A_{t-1} and A_t. Given earnings guidance G_t, we can accordingly identify the sequence of analysts' earnings forecasts for the same quarter. We define analysts' forecasts that are issued after A_{t-1} but at the latest before G_t as their initial forecast and the forecast that is issued on or after G_t but before A_t as their revised forecast. As noted above, for analysts who provide an initial forecast but do not revise, we assume that the revised forecast remains the same as the initial forecast. There are two exceptions to this general timing. First, it might be the case that G_t lies between Q_t and A_t, in which case we term the guidance a preannouncement following the convention in the literature. Second, firm management can offer more than one earnings guidance, and therefore, G_t may appear multiple times during the period. In this case, we only retain the latest guidance before A_t.

Table 1. *Forecast Error on Forecast Revision*

	Outcome Variable: Forecast Error FE_i					
	Winsorization at the 1% and 99%			Winsorization at the 2.5% and 97.5%		
	Baseline	Control	Unscaled	Baseline	Control	Unscaled
	(1)	(2)	(3)	(4)	(5)	(6)
FR_i	-0.0952***	-0.0954***	-0.0964***	-0.0926***	-0.0926***	-0.0793***
	(0.0146)	(0.0147)	(0.0124)	(0.0119)	(0.0119)	(0.0102)
Earnings of the Last Quarter		0.0023			-0.0004	
		(0.0073)			(0.0050)	
Quarter FEs	YES	YES	YES	YES	YES	YES
Analyst FEs	YES	YES	YES	YES	YES	YES
Firm FEs	YES	YES	YES	YES	YES	YES
Obs.	110,895	110,895	110,895	110,895	110,895	110,895
Adj. R-sq	0.2429	0.2429	0.2170	0.2298	0.2298	0.2236

The standard errors are clustered on firm and calendar year-quarter following Petersen (2009).*** $p<0.01$, ** $p<0.05$, * $p<0.1$

2.2. Overreaction

Our investigation of how analysts revise their forecasts starts by following the approach proposed by Bordalo, Gennaioli, Ma, and Shleifer (2020), in which they examine professional analysts' forecasts of macro variables. That is, we regress ex post analyst forecast errors on ex ante analyst forecast revisions at the individual level. To this end, we construct both forecast error FE_{ijt} and forecast revision FR_{ijt}. The former is the difference between the realized earnings per share for firm j in quarter t and the revised EPS forecast by individual analyst i for firm j in quarter t. The latter is the difference between the revised forecast after guidance and the initial forecast before guidance issued by the same analyst i for firm j in quarter t. To avoid the heterogeneity embedded in EPS across firms, both FE_{ijt} and FR_{ijt} are scaled by the stock price at the beginning of quarter t. To mitigate the impact of potential outliers, both of them are winsorized at the 1% and 99% level of their respective distributions. We estimate the following equation:

$$\text{FE}_{ijt} = b_0 + b_1 \text{FR}_{ijt} + \delta_i + \delta_j + \delta_t + \omega_{ijt}, \tag{1}$$

where we control for analyst (δ_i), firm (δ_j) and calendar year-quarter (δ_t) fixed effects. Any time-invariant analyst characteristics, time-invariant firm specific characteristics and time-series differences are absorbed and cannot explain our results. The standard errors are clustered at the firm and calendar year-quarter to adjust for both intertemporal and cross-sectional correlations, following Petersen (2009). The results from estimating equation (1) are presented in column (1) of Table 1.

We find that forecast errors are negatively correlated with forecast revisions at the individual analyst level and statistically significant at less than the 1% level. The negative coefficient indicates that analysts overreact to new information over the period that the managerial guidance is received by analysts. Despite the settings being entirely different, this result is consistent with those found in Bordalo, Gennaioli, Ma, and Shleifer (2020) and Broer and Kohlhas (2022). Both studies document the existence of overreaction to new information in analysts' forecasts of macro variables such as inflation or GDP, based on the US Survey of Professional Analysts.

We add the earnings in the last quarter ($t-1$) of firm j to the right-hand side of equation (1) and report the estimation results in column (2) of Table 1. The change in the estimated coefficient on forecast revision is negligible, and the coefficient on the earnings in the last quarter is close to zero and not significant. This suggests that the information about earnings in the last quarter is utilized by analysts to form their initial forecasts and therefore orthogonal to forecast revisions. That is the key difference from studies using Survey of Professional Forecasters (SPF) data, where initial and update forecasts are made in two separate periods.[14]

We check the robustness of our results by not scaling earnings and forecasts by stock prices, and the estimate for forecast revisions is robust, which is reported in column (3). To ensure that our results are not driven by outliers, we winsorize FE_{ijt} and FR_{ijt} at the 2.5% and 97.5% levels of their respective distributions and rerun the aforementioned exercises. Those results are reported in columns (4)-(6) of Table 1, which demonstrate the robustness of our findings.

We further perform robustness checks with different subsamples. We re-estimate equation (1) by excluding all firm-quarters with preannouncement guidance or all firm-quarters with multiple guidances or both. We report the results in Table 8, which is relegated to Appendix I.B. To ensure consistency with the results estimated locally (see sections below), we estimate equation (1) after we trim outliers from the sample and present those results in Table 9 of Appendix I.B. The estimated coefficients in the aforementioned exercises are qualitatively unchanged and only different in magnitude.

2.3. Heterogeneous Overreaction

We further explore one important feature of our empirical setting: the guidance is common for all analysts, but surprises contained in the guidance are not common across analysts because they possess heterogeneous initial forecasts. Analysts can be

[14] At each survey date, forecasters make a forecast (nowcast) for the current quarter and forecasts for the following four quarters. For example, in quarter 2005:Q3, forecasters make forecasts for 2005:Q3, 2005:Q4 and 2006:Q1-3; then, in quarter 2005:Q4, forecasters make forecasts for 2005:Q4 and 2006:Q1-4. Between the two sets of forecasts, the realized value of forecasted variable in 2005:Q3 is available to forecasters.

surprised to different extents and even in different directions. One natural question arises: Do analysts overreact differently to the same information? In fact, our data allow us to explore such heterogeneity of overreaction across analysts.

First, we construct a variable *guidance surprise* (i.e., Surprise_{ijt}) to capture the *observable* surprise in managerial guidance for individual analysts. It is defined and measured by the difference between the value of guidance (i.e., G_{jt}) issued by firm j in quarter t and analyst i's corresponding initial forecast (i.e., F_{0ijt}) for firm j in quarter t before guidance. That is, $\text{Surprise}_{ijt} \equiv G_{jt} - F_{0ijt}$.[15] For each individual analyst, the managerial guidance can be *unfavorable* or *favorable* if it falls below or exceeds the analyst's initial forecast before guidance, and the managerial guidance can be *large* or *small* if it is far from or close to the analyst's initial forecast before guidance.

Second, we remove outliers by trimming forecast errors, forecast revisions and surprises at the 2.5% and 97.5% levels of their respective distributions (to be consistent with the nonparametric estimations in the next section). We then rank surprises from the most negative to the most positive, sort them into deciles and label them from 1 to 10 according to the decile rank. To enlarge the subsample size and smooth estimates, we define a running decile window j such that (1) window j covers decile $j-1$, j, and $j+1$ if $j \neq 1$ or $j \neq 10$; (2) running decile window 1 covers deciles 1 and 2; and (3) running decile window 10 covers deciles 9 and 10.

Third, for each subsample of a running decile window, we re-estimate equation (1) (i.e., regressing forecast errors on forecast revisions). We plot the estimated coefficients and confidence intervals in Figure 2 against their window ranks. We find that analysts overreact to information in each subsample, i.e., the estimated coefficient b_1 is negative and significant. However, the degree of overreaction is not constant and is U-shaped in surprises and skewed to the left. This implies that the overreaction is stronger when the surprises are negative and the overreaction is weaker when the surprises are larger in size.

To examine whether our results are robust, we rerun the exercises with a sample where forecast errors, forecast revisions and surprises are trimmed at the 1% and 99% levels of their respective distributions. We also re-estimate equation (1) for each decile of surprises without using running windows. The patterns found are rather similar. We relegate them to Appendix I.B (see Figures 14 and 15, respectively).

In summary, on the one hand, we confirm that analysts overreact to information

[15]We stress the fact that the constructed surprise variable for managerial guidance is what is observable and accessible to the econometrician. However, it is not necessarily the real surprise for analysts, because analysts may have acquired private information, which is only observable to themselves. In this paper, we distinguish the two types of surprises both in the model setting and when making the connection between the model and the data.

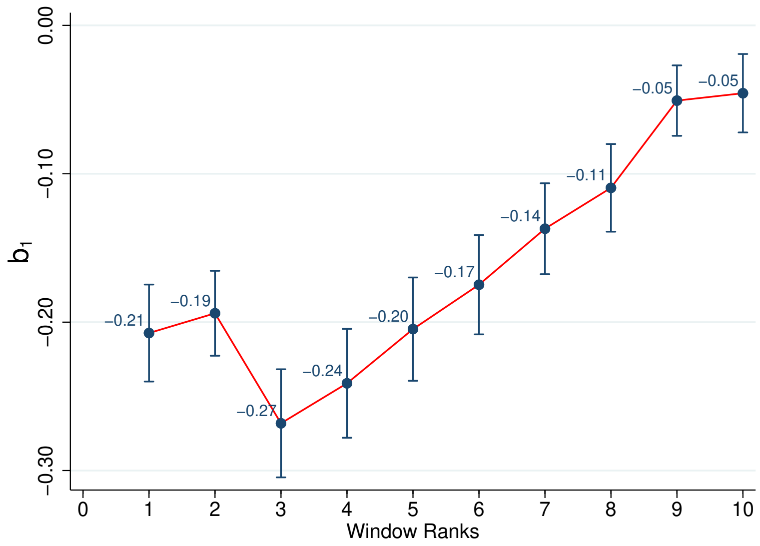


Figure 2. *Heterogeneous Overreaction. The estimated coefficients of the FE-on-FR regressions b_1 and the 95% confidence interval for each running decile window are plotted against the window rank. Running decile window j covers decile $j-1$, j, and $j+1$ if $j \neq 1$ or $j \neq 9$; running decile window 1 covers deciles 1 and 2, and running decile window 10 covers deciles 9 and 10.*

in this particular setting. Given that the forecast revisions are constructed around managerial guidance, analysts are likely to overreact to guidance surprises. On the other hand, we discover that the way that analysts react to information depends on the characteristics of the surprises that they receive, such as magnitude and favorability.

2.4. Forecast Revisions and Surprises: Mechanisms

In this section, we set out to uncover the mechanisms that underlie the heterogeneous overreaction pattern. To this end, we directly investigate the relationship between forecast revisions and surprises. Note that if forecast revisions are linear in surprises, then the degree of overreaction to new information cannot be heterogeneous (characterized in section 6.1); and if overreaction is heterogeneous in the size and direction of surprises, then forecast revisions cannot be linear in surprises.

In particular, we examine the impacts of favorability and the magnitude of guidance surprises. We begin by estimating a linear relationship between forecast revisions and surprises in guidance, controlling for the analyst, firm and quarter fixed effects, as follows,

$$\text{FR}_{ijt} = b_0 + b_1 \text{Surprise}_{ijt} + \delta_i + \delta_j + \delta_t + \omega_{ijt}, \tag{2}$$

where, as defined in the previous section, Surprise_{ijt} is the observable and measurable surprise for analyst i, contained in guidance released by firm j in quarter t. Equation (2) estimates the average effect of surprises on analysts' forecast revisions, and the

Table 2. *Forecast Revisions and Surprises in Managerial Guidance: Interactions*

	Outcome Variable: Forecast Revision FR_i					
	Winsorization at 1% and 99%			Winsorization at 2.5% and 97.5%		
	(1)	(2)	(3)	(4)	(5)	(6)
$Surprise_i$	0.1463***	0.0357*	0.4624***	0.2441***	0.1405***	0.4707***
	(0.0127)	(0.0186)	(0.0177)	(0.0129)	(0.0281)	(0.0162)
Unf		-0.0023***			-0.0014***	
		(0.0001)			(0.0001)	
$Surprise_i$ × Unf		0.1130***			0.0846***	
		(0.0231)			(0.0286)	
Large			-0.0055***			-0.0016***
			(0.0005)			(0.0003)
$Surprise_i$ × Large			-0.3666***			-0.2674***
			(0.0185)			(0.0181)
Constant	-0.0010***	0.0007***	0.0002**	-0.0005***	0.0004***	0.0001*
	(0.0001)	(0.0001)	(0.0001)	(0.0000)	(0.0001)	(0.0001)
Quarter FEs	YES	YES	YES	YES	YES	YES
Analyst FEs	YES	YES	YES	YES	YES	YES
Firm FEs	YES	YES	YES	YES	YES	YES
Obs.	110,895	110,895	110,895	110,895	110,895	110,895
Adj R-sq.	0.3943	0.4234	0.4675	0.4587	0.4723	0.4865

Notes: The standard errors are clustered on firm and calendar year-quarter following Petersen (2009). *** p<0.01, ** p<0.05, * p<0.1.

results are reported in column (1) of Table 2.[16] The significantly positive coefficient on Surprise_{ijt} suggests that individual analysts' forecast revisions are positively correlated with surprises in managerial guidances. It is intuitive that favorable surprises in guidance on average lead to upward revisions and vice versa.

However, our main interest is to explore how the positive correlation between forecast revisions and surprises varies with the favorability and magnitude of guidance surprises. To this end, we construct two dummy variables. The dummy Unf_{ijt} is equal to 1 for unfavorable guidance (i.e., Surprise_{ijt} is negative) and 0 otherwise.[17] The dummy Large_{ijt} is equal to 1 for large surprises (i.e., Surprise_{ijt} is larger or smaller than the mean value of the variable Surprise_{ijt} by one standard deviation) and 0 otherwise.

We first add the dummy Unf_{ijt} and its interaction with Surprise_{ijt} to the right-hand side of equation (2) and estimate the following regression:

$$\text{FR}_{ijt} = b_0 + b_1\text{Surprise}_{ijt} + b_2\text{Unf}_{ijt} + b_3\text{Unf}_{ijt} \times \text{Surprise}_{ijt} + \delta_i + \delta_j + \delta_t + \omega_{ijt} \quad (3)$$

[16]In the accounting and finance literature, similar specifications have been utilized, such as Hassell, Jennings, and Lasser (1988), Baginski and Hassell (1990) and Feng and McVay (2010). Hassell, Jennings, and Lasser (1988) and Baginski and Hassell (1990) document a significant positive correlation between analysts' consensus forecast revisions and the deviation of managerial guidance from the consensus forecast before guidance. Feng and McVay (2010) further investigate how this positive correlation varies with the credibility and usefulness of the guidance.

[17]Approximately 10.73% of the initial forecasts in our sample are equal to the respective managerial guidance in the corresponding quarter. We classify them as favorable to be conservative. However, our results remain qualitatively unchanged, if we exclude these confirming cases.

Such a specification allows us to compare the degrees to which analysts' forecast revisions correlate with managerial guidance surprises in each subsample. Column (2) of Table 2 shows the regression results. The positive coefficient on Surprise_{ijt} suggests that, given favorable managerial guidances, forecast revisions and surprises are still positively correlated. The coefficient on the interaction term is positive and significant, implying that the positive correlation is even more pronounced when unfavorable guidance is received.

We then add the dummy Large_{ijt} and its interaction with Surprise_{ijt} to the right-hand side of equation (2) and estimate the following regression:

$$\text{FR}_{ijt} = b_0 + b_1\text{Surprise}_{ijt} + b_2\text{Large}_{ijt} + b_3\text{Large}_{ijt} \times \text{Surprise}_{ijt} + \delta_i + \delta_j + \delta_t + \omega_{ijt} \tag{4}$$

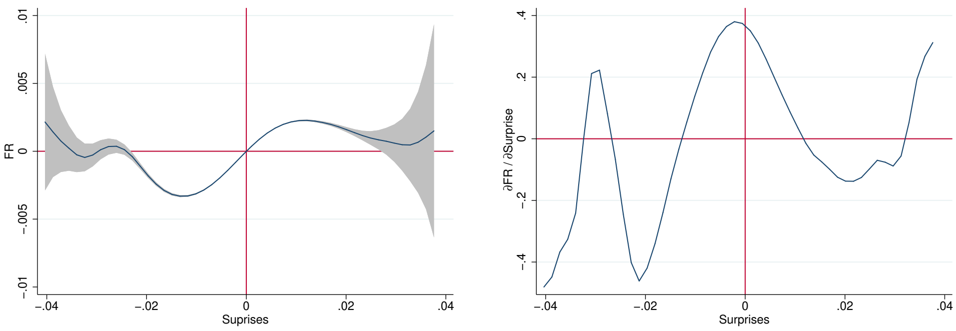
Column (3) of Table 2 shows the regression result. The coefficient on the interaction term is negative and significant, implying that the positive correlation between forecast revisions and guidance surprises is smaller when the surprises are larger. In Appendix I.B, we present regression results by using various definitions of large surprises in Table 10, which shows that our results are robust to its definition.

However, it may be argued that our results can be very sensitive to the observations in the tails of the distribution. Considering that we study the properties of "large" surprises, this concern is particularly relevant. To mitigate this, we winsorize both forecast revisions and surprises at the 2.5% and 97.5% levels of their respective distributions and re-estimate equations (2), (3) and (4). The results are displayed in columns (4), (5) and (6) of Table 2, respectively. Although the magnitude of the coefficients varies, all results are qualitatively robust.

The results in Table 2 suggest that analysts tend to react more strongly (in terms of revising their forecasts) to unfavorable surprises and that they react less strongly to large surprises. In addition, those results indicate that the underlying relationship between forecast revisions and surprises may not be linear, which motivates us to deviate from the linear regressions framework to uncover it.

To estimate the relationship in a more reliable fashion, we resort to the nonparametric estimation approach. Using the standard tool of local polynomial regression, we estimate the relationship between forecast revisions and surprises by using the Epanechnikov kernel and the third degree of the smoothing polynomial.

Because we are interested in "large" surprises and because we estimate the relationship with local polynomials, the results can be affected and biased by winsorization of the data. To alleviate this concern, we instead trim both forecast revisions and



(a) Trimming, Nonparametric estimation (b) Trimming, Derivative: marginal effect

Figure 3. *Nonparametric estimation, 5% trimming (2.5%, 97.5%). Panel (a) illustrates the relationship between forecast revisions and surprises in managerial guidances (both trimmed at 5%) that is nonparametrically estimated using the Epanechnikov kernel and the third degree of the smoothing polynomial. It is decreasing, increasing and decreasing and asymmetric around the origin. The shaded areas represent the 95% confidence intervals for the respective estimations. Panel (b) illustrates its derivatives with respect to surprises. The derivatives are negative when the surprises are large enough and positive when they are small. Forecast revisions respond more strongly to negative surprises than to positive surprises of the same magnitude.*

surprises at the 2.5% and 97.5% levels of their respective distributions and residualize them by controlling for quarter, firm and analyst fixed effects. We estimate their relationship using the local polynomial specification, and the results are presented in Figure 3(a). Forecast revisions are decreasing, increasing and decreasing in surprises and are asymmetric around the origin. Figure 3(b) illustrates its derivatives with respect to surprises. The derivatives are negative when the surprises are large enough and positive when they are small. Forecast revisions respond more strongly to negative surprises than to positive surprises of the same magnitude.

In Appendix I.B, we present various robustness checks, and the empirical findings are robust. We first present results when both forecast revisions and surprises are trimmed at the 1% and 99% levels of their respective distributions (see Figure 10). For the purpose of comparison, we also present the results by using forecast revisions and surprises that are winsorized at the 2% (1% and 99%) and 5% (2.5% and 97.5%) levels of their respective distributions. The results are all similar. In addition, we present the binscatter plots for the same set of data with various parameters. The patterns discovered with binscatter plots are consistent with those identified with local polynomial regressions.

One valid concern is that the decreasing arms of the estimated relationship might be driven by a small number of observations in the tails. Ultimately, the confidence intervals become very wide when the surprises are relatively large in magnitude. However, we find that this is not the case. For the local polynomial estimation using the trimmed data, there are 3,060 observations to the left of the trough and 7,045 observations to the right of the peak, which account for close to 10% of the total observations used in this estimation. Given the number of observations utilized, this concern is alleviated.

Another potential issue is that whether to offer earnings guidance could be strategically chosen by firms, which could affect our estimations. First, this is unlikely because firms do not make decisions about whether they disclose the earnings guidance on a quarterly basis and typically continue to provide earnings guidance for an extended period of time (Chen, Matsumoto, and Rajgopal 2011). Second, we construct a subsample in which we only include earnings forecasts conditional on firms (whose earnings are being forecasted) having to release earnings guidance for at least 12 consecutive quarters during our sample period.[18] We nonparametrically re-estimate the relationship between forecast revisions and surprises following the procedure described above. The results are presented in Figure 12 in Appendix I.B. They are rather similar to those obtained using the full sample, and the two key characteristics are even more pronounced. Therefore, the concern of strategic disclosure is inconsequential for our findings.

Our data cover the period of the 2007-2009 financial crisis, and it is not unlikely that financial market participants behaved abnormally during that period, which could affect the relationship in which we are interested. To investigate this possibility, we remove the data from 2007 to 2009, i.e., the financial crisis period, and re-estimate the relationship. The results are presented in Figure 13 in Appendix I.B, which are very similar to those obtained by using the full sample.

The facts documented in sections 2.3 and 2.4 would be puzzling if one assumed that analysts know the quality of managerial guidance with certainty. In such a case, forecast revisions would be linear in surprises, and the degree of overreaction would also be constant. Once we relax this assumption and accommodate the conjecture that the quality of information can be uncertain to analysts, those documented facts can be reasonable and consistent with each other. To account for those facts in a unifying framework, we propose a model where analysts are uncertain about the quality of information that they receive.

[18]Based on the initial full sample of management guidance, we select a quarterly management guidance for our subsample if it lies in any series of at least 12 consecutive quarters where managers provide earnings forecasts in each quarter. For example, the guidance issued in 2012Q4 is selected if it is in a series of 12 consecutive quarters from 2011Q1 to 2013Q4 with management guidance. The subsample consists of 49,116 observations with 5,601 firm-quarters. We also vary the threshold for the number of consecutive quarters, such as 8 and 16. The results are rather similar.

3. The Model

3.1. Setup

Consider a one-period static model where there exists a continuum of analysts, indexed by $i \in [0,1]$, and a firm. The firm's earnings θ are stochastic. Analyst i makes a forecast F_{0i} about the earnings at the beginning of the period and makes an updated forecast F_i at the end of the period.

Utility function. In the context of forecasting problems, we impose one restriction that analysts' optimal forecast is precisely $F^* = \theta$, conditional on analysts' information being complete (i.e., the earnings θ are known to the analysts). Any utility functions that satisfy this restriction can be approximated by a utility function $U(\cdot,\cdot)$ that is quadratic in both forecasts and earnings. In the main text, we consider one particular case among this class of quadratic utility functions, which is given by:

$$U(F,\theta) = -(F-\theta)^2 + \beta\theta, \tag{5}$$

where β is a constant. To interpret parameter β, consider the scenario where analysts have complete information. They can minimize the forecasting errors to zero, but the realized earnings may still matter for analysts in our model. The parameter $\beta > 0$ ($\beta < 0$) implies that analysts would be better (worse) off if the realized earnings θ were higher. The parameter β will be estimated and interpreted in section 5.2.[19]

This utility function is used for ease of exposition and highlighting our new mechanisms. In Online Appendix B, we present a full characterization of the model with the most general quadratic utility function of this class. We show that it is qualitatively similar and provide evidence that the additional parameters in the general case are empirically irrelevant in this setting.

Information structure. We assume that the earnings follow a normal distribution with mean 0 and variance σ_θ^2, i.e., $\theta \sim N(0,\sigma_\theta^2)$; let $\tau_\theta = 1/\sigma_\theta^2$. The distribution of earnings is known to all analysts. To have a direct mapping with the data, we allow each analyst i to be endowed with private information about the earnings before making the initial forecasts, as follows:

$$z_{0i} = \theta + \iota_i,$$

where ι_i is normally distributed with mean 0 and variance σ_z^2, i.e., $\iota_i \sim N(0,\sigma_z^2)$; let $\tau_z = 1/\sigma_z^2$. Analyst i makes forecast F_{0i} with heterogeneous information z_{0i}.

[19] Section 5.2 provides discussions on empirical evidence that analysts' utility can be dependent on earnings. In this section, we provide a characterization in which β can take any value.

Each analyst then receives a set of new information. First, analysts receive managerial guidance released by the firm, which is a noisy signal about earnings:

$$y = \theta + \eta.$$

where η is normally distributed with mean 0 and variance σ_Y^2, i.e., $\eta \sim N\left(0, \sigma_Y^2\right)$; let $\tau_Y = 1/\sigma_Y^2$. The managerial guidance is a public signal and can be accessed by the econometrician. Second, each analyst also receives a noisy private signal:

$$x_i = \theta + \varepsilon_i.$$

where ε_i is normally distributed with mean 0 and variance σ_x^2, i.e., $\varepsilon_i \sim N\left(0, \sigma_x^2\right)$; let $\tau_x = 1/\sigma_x^2$. It can be interpreted as a sufficient statistic for all the new information analyst i receives prior to making forecast F_i, except the managerial guidance.[20] Such a private signal is not observable to other analysts and thereby not observable to the econometrician. After analysts have made their updated forecasts, the earnings announcement is made, and the payoffs to analysts are realized.

The information structure in this model warrants discussion. First, in this paper, we focus on a static model without modeling the dynamics of earnings across periods. As discussed in section 2.1, analysts have perfect information about earnings in the last quarter. In fact, we show in section 2.2 that earnings in the last quarter cannot predict forecast errors in the current quarter conditional on forecast revisions and are orthogonal to forecast revisions in the data. Second, both the initial and updated forecasts in the data are made after the earnings in the last quarter are known to analysts. In this case, forecasts of the last period's earnings are not relevant in this period, conditional on the last quarter's earnings themselves. Note that the updated earnings forecasts of the last period are *not* the initial forecasts for earnings in this period.

Ambiguity-averse preferences. The key departure of this model from the existing forecasting literature is that we assume that analysts are uncertain or ambiguous about the quality of the managerial guidance or their objective precision (i.e., τ_Y). Therefore, they have to form their own subjective belief about its precision (i.e., τ_y). Such an assumption is reasonable. Analysts may not know the quality of the guidance with complete certainty because management has incentives not to release the best possible information at hand and because even the best possible estimates from the management can

[20]For example, this can include new information generated from analysts' own research or private information acquired from other sources. We also allow the analysts to have access to others' initial forecasts, either a subset of them or all of them. In the latter case, to prevent analysts from learning about the earnings, we can assume that there exists a common noise in z_{0i}, such that aggregation does not guarantee full revelation. Our qualitative and quantitative results will not be affected by this assumption.

be plagued with noise but analysts are not certain about its structure.

Specifically, we let Γ_y be the range of support for the possible precision τ_y of managerial guidance. Analysts believe that $\tau_y \in \Gamma_y$ and possess some prior belief over Γ_y, whose density distribution is given by $p\left(\tau_y\right)$. We say that one particular τ_y represents a *model* that generates the managerial guidance y.

Furthermore, we assume that analysts dislike uncertainty in the quality of the managerial guidance or are ambiguity averse. In this model, we capture such a preference of analysts by using the *smooth model of ambiguity* as proposed in Klibanoff, Marinacci, and Mukerji (2005). That is, analyst i maximizes the objective function:

$$\int_{\Gamma_y} \phi\left(\mathbb{E}^{\tau_y}\left[U\left(F_i,\theta\right)|z_{0i},x_i,y\right]\right) p\left(\tau_y|z_{0i},x_i,y\right) \mathrm{d}\tau_y, \tag{6}$$

where $\phi\left(\cdot\right)$ is some increasing, concave and twice continuously differentiable function. In addition, $\mathbb{E}^{\tau_y}\left[U\left(F_i,\theta\right)|z_{0i},x_i,y\right]$ denotes the mathematical expectation conditional on analyst i's information set $\left(z_{0i},x_i,y\right)$ for a particular model τ_y (or a certain precision of managerial guidance). In what follows, we use $\mathbb{E}_i^{\tau_y}\left[U\left(F_i,\theta\right)\right]$ to denote the expected utility of analyst i, unless it causes confusion. The density of the posterior belief over possible models is assumed to be Bayesian and denoted by $p\left(\tau_y|z_{0i},x_i,y\right)$.

The curvature of function $\phi\left(\cdot\right)$ captures an aversion to mean-preserving spreads in $\mathbb{E}_i^{\tau_y}$ induced by ambiguity in τ_y.[21] The more concave the function $\phi\left(\cdot\right)$ is, the stronger the ambiguity aversion. In other words, it characterizes analysts' taste for ambiguity. In this paper, we consider a function $\phi\left(\cdot\right)$ that features constant absolute ambiguity aversion (CAAA) following Cerreia-Vioglio, Maccheroni, and Marinacci (2022) throughout:

$$\phi\left(t\right) = -\frac{1}{\lambda}e^{-\lambda t}, \tag{7}$$

where $\lambda \geq 0$ measures the degree of ambiguity aversion. Two special cases are nested. When $\lambda = 0$ and $\phi\left(\cdot\right)$ is linear, this corresponds to the case where analysts are ambiguity neutral or fully Bayesian. When $\lambda \to +\infty$, this corresponds to the case where analysts' aversion to ambiguity is infinite, which is the classic Wald (1950) maxmin criterion.[22]

[21] Ambiguity aversion differs from risk aversion, which is implicitly captured by $U\left(F_i,\theta\right)$. In this model, it is the aversion to ambiguity rather than the aversion to risk that drives our results.

[22] The model with extreme ambiguity aversion is a special case of the multiple priors preference proposed by Gilboa and Schmeidler (1989), where the priori set of priors include all Dirac measures of each model.

3.2. Noisy Information Benchmark: A Special Case

Our framework is a generalized version of the standard forecasting problem in which analysts possess noisy information and minimize the mean-squared error of their forecasts of the random variable. In other words, the noisy information benchmark is a special case of our model when agents are ambiguity neutral (i.e., $\lambda = 0$) and there exists no uncertainty in information quality (i.e., Γ_y is singleton).[23] In this section, we characterize such a special case and illustrate why it fails to account for the empirical patterns documented in section 2.3 and 2.4 and why deviations from this benchmark are necessary.

With noisy information expectations, the optimal initial and updated forecasts are such that

$$\mathrm{F}_{0i}^{\mathrm{NI}} = \mathbb{E}\left[\theta | z_{0i}\right]; \qquad \mathrm{F}_{i}^{\mathrm{NI}} = \mathbb{E}\left[\theta | z_{0i}, x_i, y\right],$$

where $\mathbb{E}\left[\theta|\mathcal{I}_i\right]$ denotes the conditional expectations (i.e., Bayesian posterior). The relationship between $\mathrm{F}_{0i}^{\mathrm{NI}}$ and $\mathrm{F}_{i}^{\mathrm{NI}}$ is therefore given by:

$$\mathrm{F}_{0i}^{\mathrm{NI}} = \left(1 - \kappa_x - \kappa_y\right) \mathrm{F}_{0i}^{\mathrm{NI}} + \kappa_x x_i + \kappa^{\mathrm{RE}} y,$$

where κ_x and κ^{RE} are the relevant weights assigned to the private and public information:

$$\kappa_x \equiv \frac{\tau_x}{\tau_\theta + \tau_z + \tau_x + \tau_Y} > 0; \qquad \kappa^{\mathrm{RE}} \equiv \frac{\tau_Y}{\tau_\theta + \tau_z + \tau_x + \tau_Y} > 0. \tag{8}$$

Therefore, the relevant forecast revision is given by

$$\mathrm{FR}_{i}^{\mathrm{NI}} \equiv \mathrm{F}_{i}^{\mathrm{NI}} - \mathrm{F}_{0i}^{\mathrm{NI}} = \kappa^{\mathrm{RE}} \left(y_i - \mathrm{F}_{0i}^{\mathrm{NI}}\right) + \kappa_x \left(x_i - \mathrm{F}_{0i}^{\mathrm{NI}}\right), \tag{9}$$

and forecast error is given by

$$\mathrm{FE}_{i}^{\mathrm{NI}} \equiv \theta - \mathrm{F}_{i}^{\mathrm{NI}} = \kappa_\theta \theta + \kappa_z \iota_i + \kappa_x \epsilon_i + \kappa^{\mathrm{RE}} \eta, \tag{10}$$

where $\kappa_\theta \equiv \frac{\tau_\theta}{\tau_\theta+\tau_z+\tau_x+\tau_y} > 0$ and $\kappa_z \equiv \frac{\tau_z}{\tau_\theta+\tau_z+\tau_x+\tau_y} > 0$.

Lemma 1 (FR-on-Surprise and FE-on-FR). *In the noisy expectation benchmark, forecast revisions are linear in guidance surprises and uncorrelated with forecast errors,*

$$\mathbb{COV}\left(\mathrm{FE}_{i}^{\mathrm{NI}}, \mathrm{FR}_{i}^{\mathrm{NI}}\right) = 0.$$

[23] In the noisy information benchmark, the parameter β in Equation (5) plays no role at all. However, it is important for the optimal forecasts when agents have ambiguity averse preferences.

Observe that the term $(y - \mathrm{F}_{0i}^{\mathrm{NI}})$ in equation (9) is the theory counterpart of managerial guidance surprises in our empirical exercise. Equation (9) predicts that forecast revisions should be linear in guidance surprises and the unobservable surprises contained in the private information $(x_i - \mathrm{F}_{0i}^{\mathrm{NI}})$ is the white noise to the FR-on-Surprise relation. However, this prediction contradicts the non-monotone and asymmetric relationship documented in section 2.4.

Further, using Equations (9) and (10), it is evident that forecast revisions and forecast errors are uncorrelated. It then predicts that the estimated coefficient in the FE-on-FR regression should be 0, i.e., no over-reaction at the individual level. This prediction contradicts evidence that analysts overreact to new information (documented in section 2.2) and that such overreaction varies in a non-monotonic and asymmetric fashion (documented in section 2.3).

The key to the failure that the noisy information benchmark cannot capture the empirical patterns, is that the optimal forecasting rule is state-independent and determined by constant signal-to-noise ratios. That is, the weight κ^{RE} assigned to the public signal (i.e., managerial guidance in this context) is constant and independent of the realization of the public signal. However, evidence suggests that the weight should vary depending on the realization of public signal in a particular way: the weight should be larger when the surprise is negative than when it is positive but of the same magnitude; and the weight should be negative (instead of positive) when surprises are large enough. In the following section, we demonstrate that our framework, featuring the ambiguous information quality and ambiguity aversion towards uncertainty, can generate a state-dependent forecasting rule that is consistent with data.

3.3. Equilibrium Characterization

In this section, we turn to the characterization of analysts' optimal forecasts. The initial forecast of each analyst F_{0i}^* is derived by Bayes' rule:

$$F_{0i}^* = \frac{\tau_z}{\tau_z + \tau_\theta} z_{0i}.$$

To choose the optimal updated forecast F_i^* after obtaining a new set of information, analysts maximize the objective in equation (6). That is, the optimal forecast F_i^* is such that the first-order condition holds:

$$F_i = \int_{\Gamma_y} \left(\frac{\tau_z z_{0i} + \tau_x x_i + \tau_y y}{\tau_\theta + \tau_z + \tau_x + \tau_y} \right) \tilde{p} \left(\tau_y | z_{0i}, x_i, y; F_i \right) \mathrm{d}\tau_y, \tag{11}$$

where the distorted posterior belief $\tilde{p}$ is such that

$$\tilde{p}\left(\tau_y|z_{0i},x_i,y;F_i\right) \propto \underbrace{\phi'\left(\mathbb{E}_i^{\tau_y}\left[U\left(F_i,\theta\right)\right]\right)}_{\text{Pessimistic Distortion}} \underbrace{p\left(z_{0i},x_i,y|\tau_y\right) p\left(\tau_y\right)}_{\text{Bayesian Kernel}}. \tag{12}$$

The term with the combined fraction in equation (11) captures the posterior mean of the random variable θ for a particular model τ_y, where the weights assigned to observations (z_{0i}, x_i, y) are dictated by Bayes' rule.

The distribution of τ_y is updated by following equation (12). When analysts are ambiguity neutral (i.e., $\lambda = 0$), $\phi'(\cdot)$ is constant and the posterior distribution of τ_y simply follows Bayes' rule. When analysts are ambiguity averse (i.e., $\lambda > 0$), the posterior distribution of τ_y is distorted by their pessimistic attitude: its density is reweighted by the term $\phi'\left(\mathbb{E}_i^{\tau_y}\left[U\left(F_i,\theta\right)\right]\right)$.

To understand such pessimism, consider analyst i who obtains observations (z_{0i}, x_i, y) and contemplates releasing a forecast F_i. She views model τ_y as the more likely model if she is worse off under such a model. That is, a model with τ_y that generates a lower expected utility for analyst i is given a higher weight in her distorted posterior belief. Recall that $\phi'(\cdot) > 0$ and $\phi''(\cdot) < 0$. Consequently, the posterior belief $\tilde{p}\left(\tau_y|z_{0i},x_i,y;F_i\right)$ depends on her forecast F_i. Such a dependence is the key difference from the standard forecasting problems.

To facilitate the subsequent analysis and characterize the pessimism, we represent the first-order condition by orthogonalizing the information set, which has a natural interpretation. Analyst i who receives x_i, updates her belief about θ, and then her posterior belief will be:

$$X_i = F_{0i} + \frac{\tau_x}{(\tau_\theta + \tau_z + \tau_x)}\left(x_i - F_{0i}\right).$$

The analyst next receives the managerial guidance y, and the *surprise* for analyst i is denoted by $s_i \equiv y - X_i$, i.e., the difference between the guidance y and the analyst's posterior belief X_i. The optimality condition of equation (11) is represented by:

$$F_i = X_i + \kappa\left(X_i, s_i, F_i\right) \cdot s_i, \tag{13}$$

where

$$\kappa\left(X_i, s_i, F_i\right) \equiv \left[\int_{\Gamma_y}\left(\frac{\tau_y}{\tau_\theta + \tau_z + \tau_x + \tau_y}\right)\tilde{p}\left(\tau_y|X_i, s_i; F_i\right) \mathrm{d}\tau_y\right], \tag{14}$$

and the distorted posterior belief is such that

$$\tilde{p}\left(\tau_y | X_i, s_i; F_i\right) \equiv \tilde{p}\left(\tau_y | z_{0i}, \frac{\tau_\theta + \tau_z + \tau_x}{\tau_x}\left(X_i - F_{0i}^*\right) + F_{0i}^*, s_i + X_i; F_i\right). \quad (15)$$

For any particular model τ_y, the optimal response to the surprise s_i is $\frac{\tau_y}{\tau_\theta + \tau_z + \tau_x + \tau_y}$, which is dictated by Bayes' rule and increasing in τ_y (the quality of managerial guidance). The response to the surprise (represented by κ) is a weighted average over the model space by using the distorted distribution $\tilde{p}\left(\tau_y | X_i, s_i; F_i\right)$, and therefore it is bounded between 0 and 1. In this representation, the pessimistic preference of analysts is specifically captured by the following lemma.

Lemma 2 (Pessimism). *Consider any $F_i' > F_i$ and the likelihood ratio*

$$L\left(\tau_y\right) \equiv \frac{\tilde{p}\left(\tau_y | X_i, s_i; F_i'\right)}{\tilde{p}\left(\tau_y | X_i, s_i; F_i\right)}.$$

If the surprise s_i is positive, $L\left(\tau_y\right)$ decreases in τ_y; if it is negative, $L\left(\tau_y\right)$ increases in τ_y.

All proofs are collected in Appendix B. Suppose that the surprise s_i is positive. An analyst i who contemplates a higher forecast F_i' would consider the positive surprise to be less likely to be informative and assign a lower probability density for models with a high τ_y in her distorted belief $\tilde{p}$. Therefore, κ is decreasing in F_i. In contrast, suppose that the surprise s_i is negative. An analyst i who contemplates a higher forecast would consider the negative surprise to be more likely to be informative and therefore assign a higher probability density to models with high τ_y in her distorted belief. Therefore, κ is increasing in F_i.

As implied by Lemma 2, the right-hand side of equation (13) always decreases in F_i. The optimal forecast F_i^* is the fixed point of equation (13). The following proposition summarizes the equilibrium existence and uniqueness of the forecasting problem.

Proposition 1 (Existence and Uniqueness). *If analysts are ambiguity averse ($\lambda > 0$), there always exists a unique optimal forecast $F_i^*\left(X_i, s_i\right)$ that satisfies (13) and a unique optimal response $\kappa^*\left(s_i\right) \equiv \kappa\left(X_i, s_i, F_i^*\right)$ associated with it.*

An interesting special case is nested in this framework: if analysts are ambiguity neutral, there is no dependence of analyst i's posterior belief $\tilde{p}$ on F_i. Bayes' rule dictates that the posterior distribution of τ_y only depends on the magnitude of the surprise, but not its sign. Therefore, the response to surprises in managerial guidance should always be symmetric.

4. Equilibrium Analysis

This section presents a set of equilibrium analyses corresponding to the empirical facts documented in section 2. We demonstrate that the two basic model mechanisms (uncertainty in quality and aversion to uncertainty) and their interaction can help account for those empirical patterns.

4.1. Asymmetry

We first characterize the impacts of ambiguity aversion on analysts' asymmetric responses to negative and positive surprises in managerial guidance. To state this formally, let a pair of surprises be $\left(s_i^-, s_i^+\right)$, such that $s_i^- < 0 < s_i^+$ and $s_i^- + s_i^+ = 0$.

Proposition 2. *If analysts are ambiguity averse, forecast revisions in response to surprises are asymmetric. Specifically,*

$$\left(\kappa^*\left(s_i^-\right) - \kappa^*\left(s_i^+\right)\right)\beta \geq 0,$$

where the equality holds if and only if $\beta = 0$.

To illustrate this, consider the case where analysts are better off when the earnings realization is high (i.e., $\beta > 0$). That is, analysts consider the news that suggests higher realizations of earnings to be favorable.

Proposition 2 states that analysts would always be less responsive to positive surprises (i.e., s_i^+, favorable news) than to negative surprises (i.e., s_i^-, unfavorable news). The mechanism is as follows. In this model, analysts are uncertain about the quality of the information source and, therefore, need to assess its quality based on the news itself. Given that favorable news improves analyst i's expected utility, she would behave with more caution (due to her ambiguity-averse preferences) and "discount" the quality of favorable news. Conversely, given that negative surprises or unfavorable news reduce her expected utility, she would "over-count" its quality, i.e., assign a high probability density to models with high quality τ_y. Therefore, analyst i responds to negative surprises to a larger extent than to positive surprises of the same magnitude, that is, $\kappa^*\left(s_i^-\right) > \kappa^*\left(s_i^+\right)$.

4.2. Nonmonotonicity

Next, we show that the model also features a nonmonotonic relationship between forecast revisions and surprises. Two key take-away messages are as follows. First, the nonmonotonicity does not rely on ambiguity aversion but instead on ambiguity (uncertainty) in quality. Second, in fact, the nonmonotonicity disappears when the degree of ambiguity aversion becomes extreme. Proposition 3 formalizes the former, and proposition 4 characterizes the latter. To simultaneously capture both nonmonotonicity and asymmetry, neither ambiguity-neutral preferences nor extreme ambiguity aversion is feasible.

Proposition 3. *If analysts are ambiguity neutral ($\lambda = 0$), the optimal forecast revision $F_i^* - X_i$ increases in s_i conditional on surprise s_i being small in magnitude and decreases in s_i conditional on surprise s_i being sufficiently large in magnitude. The forecast revision at the individual level $F_i^* - X_i$ is always symmetric around the origin.*

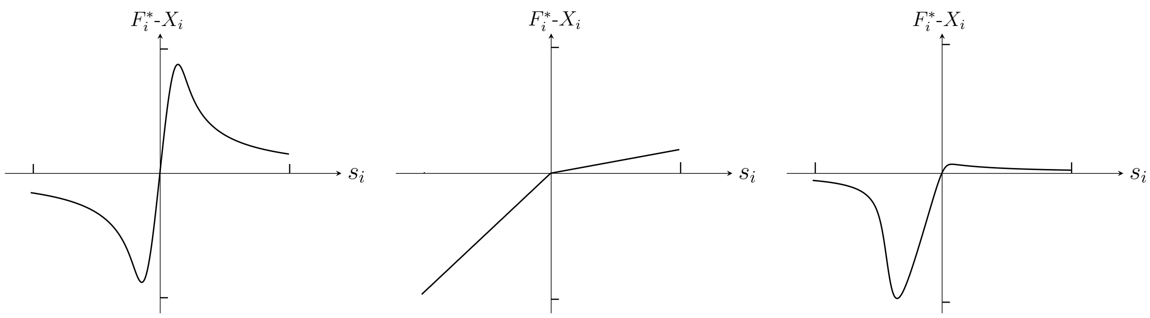
Given that the quality of guidance is uncertain, analyst i updates her belief through two mechanisms. First, for any given quality τ_y, analyst i updates her belief about the earnings upon receiving the guidance. This mechanism dictates that positive (negative) surprises raise (suppress) forecasts. Second, she also updates her belief about the distribution of quality. When the surprise is large, Bayesian analysts will assign a higher probability density to low qualities. That is, they tend to believe that large surprises are of low quality. Crudely, this is because low-quality information sources would have fatter tails and be more likely to generate large surprises. In other words, the posterior distribution of information quality given a small surprise first-order stochastically dominates the posterior distribution given a large surprise. Therefore, this mechanism implies that forecast revisions can be less responsive to surprises when they are larger.

For small enough surprises, the second mechanism (i.e., updating the distribution of quality) is less consequential, and therefore forecast revisions increase in surprises. For large enough surprises, the second mechanism dominates the first, and, as a result, forecast revisions decrease in surprises. Figure 4(a) illustrates this pattern that forecast revisions decrease and increase and then decrease in surprises. The symmetry is trivial given that analysts are Bayesian.

Now, we turn to the other polar cases: extreme ambiguity aversion ($\lambda \to +\infty$) or the classic max-min criterion.

Proposition 4. *If analysts have extreme degree of ambiguity aversion ($\lambda \to +\infty$), the optimal forecast revision $F_i^* - X_i$ is increasing in surprise s_i.*

When surprises are relatively small in magnitude, the Bayesian mechanism dictates that forecast revisions increase in surprises (Proposition 3). Furthermore, the ambiguity aversion mechanism also dictates an increasing relationship. To see this and ease exposition, we assume that $X_i = 0$. Then, analyst i tends to believe that negative surprises are of higher (lower) quality than positive surprises of the same magnitude if $\beta > 0$ ($\beta < 0$). Given that the ambiguity aversion is extreme, analysts believe that the



(a) Ambiguity Neutral $\lambda = 0$ (b) Max-min $\lambda \to +\infty$ (c) Ambiguity Aversion $\lambda > 0$

Figure 4. *Monotonicity and the degree of ambiguity aversion. Panel (a) illustrates the case where analysts are ambiguity neutral. Forecast revisions are decreasing, increasing and decreasing in surprises. Panel (b) illustrates the case where analysts have extreme degree of ambiguity aversion ($\lambda \to +\infty$). Note that $X_i = 0$ and $\beta > 0$. Forecast revisions are increasing in surprises and asymmetric. Panel (c) illustrates the case where analysts' ambiguity aversion is moderate. Both asymmetry and nonmonotonicity are present.*

quality of negative news is of the highest possible value and that of positive news is of the lowest possible value if $\beta > 0$ and vice versa. Figure 4(b) illustrates the case where $X_i = 0$, $\beta > 0$ and $\lambda \to +\infty$. In this case, analyst i with $X_i = 0$ believes that negative surprises are of the highest quality and positive surprises are of the lowest quality. Therefore, forecast revisions increase in surprises with a flatter slope when surprises are positive and with a steeper slope when surprises are negative.

When surprises are very large in magnitude, the Bayesian mechanism dictates that forecast revisions decrease in surprises (Proposition 3). However, this is dominated by the impact of extreme ambiguity aversion. Therefore, forecast revisions always increase in surprises, despite the sign of β.

In summary, the contrast of the two polar cases reveals (i) that ambiguity in guidance quality gives rise to nonmonotonicity in surprises and (ii) that aversion to such ambiguity leads to asymmetric responses to negative and positive surprises. Our model of finite ambiguity aversion lies in between. Figure 4(c) illustrates the relationship between forecast revisions and surprises when the degree of ambiguity aversion is moderate. The optimal forecast revision is not monotonically increasing, which resembles the case of ambiguity neutrality. Nevertheless, it is also asymmetric, which resembles the case of extreme ambiguity aversion.

4.3. Overreaction and Ambiguity Aversion

Do ambiguity-averse preferences contribute to analysts' overreaction to information in our model and its asymmetrical and nonmonotonic pattern documented in section 2.3? We take two steps to analyze this question. First, we study a special case of the model where analysts are ambiguity neutral ($\lambda = 0$) and show that analysts can either overreact or underreact to information in this model given the uncertainty in guidance quality. It predicts heterogenous degrees of overreaction across analysts, but the pattern is symmetric. Second, we illustrate how ambiguity aversion can amplify the overreaction to information and skew such a distribution in a negative direction.

We start our investigation by constructing a theoretical counterpart of the FE-on-FR coefficient, i.e., b_1 in equation (1). Specifically, let

$$\hat{b}_1 \equiv \frac{\text{Cov}\left(\text{FE}_i, \text{FR}_i\right)}{\text{Var}\left(\text{FR}_i\right)}.$$

To construct a benchmark for overreaction and underreaction, recall the rational expectation case where analysts know the actual quality of guidance τ_Y (discussed on page 24) and the analyst's response to surprises κ^{RE} is characterized by Bayes' rule (see Equation (8)). In this case, there is neither any overreaction nor any underreaction to information.

Proposition 5. *If analysts are ambiguity neutral (i.e., $\lambda = 0$), analysts may on average either over- or underreact to information, depending on their prior beliefs $p\left(s_i\right)$. That is,*

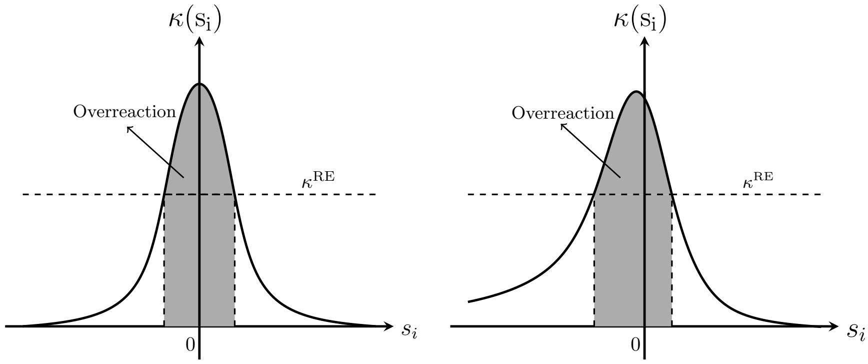
$$\text{sgn}\left\{\hat{b}_1\right\} = \text{sgn}\left\{\kappa^{\text{RE}} - \hat{\mathbb{E}}\left[\kappa\left(s_i\right)\right]\right\}.$$

where $\hat{\mathbb{E}}$ is an expectation operator under the adjusted belief $\hat{p}\left(s_i\right)$,

$$\hat{\mathbb{E}}\left[\kappa\left(s_i\right)\right] \equiv \int_{\mathbb{R}} \kappa\left(s_i\right) \hat{p}\left(s_i\right) \mathrm{d}s_i; \qquad \hat{p}\left(s_i\right) \propto \Omega(s_i) p\left(s_i\right); \qquad \Omega(s_i) \equiv \frac{\kappa\left(s_i\right) s_i^2}{\mathbb{E}\left[\kappa\left(s_i\right) s_i^2\right]}.$$

The term $\Omega(s_i)$ is an adjusted term for the transformed belief. If the average response of analysts (i.e., $\hat{\mathbb{E}}\left[\kappa\left(s_i\right)\right]$) is higher than the rational expectation benchmark κ^{RE}, analysts appear to be overreacting to information (i.e., $\hat{b}_1 < 0$); if it is lower than κ^{RE}, analysts appear to be underreacting to information (i.e., $\hat{b}_1 > 0$).

A special case is nested in this proposition. Suppose that analysts believe the quality is one particular τ_y (different from τ_Y). Then, their response to surprises would be constant and not depend on surprises, and therefore the average response is such that $\hat{\mathbb{E}}\left[\kappa\left(s_i\right)\right] = \kappa$, where $\kappa = \tau_y / \left(\tau_\theta + \tau_z + \tau_x + \tau_y\right)$. If analysts' belief is such that $\tau_Y < \tau_y$, i.e., a prior belief consistent with the diagnostic belief (Bordalo, Gennaioli, Ma, and Shleifer 2020) or overconfidence (Broer and Kohlhas 2022), then $\kappa^{\text{RE}} < \kappa$. Therefore, analysts appear to be overreacting to information (i.e., $\hat{b}_1 < 0$). Suppose that analysts know the quality of guidance τ_Y; then, the distorted expectation $\hat{\mathbb{E}}[\kappa]$ degenerates to precisely κ^{RE}. This is consistent with the prediction that there is no over- or underreaction in a rational expectation model, i.e., $\hat{b}_1 = 0$.



(a) $\lambda = 0$ (b) $\lambda > 0$

Figure 5. *Distribution of analysts' responses to guidance. Panel (a) displays the cross-sectional distribution of $\kappa\left(s_i\right)$ when analysts are ambiguity neutral. Panel (b) displays the cross-sectional distribution of $\kappa\left(s_i\right)$ when analysts are ambiguity averse.*

Observe that $\hat{\mathbb{E}}$ depends on the prior belief about the information quality, i.e., $p\left(\tau_y\right)$. Therefore, corresponding to various prior beliefs, analysts may either over- or underreact to information. To illustrate this, consider one more special case in which analysts entertain a set of possible models such that $\tau_y > \tau_Y$ for any $\tau_y \in \Gamma_y$ and $\tau_Y \notin \Gamma_y$. That is, the actual quality of guidance is lower than all the possible values in the analysts' belief set. It is straightforward to show that $\kappa^{\text{RE}} < \hat{\mathbb{E}}\left[\kappa\left(s_i\right)\right]$, and therefore, all analysts overreact to information.

Interestingly, when analysts' belief set includes the actual quality, i.e., $\tau_Y \in \Gamma_y$, it may be the case that some analysts underreact to the guidance and others overreact to it. In other words, our model predicts a cross-sectional distribution of analysts who may over- and underreact to the same guidance. Figure 5(a) illustrates this case. For analysts that receive surprises of a smaller magnitude from the guidance, they tend to believe that the quality of guidance is relatively high and therefore react more strongly to the news. For analysts who receive surprises of a larger magnitude from the guidance, they tend to believe that the quality of guidance is relatively low and therefore react less strongly to the news. The shaded area illustrates the fraction of analysts that overreact to the guidance, and the remaining area represents the fraction of analysts that underreact.

In general, the average response can be either higher or lower than κ^{RE} depending on $p\left(\tau_y\right)$. That is, if we regress forecast errors on forecast revisions at the analyst level, we may conclude that the population on average over- or underreacts to information, without noticing that some analysts overreact and others underreact to information.

Figure 5(b) illustrates the case when ambiguity-averse preferences are present in the model. As predicted by proposition 2, if analysts are ambiguity averse, forecast revisions in response to surprises are asymmetric, i.e., it skews the distribution in a negative direction, and analysts react even more strongly to new information when it contains unfavorable news than when it contains favorable news. This is consistent with our finding in section 2.3 that analysts' overreaction to new information is stronger when the managerial guidance is negative.

Combining these findings, the pattern emerging from our model indicates that for very large and positive surprises, analysts overreact to them mildly or even underreact; that relative to positive surprises, analysts overreact more to negative ones; and that for very large negative surprises, analysts overreact less strongly or even underreact. The cross-sectional profile of overreaction present in our model is broadly consistent with the empirical pattern illustrated by Figure 2.

5. Quantitative Analysis

While the patterns of asymmetry, nonmonotonicity and overreaction in our model qualitatively correspond to their counterparts in the data, is the model indeed informative about the empirical findings? In this section, we further pursue a quantitative analysis. We estimate the model using the simulated method of moments to match the relationship between forecast revisions and surprises that is empirically estimated in section 2.4. The estimated model will be interpreted and then used to revisit the pattern of heterogeneous overreaction (documented in sections 2.2 and 2.3).

5.1. Connecting Theory to Data

In the model, we construct surprises in managerial guidances with the analyst's private information, i.e., $s_i \equiv y - X_i$, and implicitly assume its availability. However, in our empirical setting, the econometrician cannot have access to private information available to analysts and can only construct observable surprises with guidance and initial forecasts, i.e., $S_i \equiv y - F_{0i}$.[24] To directly relate the relationship characterized in our model (section 3) to that in the data (section 2.4), we need to carefully distinguish the two surprises. First, in our quantitative exercises, we construct and work with surprises observable to the econometrician (i.e., S_i) from our simulated data and estimate the relationship between forecast revisions and surprises in the same way as we do with the data. Second, we also show that, in our model, forecast revisions can be explicitly approximated by a cubic function of observable surprises in closed form. Our model predictions for the signs of polynomial coefficients imply that forecast revisions may decrease, increase and decrease in observable surprises and that the relationship is asymmetric, a pattern that is consistent with our empirical findings. We relegate the relevant characterization and discussion to Online Appendix A.

[24] The observable surprise to the econometrician is the sum of the surprise to the analyst and a noise term. That is, $S_i = s_i + \frac{\tau_x}{(\tau_\theta+\tau_z+\tau_x)}\,(x_i - F_{0i})$.

Furthermore, our model features a finite degree of ambiguity aversion, i.e., λ is in the range $(0, +\infty)$. On the one hand, our analysis shows that the degree of ambiguity aversion matters for qualitative predictions of the model. On the other hand, it is a quantitative question how much ambiguity aversion is needed to generate a relationship between forecast revisions and guidance surprises that is close to the data. We uncover the degree of ambiguity aversion by estimating our model to match moments in the empirical relationship estimated nonparametrically from the data (section 2.4). Various quantitative exercises using this estimated model will reveal the roles of key mechanisms, such as ambiguity averse preferences and prior beliefs.

5.2. Estimation

The model characterized in section 3.1 is fully specified by two sets of parameters and one distribution. First, two parameters characterize the preferences of analysts, i.e., ambiguity aversion λ, and analysts' attitude toward earnings β.

Second, there is a set of volatilities, i.e., the objective volatility of earnings σ_θ, the objective volatility of managerial guidance σ_Y, the volatility of initial endowed information about earnings before the initial forecast σ_z, and the volatility of private information σ_x.

Third, the analysts' prior belief about guidance quality $p\left(\tau_y\right)$ defined in section 3.1 also needs to be specified. We assume that the ratio $\delta \equiv \tau_y/(\tau_\theta + \tau_z + \tau_x + \tau_y)$ is a uniform distribution over $[L, U]$, where $0 \leq L < U \leq 1$. The advantage of this transformation is that we can entertain the possibility that τ_y is very large, without dealing with a very wide support for τ_y, which economizes our computation. The upper bound U (lower bound L) regulates the perceived largest (smallest) possible precision for managerial guidance.

To estimate the set of parameters $\Theta = \{\lambda, \beta, L, U, \sigma_\theta, \sigma_x, \sigma_Y, \sigma_z\}$, we follow Chernozhukov and Hong (2003) in computing Laplace type estimators (LTE) with an MCMC approach, and the "distance" between the empirical and simulated revision-surprise relationships is constructed in the fashion of the method of simulated moments.

To define the distance, we first choose $N = 50$ equally spaced points for surprises between $[-0.025, 0.03]$, within which the empirical relationship (nonparametrically estimated in section 2.4) decrease, increase and then decrease. Then, we derive the corresponding values of forecast revisions from the estimated revision-surprise relationship and denote them with the vector $\hat{m}$. We further construct the vector m, i.e., the model counterpart of $\hat{m}$, which is estimated by using our simulated dataset. Specifically, for each set of model parameters, we simulate our model and estimate the revision-surprise relationship with the same nonparametric regression as for the empirical data

Table 3. Estimated Model Parameters

	Mean	90% HPDI	95% HPDI
λ	449.9	$(411.9, 504.0)$	$(379.5, 504.2)$
β	1.379	$(0.773, 1.971)$	$(0.694, 2.092)$
U	0.772	$(0.676, 0.855)$	$(0.674, 0.875)$
L	0.082	$(0.036, 0.119)$	$(0.030, 0.121)$
$100\sigma_x$	0.472	$(0.332, 0.593)$	$(0.305, 0.625)$
$100\sigma_z$	0.186	$(0.140, 0.234)$	$(0.137, 0.240)$
$100\sigma_Y$	0.435	$(0.416, 0.453)$	$(0.411, 0.453)$

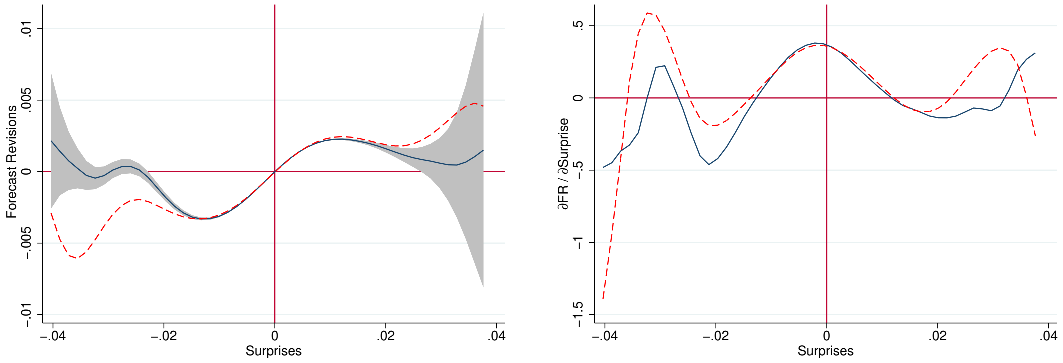
(see section 2.4). We then obtain the vector m from the estimated relationship between forecast revisions and surprises observable to the econometrician. The distance that we construct is:

$$\Lambda(\Theta) = \frac{1}{N} \left[m(\Theta) - \hat{m}\right]' \hat{W} \left[m(\Theta) - \hat{m}\right].$$

where $N = 50$ is the length of the vector of targeted moments $\hat{m}$ and $\hat{W}$ is the weighting matrix with diagonal elements being the precision of moments $\hat{m}$. Our goal is to choose model parameters to "minimize" the distance $\Lambda(\Theta)$ in a pseudo Bayesian manner by using MCMC with the Metropolis-Hastings algorithm.

A few remarks regarding the simulation procedure are in order. First, we choose σ_θ, i.e., the standard deviation of θ, to exactly match the empirical counterpart of an unconditional standard deviation of realized earnings (after removing the firm and time fixed effects). As a result, the calibrated value of $100\sigma_\theta$ is 0.985. Second, when we simulate the model, we feed surprises (to the econometrician) uncovered from the empirical data into our simulation. We recover the corresponding surprises to the analysts and then obtain updated forecasts by using decision rules in our model. Third, in this model, the unconditional volatility of surprises to the econometrician is determined by both σ_Y and $1/\sigma_\theta^2 + 1/\sigma_z^2$. We directly estimate $1/\sigma_\theta^2 + 1/\sigma_z^2$ in the estimation and back out σ_Y by requiring that the unconditional volatility of surprises matches its empirical counterpart, an internal consistency condition for our estimation strategy.

The estimated parameters are reported in Table 3 together with the 90% and 95% high posterior density interval (HPDI). The relative magnitude of the estimated volatilities appears to be reasonable. The volatility of private information is larger than that of earnings. The managerial guidance is much more precise than the private information. This is likely because there may not be much private information that arrives during the time window that we construct (i.e., between the two forecasts around the



(a) $\lambda = 449.9$, Nonparametric estimation (b) $\lambda = 449.9$, Derivative

Figure 6. *The revision-surprise relationship nonparametrically estimated with simulation data. We simulate the model with the set of parameters reported in Table 3 and, in particular,* $\lambda = 449.9$*. The dashed line in panel (a) illustrates the revision-surprise relationship estimated with simulation data. The empirical counterpart (i.e., the solid line) and its confidence interval (i.e., the shaded area) are also plotted for comparison. Panel (b) illustrates the derivative of the revision-surprise relationship with the dashed line. Its empirical counterpart is illustrated with the solid line.*

date of managerial guidance release). Based on this set of parameters, the response of forecast revisions to surprises under noisy rational expectation (i.e., κ^{RE}) is 0.132. The upper bound for the subjective belief on managerial guidance precision is 0.772, and the lower bound is 0.082. The support is large enough to allow sufficient ambiguity and encompass κ^{RE}. The degree of ambiguity aversion is the key, and its value is estimated to be $\lambda = 449.9$.

Furthermore, the parameter β that characterizes the utility function is estimated to be positive (i.e., slightly larger than 1), indicating that analysts are likely to care about the earnings performance of firms that they cover. Prior empirical studies suggest that it is plausible that β is positive. There are multiple channels through which financial analysts would benefit from better earnings performance of the firms that they cover and therefore view positive surprises in managerial guidance as favorable. For example, stronger earnings performance can be rewarding to financial analysts who make earnings forecasts and recommendations for the underlying stocks through the trading commissions channel.[25]

Using this set of estimated parameters, we simulate the model and nonparametrically estimate the revision-surprise relationship with the simulation data. In Figure 6(a), we display the relationship, together with its empirical counterparts (previously

[25]One goal of financial analysts is to generate stock trading and bring in trading commissions to the brokerage houses for which they work. Analysts' positive recommendations, based on earnings expectations, are more likely to generate a larger trading volume that, in turn, is beneficial for the analysts. Barber and Odean (2008) provide evidence that investors are more likely to follow analysts' positive recommendations (i.e., "buy-type") and tend to be net buyers.

Table 4. *Regress Forecast Errors on Forecast Revisions and Regress Forecast Revisions on Surprises*

		(1)	(2)	(3)	(4)
	Coff.	Data	$\lambda = 449.9$	$\lambda = 0$	$\lambda \to \infty$
(A) $\text{FE}_i = b_0 + b_1 \text{FR}_i + \omega_i$	b_1	-0.0950^{***}	-0.4603^{***}	-0.4387^{***}	-0.6938^{***}
			$(-0.4686, -0.4512)$	$(-0.4475, -0.4304)$	$(-0.6998, -0.6882)$
(B) $\text{FR}_i = b_0 + b_1 \text{Surp}_i + \omega_i$	b_1	0.2441^{***}	0.3007^{***}	0.2917^{***}	0.4391^{***}
			$(0.2992, 0.3020)$	$(0.2903, 0.2932)$	$(0.4366, 0.4423)$
(C) $\text{FR}_i = b_0 + b_1 \text{Surp}_i + b_2 \text{Unf}_i + b_3 \text{Surp}_i \times \text{Unf}_i + \omega_i$	b_1	0.1405^{***}	0.2159^{***}	0.2537^{***}	0.1056^{***}
			$(0.2126, 0.2190)$	$(0.2505, 0.2569)$	$(0.1034, 0.1079)$
	b_3	0.0846^{***}	0.0914^{***}	-0.0001	0.6669^{***}
			$(0.0864, 0.0957)$	$(-0.0042, 0.0048)$	$(0.6644, 0.6690)$
(D) $\text{FR}_i = b_0 + b_1 \text{Surp}_i + b_2 \text{Large}_i + b_3 \text{Surp}_i \times \text{Large}_i + \omega_i$	b_1	0.4707^{***}	0.3565^{***}	0.3459^{***}	0.4391^{***}
			$(0.3550, 0.3584)$	$(0.3441, 0.3476)$	$(0.4365, 0.4417)$
	b_3	-0.2674^{***}	-0.0697^{***}	-0.0677^{***}	0.0000
			$(-0.0719, -0.0672)$	$(-0.0699, -0.0654)$	$(-0.0032, 0.0031)$

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. In columns (2), (3) and (4), we report the average of point estimates and a 99% high posterior density interval (HPID) in the bracket (calculated based on the posterior distribution of point estimates for $N = 1000$ estimations). For each simulation, there are 101,086 observations of analysts, and the size is consistent with our data.

shown in Figure 3(a)). In Figure 6(b), we illustrate its implied derivative with respect to surprises. Our model can successfully capture both features of nonmonotonicity and asymmetry.

The estimated value of λ is not too high or too low, indicating that neither extreme ambiguity aversion ($\lambda \to +\infty$) nor ambiguity-neutral preferences ($\lambda = 0$) would be realistic for analysts in this setting. To illustrate this, we simulate the model $N = 1000$ times, and in each simulation there are 101,086 observations of analysts (i.e., the size is consistent with our empirical data). In each simulation, we estimate equations (A), (B), (C) and (D) in Table 4, which are counterparts of equations (1), (2), (3) and (4) in sections 2.2 and 2.4.[26] For comparison, we repeat the aforementioned exercises twice with the same set of parameters, except allowing λ to be 0 or $+\infty$. Then, we estimate equations (A), (B), (C) and (D) with the two additional datasets. We report the average of the point estimates from each simulation and its 99% high posterior density interval (HPID) in columns (2), (3) and (4), corresponding to $\lambda = 449.9$, $\lambda = 0$ and $\lambda \to \infty$, respectively. We reproduce the estimates from the data in column (1).

Column (2) shows that the average response of forecast revisions to surprises in our estimated model (the coefficient on b_1 in equation (B)) is fairly close to that in the data. Our model can also generate differential responses to positive and negative surprises, as well as to large and small surprises. That is, both interaction terms in equations (C) and (D) are highly significant. In other words, our estimated model is consistent with the empirical results obtained using linear regressions.

[26]For this purpose, we do not need to use the surprises revealed in the data to simulate the model. Instead, we resample them for each simulation.

Column (3) presents the results when $\lambda = 0$. It is evident that the revision-surprise relationship is symmetric, i.e., the interaction term in equation (C) is almost zero and insignificant. However, the size of surprises matters for the response of forecast revisions to surprises, i.e., the interaction term in equation (D) is highly significant and negative.

Column (4) presents the results when $\lambda \to \infty$. It is evident that the revision-surprise relationship is asymmetric, i.e., the interaction term in equation (C) is positive and significant. However, the size of surprises does not matter for the response of forecast revisions to surprises, i.e., the interaction term in equation (D) is zero and insignificant.

In sum, to simultaneously capture both qualitative features of nonmonotonicity and asymmetry, we do need a moderate amount of ambiguity aversion.

5.3. Heterogeneous Overreaction and Ambiguity Aversion

In this section, we examine whether our estimated model can produce the pattern of heterogeneous overreaction found in the data (in section 2.3). To investigate this, we utilize the simulated data and construct the surprises observable to the econometrician in the same way as we do with the empirical data. We rank surprises from the most negative to the most positive and sort them into deciles, labelling them from 1 to 10 according to the decile rank. We further define a running decile window j, such that (1) the window j covers decile $j-1$, j, and $j+1$ if $j \neq 1$ or $j \neq 10$; (2) running decile window 1 covers deciles 1 and 2; and (3) running decile window 10 covers deciles 9 and 10. For each subsample, we re-estimate equation (A). We plot the estimated coefficients and confidence intervals in Figure 7 against their window ranks. In the simulated data, we find that the pattern of heterogeneous overreaction is U-shaped and skewed to the left, which is consistent with our model predictions in section 4.3 and also close to the pattern in the empirical data (Figure 2).

While our model can predict the pattern of cross-sectional variation in overreaction, can it also produce overreaction to information at the aggregate level? To examine the average extent of overreaction implied by the estimated model, we first estimate the regression specified in equation (A) (in Table 4) with our simulation data ($\lambda = 449.5$). It is the counterpart of equation (1) in section 2.2. We contrast the results from the empirical and simulated data in columns (1) and (2) of Table 4, respectively. All of the reported coefficients from our simulate data are highly significant. In regression (A), we observe that the estimated coefficient on b_1 is negative and significantly different from zero, indicating that analysts do, on average, indeed overreact to information in our model, even though it is not targeted in the estimation.

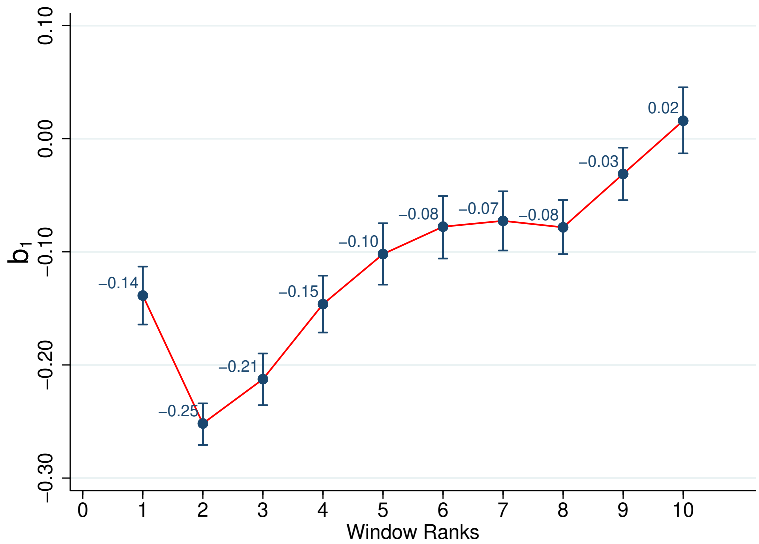


Figure 7. *Overreaction by surprise deciles with simulated data. Using simulated data, we report the estimated coefficients of the FE-on-FR regressions* b_1 *for each running decile window, and we plot them against the window rank. Running decile window* j *covers decile* $j-1$, j, *and* $j+1$ *if* $j \neq 1$ *or* $j \neq 10$; *running decile window* 1 *covers deciles* 1 *and* 2, *and running decile window* 10 *covers deciles* 9 *and* 10.

Why is it the case that analysts overreact to information in our estimated model? Recall that analysts act as if their posterior beliefs are governed by equation (12). It has two key components: a Bayesian kernel and pessimistic distortion, both of which contribute to the observed overreaction.

To illustrate the role of prior beliefs and ambiguity averse preferences, we simulate the model by varying the prior distribution of τ_y or, specifically, the upper bound of the prior belief U (defined in section 5.2) from a value larger than the lower bound L to 1, while we keep other parameters unchanged. Recall that the distribution of τ_y is such that the ratio $\delta \equiv \tau_y/(\tau_\theta + \tau_z + \tau_x + \tau_y)$ is a uniform distribution over $[L, U]$, where $0 \leq L < U \leq 1$. A higher U implies that a larger fraction of probabilistic density in the prior belief is allocated to the right of the objective quality τ_Y, corresponding to a situation where analysts are more likely to overestimate the quality of guidance.

We re-estimate regression (A) for each simulation and plot the average value of estimated coefficients for b_1 against the corresponding upper bound U. The solid line in Figure 8 illustrates the relationship: when U is small, the estimated coefficients on b_1 are small in magnitude; when U is large, they are negative and larger in magnitude. This is intuitive: when analysts' prior belief is such that the guidance is, on average, more precise than it actually is, they overreact.

We repeat the same exercise by using simulation data that sets $\lambda = 0$ and plot the counterpart in Figure 8 with the dashed line. The contrast between the cases of

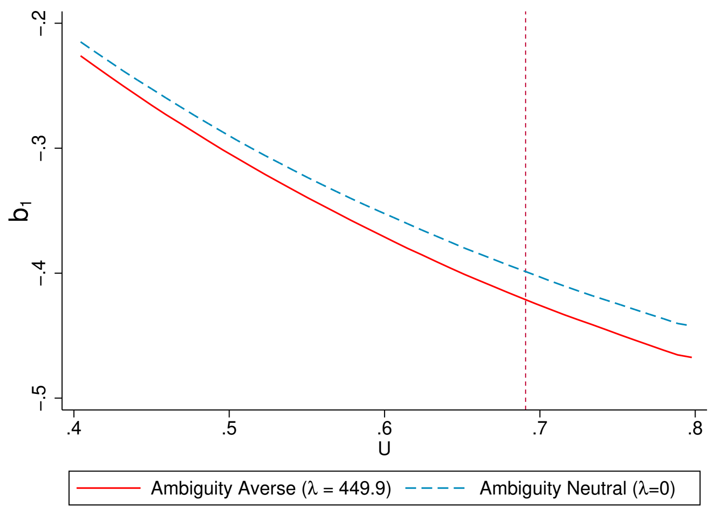


Figure 8. *Ambiguity aversion and prior beliefs. The solid line illustrates the relationship between estimated coefficients b_1 and the upper bound of the prior belief, i.e., U. The dashed line corresponds to the case where $\lambda = 0$. For the same value of U, a larger degree of ambiguity aversion leads to a larger overreaction. The vertical dashed line indicates the value of U in our benchmark estimation.*

$\lambda = 0$ and $\lambda = 449.5$ shows that ambiguity-averse preferences indeed contribute to the overreaction. For the same prior belief (represented by U here), analysts tend to overreact more when they are ambiguity averse.

The comparison of the results with and without ambiguity aversion reveals mechanisms that drive the overreaction to information in this model. First, the prior belief $p(\tau_y)$ plays a role in the observed overreaction: the estimated prior allocates a sufficiently large density to the right of objective quality τ_Y. That is, our structural estimation shows that analysts tend to overreact to information even without the distortion of ambiguity aversion preferences. Such an estimated prior is hardly surprising: it can be interpreted as a version of diagnostic belief (Bordalo, Gennaioli, Ma, and Shleifer 2020) or overconfidence (Broer and Kohlhas 2022) in the context that model uncertainty exists.[27]

Second, the ambiguity aversion (or pessimistic distortion) also contributes to the observed overreaction. On the one hand, analysts may overreact to information because they treat unfavorable surprises as higher quality news than they actually are. On the other hand, analysts may underreact to information because they discount the quality of large surprises more heavily. In this estimated model, we observe that the former dominates the latter, because large surprises by definition account for only a

[27]Put it crudely, analysts may appear to be over-reacting to information in managerial guidance, when they allocate higher probabilistic density in their posterior beliefs to the states of the world where the managerial guidance's quality is higher than its objective value τ_Y. Conversely, analysts may appear to be under-reacting to information in managerial guidance, when they allocate lower probabilistic density in their posterior beliefs to such states.

smaller fraction of the sample.

We need to stress that this model does not have a priori about the *aggregate level* of overreaction and only predicts the *cross-sectional pattern*. While the aggregate level of overreaction has been documented and studied by a number of aforementioned contributions, we discover and rationalize heterogeneous overreactions to shocks of different properties, which is one step further from the existing literature.

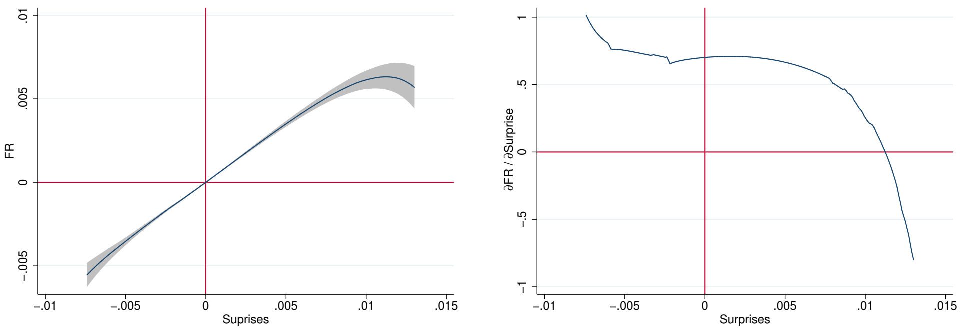
5.4. Auxiliary Predictions

In this paper, we provide a theory about how the expectation is formed when forecasters are not completely certain of the quality of the information that they receive. Our theory organizes a number of facts that we document with the earnings forecasts data. Our underlying assumption is that firms' earnings guidance is of uncertain quality. However, the extent of uncertainty is likely heterogenous across firms. For example, there should be established firms with a good reputation whose managerial guidance is of high quality and analysts have little doubt about its quality.

For firms with low or no uncertainty in earnings guidance quality, our theory predicts that analysts' forecast revisions should be close to linear in guidance surprises, i.e., the relationship is monotonic and symmetric. That is because once the uncertainty in quality has been removed, analysts only update their beliefs based on the guidance and do not need to update their beliefs on the quality.

There is a conceptual hurdle to testing this auxiliary prediction with our data: the perceived uncertainty in guidance quality is not observable and therefore not measurable. To circumvent this issue, we proxy for it with the observed average quality in the data, i.e., the ex post variance of the differences between guidance and actual earnings in the data. Our assumption is that the perceived uncertainty in quality is low if the observed average quality is high.

We construct a subsample that only includes firms that deliver very precise earnings guidance whose uncertainty with respect to quality is supposed to be low. The first step is to rank firms in terms of their average guidance quality. Our full sample of 110,895 individual analyst forecasts consists of 16,241 firm-quarter observations, based on which we trim realized earnings and management guidance (both scaled by the stock price at the prior-quarter end) at the 2.5% and 97.5% percentiles of their respective distributions. With the remaining 15,427 firm-quarter observations, we regress management guidance on the realized earnings of the same quarter by controlling for year-quarter fixed effects to obtain the residuals. We drop firms present for fewer than 5 quarters, which reduces the sample to cover 1,035 firms. Then, we compute the standard deviations of the residuals for each remaining firm and sort those firms according



(a) Nonparametric estimation (b) Derivative: marginal effect

Figure 9. *Nonparametric estimation using a subsample with the top 5% of firms in terms of guidance precision. Panel (a) illustrates the relationship between forecast revisions and surprises in managerial guidance that is nonparametrically estimated using the Epanechnikov kernel and the third degree of the smoothing polynomial. The shaded areas represent the 95% confidence intervals for the respective estimations. Panel (b) illustrates the derivative of forecast revisions with respect to surprises.*

to the calculated standard deviations. Second, we focus on the top 5% of firms with the highest average guidance quality and construct a corresponding subsample of 2,521 individual analyst forecast revisions and guidance surprises.

Using this subsample, we nonparametrically re-estimate the relationship between forecast revisions and guidance surprises by following the same procedure as detailed in section 2.4. The results are shown in Figure 9(a). The relationship between forecast revisions and surprises is almost linear, unless the surprises are relatively very large and positive. The derivative estimated and shown in Figure 9(b) is close to a constant when the surprises are not too large, thus contrasting with the derivative estimated using the full sample (shown in Figure 3(b)). This exercise strengthens our confidence in our model mechanisms. Furthermore, it helps us stress that uncertainty in information quality does have an impact on how analysts update their beliefs.

6. Discussions on Alternative Hypotheses

In this paper, we provide a simple unified framework to account for new facts regarding how analysts update their forecasts or form expectations. It is important that our estimated model can generate the skewed U-shaped pattern of overreaction that is consistent with the data. This paper is the first that discovers and rationalizes this set of facts in the literature of expectation formation. Nevertheless, we acknowledge that there could be other mechanisms that simultaneously contribute to the observed patterns. We examine several likely candidates in sequence, which helps differentiate our theory from those in the existing literature.

6.1. Diagnostic Expectations and Over-confidence

Bordalo, Gennaioli, Ma, and Shleifer (2020) show that forecasters with diagnostic expectations over-react to new information at the individual level. Diagnostic expectations proposed by Bordalo, Gennaioli, and Shleifer (2018) is a non-Bayesian model of belief formation that formalizes representativeness heuristic (Kahneman and Tversky 1972): agents overweight states that are more likely in light of the arrival of new signals. As a consequence, agents over-react to new information when forming expectations. Specifically, agents update their beliefs using the following distorted posterior density $f^{\psi}\left(\theta|\mathcal{I}_{it}\right)$:

$$f^{\psi}\left(\theta|\mathcal{I}_{it}\right) \propto f\left(\theta|\mathcal{I}_{it}\right)\left(\frac{f\left(\theta|\mathcal{I}_{it}\right)}{f\left(\theta|\mathcal{I}_{it-1}\right)}\right)^{\psi}$$

where $f\left(\theta|\mathcal{I}_{it}\right)$ denotes the Bayesian posterior density and the constant $\psi \geq 0$ measures the extent to which the posterior of agents with diagnostic expectations are distorted away from Bayesian benchmark. When $\psi = 0$, the model with diagnostic expectations reduces to the noisy information benchmark. Observe that $R_t\left(\theta\right) \equiv \frac{f(\theta|\mathcal{I}_{it})}{f(\theta|\mathcal{I}_{it-1})}$ measures the representativeness of the new information defined as the gap between $\mathcal{I}_{it}$ and $\mathcal{I}_{it-1}$. When $\psi > 0$, the distorted posterior belief overweights states θ featuring $R_t\left(\theta\right) \geq 1$, which leads to overreaction to the arrival of new information (Bordalo, Gennaioli, Ma, and Shleifer 2020).

In this section we investigate the optimal forecasting rule in a setting where we allow for diagnostic expectations in the benchmark model specified in section 3.2. In this case, the optimal initial and updated forecasts are given by:

$$\mathrm{F}_{0i}^{\mathrm{DE}} = \mathbb{E}\left[\theta|z_{0i}\right] + \psi\left(\mathbb{E}\left[\theta|z_{0i}\right] - \mathbb{E}\left[\theta\right]\right),$$

and

$$\mathrm{F}_{i}^{\mathrm{DE}} = \mathbb{E}\left[\theta|z_{0i}, x_i, y\right] + \psi\left(\mathbb{E}\left[\theta|z_{0i}, x_i, y\right] - \mathbb{E}\left[\theta|z_{0i}\right]\right),$$

where $\mathbb{E}\left[\theta|\mathcal{I}_i\right]$ denotes the conditional expectations (i.e., Bayesian posterior). As a consequence, forecast revisions under diagnostic expectations are given by:

$$\mathrm{FR}_{i}^{\mathrm{DE}} = \left(1+\psi\right)\kappa^{\mathrm{RE}}\left(y - \mathrm{F}_{0i}^{\mathrm{DE}}\right) + \left(1+\psi\right)\kappa_x\left(x - \mathrm{F}_{0i}^{\mathrm{DE}}\right) + \psi\left[\left(\kappa_x + \kappa^{\mathrm{RE}}\right) - \frac{1}{1+\psi}\right]\mathrm{F}_{0i}^{\mathrm{DE}}, \tag{16}$$

where $\kappa_x \equiv \frac{\tau_x}{\tau_\theta+\tau_z+\tau_x+\tau_Y} > 0$ and $\kappa^{\mathrm{RE}} \equiv \frac{\tau_Y}{\tau_\theta+\tau_z+\tau_x+\tau_Y} > 0$ are the respective weights for the private and public information in the noisy information benchmark. Similar to the noisy information benchmark, the term $y - \mathrm{F}_{0i}^{\mathrm{DE}}$ is the theoretical counterpart to the manager guidance surprises in our empirical exercise. Observe that the relationship between forecast revisions and surprises in Equation (16) remains linear and state-independent.

Broer and Kohlhas (2022) show that over-confidence can help rationalize over-reaction documented with SPF data, in which they assume that forecasters subjectively believe that new signals are more precise than they actually are. Interestingly, once we allow for such behavioral feature in the noisy information benchmark specified in section 3.2, the FR-on-Surprise relation is still linear, while overreaction to new information appears. To see this, specifically we assume that analysts subjective believe that $\eta \sim \mathbb{N}\left(0, 1/\bar{\tau}_y\right)$ such that $\bar{\tau}_y > \tau_Y$ and derive the relationship between forecast revisions and surprises as follows:

$$\mathrm{FR}_i^{\mathrm{OC}} \equiv \mathrm{F}_i^{\mathrm{OC}} - \mathrm{F}_{0i} = \bar{\kappa}_y \left(y_i - \mathrm{F}_{0i}\right) + \bar{\kappa}_x \left(x_i - \mathrm{F}_{0i}\right),$$

where $\bar{\kappa}_x \equiv \frac{\tau_x}{\tau_\theta + \tau_z + \tau_x + \bar{\tau}_y} > 0$ and $\bar{\kappa}_y \equiv \frac{\bar{\tau}_y}{\tau_\theta + \tau_z + \tau_x + \bar{\tau}_y} > 0$.

Corollary 1: If forecast revisions are linear in surprises, forecast errors are linear in forecast revisions, or the degree of overreaction does not depend on realizations of surprises.

When forecast revisions are linear in surprises, it must be the case that both forecast errors and forecast revisions are linear in Gaussian noises. Such a property implies that $\frac{\mathbb{COV}(\mathrm{FE},\mathrm{FR})}{\mathbb{V}(\mathrm{FR})}$ can be zero or a non-zero constant but is always independent of the realizations of surprises.

6.2. Loss Aversion

Another plausible conjecture is that analysts are loss-averse, which may also likely generate the empirical pattern documented in section 2.4, given that they behave in a pessimistic way. To investigate this possibility, we consider two commonly used specifications of loss aversion: one parsimonious setup with analytical solutions (Capistrán and Timmermann 2009) and another flexible setup with more quantitative potentials (Elliott, Komunjer, and Timmermann 2008, Elliott and Timmermann 2008). In this section, we show that (1) the parsimonious setup predicts a linearly increasing relation between forecast revisions and surprises and that (2) the flexible setup predicts a monotone increasing relation between forecast revisions and surprises.

The Parsimonious Setup. We follow Capistrán and Timmermann (2009) and specify the loss function of analysts to be:

$$L\left(\mathrm{F}_i, \theta; \phi\right) = \frac{1}{\phi^2}\left[\exp\left(\phi\left(\theta - \mathrm{F}_i\right)\right) - \phi\left(\theta - \mathrm{F}_i\right) - 1\right],$$

where F_i stands for the forecast of analyst i and the parameter ϕ is a constant that captures asymmetries in the loss function. If $\phi > 0$, analysts dislike negative forecast error $\theta - \mathrm{F}_i < 0$ more than positive forecast error $\theta - \mathrm{F}_i > 0$. If ϕ goes to zero, the loss function is reduced to the standard MSE function. Information structure is the same as that of the noisy information benchmark.

Analyst i chooses the optimal forecasts F_i^L to minimax the loss function conditional on her information set, which leads to her decision rule:

$$\mathrm{F}_i^L = \mathbb{E}_i\left[\theta\right] - \frac{1}{2}\phi \mathrm{Var}_i\left[\theta\right].$$

Relative to the noisy information expectations benchmark ($\phi = 0$), the loss-averse analyst i ($\phi > 0$) would like to inflate the forecast errors and bias her forecast downward by $\frac{1}{2}\phi \mathrm{Var}_i\left[\theta\right]$.

Accordingly, the initial optimal forecast is given by:

$$\mathrm{F}_{0i}^L = \mathbb{E}\left[\theta | z_{0i}\right] - \frac{1}{2}\phi \mathrm{Var}\left[\theta | z_{0i}\right] = \frac{\tau_z}{\tau_\theta + \tau_z} z_{0i} - \frac{1}{2}\phi \frac{1}{\tau_\theta + \tau_z}.$$

and the updated optimal forecast is given by:

$$\mathrm{F}_i^L = \mathbb{E}\left[\theta | z_{0i}, x_i, y\right] - \frac{1}{2}\phi \mathrm{Var}\left[\theta | z_{0i}, x_i, y\right] = \frac{\tau_z z_i + \tau_x x_i + \tau_Y y}{\tau_\theta + \tau_z + \tau_x + \tau_Y} - \frac{1}{2}\phi \frac{1}{\tau_\theta + \tau_z + \tau_x + \tau_Y}.$$

Therefore, the forecast revision is such that

$$\mathrm{FR}_i^{\mathrm{L}} \equiv \mathrm{F}_i^L - \mathrm{F}_{0i}^L = \kappa_y\left(y - \mathrm{F}_{i0}^{\mathrm{L}}\right) + \kappa_x\left(x_i - \mathrm{F}_{i0}^{\mathrm{L}}\right). \tag{17}$$

where $\kappa_x \equiv \frac{\tau_x}{\tau_\theta + \tau_z + \tau_x + \tau_y} > 0$ and $\kappa_y \equiv \frac{\tau_Y}{\tau_\theta + \tau_z + \tau_x + \tau_y} > 0$ are the relevant weights for the private and public information under noisy information benchmark. Observe that the relation between forecast revisions and guidance surprises are still linear.

The Flexible Setup. Following Elliott, Komunjer, and Timmermann (2008) and Elliott and Timmermann (2008), we specify the loss function of analysts to be:

$$L_p\left(\mathrm{F}_i, \theta; \alpha\right) = \left[\alpha + (1 - 2\alpha)\mathbb{1}\left\{\theta - \mathrm{F}_i < 0\right\}\right]\left|\theta - \mathrm{F}_i\right|^p, \tag{18}$$

where $\mathbb{1}\{\cdot\}$ denotes an indicator function, the parameter $\alpha \in (0,1)$ is a constant captures asymmetries in the loss function, and the parameter $p \geq 1$ is another constant that controls the curvature of the loss functions.

This specification of loss function is flexible and can be reduced to many commonly used loss functions in the literature (as shown in Elliott, Komunjer, and Timmermann 2008 and Elliott and Timmermann 2008). For example, if $\alpha = \frac{1}{2}$ the loss function is symmetric. It can be further reduced to the standard MSE loss function if $p = 2$ or mean absolute error function if $p = 1$. In particular, when $\alpha < \frac{1}{2}$, negative forecast error ($\theta - \mathrm{F}_i < 0$) disproportionately induces more loss than positive forecast error ($\theta - \mathrm{F}_i > 0$), indicating that analysts are loss averse.

For the ease of exposition, we focus our analysis on the parameter space that $p = 2$, that is, the loss function is of a generalized MSE form. However, it is noted that all results presented below can be extended to the general case that $p \geq 1$.

Analyst i chooses the optimal forecasts F_i^L to minimax the loss function conditional on her information set. Implicitly, her optimal decision rule is characterized by:

$$\int_{-\infty}^{+\infty} \left(\theta - \mathrm{F}_i^L\right) f\left(\theta|\mathcal{I}_i\right) \mathrm{d}\theta + \frac{1-2\alpha}{\alpha} \int_{-\infty}^{\mathrm{F}_i^L} \left(\theta - \mathrm{F}_i^L\right) f\left(\theta|\mathcal{I}_i\right) \mathrm{d}\theta = 0, \tag{19}$$

where $f\left(\theta|\mathcal{I}_i\right)$ denotes the posterior density of the fundamental θ with respect to the information set $\mathcal{I}_i$.

It is worth noting that when $\alpha = \frac{1}{2}$ (i.e., the loss function is symmetric), only the first term of the LHS of (19) is relevant. Therefore, the optimal forecast is just the conditional expectation as in the noisy information setup:

$$\int_{-\infty}^{+\infty} \left(\theta - \mathrm{F}_i^L\right) f\left(\theta|\mathcal{I}_i\right) \mathrm{d}\theta = 0 \Rightarrow \mathrm{F}_i^L = \mathbb{E}\left[\theta|\mathcal{I}_i\right].$$

where $\mathbb{E}\left[\theta|\mathcal{I}_i\right]$ denotes the conditional mean under Bayesian posterior.[28]

Furthermore, observe that the second term of the LHS of (19) is negative if and only if analysts are loss averse ($\alpha < \frac{1}{2}$). Therefore, as in the parsimonious setup, loss averse analysts would like to inflate the forecast errors by biasing their forecasts downwards

$$\mathrm{F}_i^L < \mathbb{E}\left[\theta|\mathcal{I}_i\right].$$

Proposition 6. *With the flexible specification of the loss function* (18), *forecast revisions are globally monotone in surprises.*

[28] For the general case $p \geq 1$, the same result holds. Intuitively, as long as the information structure is symmetric, any symmetric loss function implies that optimal forecasts are the conditional expectations (Bhattacharya and Pfleiderer 1985).

To understand this lemma, we note that a sufficient condition for the global monotonicity is that the optimal forecast F_i^L is globally monotone in signals. According to (19), the optimal forecasts can be written as the summation of the Bayesian posterior mean and a bias:

$$\mathrm{F}_i^L = \underbrace{\mathbb{E}\left[\theta|\mathcal{I}_i\right]}_{\text{Bayesian}} + \underbrace{\frac{1-2\alpha}{\alpha}\int_{-\infty}^{\mathrm{F}_i^L}\left(\theta - \mathrm{F}_i^L\right) f\left(\theta|\mathcal{I}_i\right) \mathrm{d}\theta}_{\text{Bias}}. \tag{20}$$

In the proof of Lemma 6, we demonstrate that both the Bayesian posterior mean and the bias are increasing in signals of the fundamental θ, which is the same as the parsimonious setup. Though the FR-on-Surprise relation can be non-linear in this case, it is still globally monotone, which is still inconsistent with the documented non-monotone FR-on-Surprise relation in section 2.4.

6.3. Dynamic models.

Using the Survey of Professional Forecasters (SPF), Kohlhas and Walther (2021) show that forecasters' expectations overreact to recent realizations of the output growth and therefore display a pattern of extrapolation. To explain this, they propose a model of "asymmetric attention", where Bayesian agents pay more attention to the procyclical component and less attention to the countercyclical component. Afrouzi, Kwon, Landier, Ma, and Thesmar (2022) design an experiment where participants who observe a large number of past realizations of a given AR(1) process make forecasts about future realizations. They show a pattern of overreaction, i.e., the perceived persistence of the AR(1) process is larger than the actual persistence. They propose a "top-of-mind"model, where agents rely excessively on or overreact to the recent realizations, relative to the rational benchmark.

In our empirical setting, both initial and updated forecasts are made within the same period, which encompass the earnings guidance for the current period. We use the variations of surprises contained in the earnings guidance across analysts to explore impacts of surprises' characteristics on forecast revisions. Dynamic models such as Kohlhas and Walther (2021) and Afrouzi, Kwon, Landier, Ma, and Thesmar (2022) are not informative about the cross-sectional heterogeneity of overreaction.

To illustrate, recall that we show in section 2.2 that earnings in the last quarter cannot predict forecast errors conditional on forecast revisions (see Table 1). How do forecast errors react to earnings in the last quarter without controlling forecast revisions? We run a regression of forecast errors (i.e., analyst i's forecast error in quarter t for firm j) on earnings in the last quarter (quarter $t-1$ and firm j) with a full set of

Table 5. Forecast Errors, Forecast Revisions and Earnings in the Last Quarter

	Outcome Variable: in Quarter t for Firm j, analyst i's			
	Forecast Errors		Forecast Revisions	
	1% and 99%	2.5% and 97.5%	1% and 99%	2.5% and 97.5%
	(1)	(2)	(3)	(4)
Earnings in the Last Quarter ($t-1$)	0.0008	-0.0010	-0.0048	-0.0058
	(0.0070)	(0.0049)	(0.0066)	(0.0053)
Surprise$_i$			0.1468***	0.2445***
			(0.0125)	(0.0128)
Constant	-0.0000	0.0001**	-0.0009***	-0.0004***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Quarter FEs	YES	YES	YES	YES
Analyst FEs	YES	YES	YES	YES
Firm FEs	YES	YES	YES	YES
Obs.	110,895	110,895	110,895	110,895
Adj. R-sq	0.2341	0.2202	0.3943	0.4588

The standard errors are clustered on firm and calendar year-quarter following (Petersen 2009).
*** p<0.01, ** p<0.05, * p<0.1

fixed effects as in equation (1). We report the estimation results in Table 5. Column (1) shows that the estimated coefficient is very small and insignificant, suggesting that earnings in the last quarter cannot predict analysts' forecast errors. To ensure robustness, we winsorize the FE_{ijt} and the last quarter earnings at 2.5% and 97.5% of their respective distributions and re-estimate equation (1). We report the results in column (2), which are consistent with those in column (1). This result is in contrast with both Kohlhas and Walther (2021) and Afrouzi, Kwon, Landier, Ma, and Thesmar (2022).

Furthermore, in this setting, we predict that forecast revisions would not be affected by earnings in the last quarter. To confirm this, we add earnings in the last quarter to equation (2) and re-estimate it. The results are reported in columns (3) and (4) of Table 5 for different levels of winsorization. The estimated coefficient on earnings in the last quarter is very small and insignificant, suggesting that they do not affect analysts' forecast revisions in the current period either.

This set of results is intuitive: the initial forecast in this setting absorbed information contained in earnings in the last quarter, which do not impact forecast revisions that take place in the current quarter. Therefore, forecast revisions reflect the impact of earnings guidance, instead of the impact of earnings in pervious quarters. By contrast, in studies using SPF data, "initial forecasts" for a random variable x_{t+k} in period $t+k$ are made in period $t-1$ and "updated forecasts" are made in period t after observing the current realization of the variable x_t.

6.4. Agency issues

This empirical setting is new to the literature and informative about expectation formation. However, one may worry about the role of agency issues between analysts

Table 6. *Guidance Quality and Negativity*

	Outcome Variable: Absolute Difference between Guidance and Earnings			
	Sample: Full		Exclude Conforming	
	1% and 99%	2.5% and 97.5%	1% and 99%	2.5% and 97.5%
	(1)	(2)	(3)	(4)
Negative Guidance	0.0012***	0.0008***	0.0010***	0.0003
	(5.1339)	(3.7038)	(3.1044)	(1.2632)
Constant	0.0050***	0.0048***	0.0057***	0.0056***
	(35.1519)	(38.5143)	(26.4874)	(30.2028)
Quarter FEs	YES	YES	YES	YES
Firm FEs	YES	YES	YES	YES
Observations	15,528	15,528	13,476	13,500
Adjusted R-squared	0.6105	0.5395	0.6151	0.5428

Notes: The observation numbers in columns (3) and (4) vary because the numbers of conforming cases vary due to Winsorization. The standard errors are clustered on firm and year-quarter. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

and the managerial teams who might have incentives to misrepresent information. Managers could have both incentives to overstate or understate information related to earnings (discussed below). In fact, that is one of the crucial factors underlying our assumption that the quality of managerial guidance is uncertain. Nevertheless, we examine two such likely mechanisms, respectively.

Skewed information reliability. In the literature, it is often shown that managers spin information in self-serving ways to cater to investors and analysts (e.g. Solomon 2012). Given managerial guidance is an important information protocol provided by managers, it is reasonable to conjecture that managers have an incentive to bias their guidance positively, which makes positive managerial guidance less reliable than negative managerial guidance. This skewed information reliability, if exists, may lead to the asymmetry we documented.

It is empirically testable whether positive guidance is of lower quality on average.[29] We regress a measure of guidance quality on guidance negativity and report the results in Table 6. Specifically, the dependent variable is the absolute difference between managerial guidance and actual realized earnings per share for firm i in quarter t, scaled by the stock price at the beginning of quarter t. The independent variable of our interest, Negative Guidance, is an indicator, which is equal to 1 if the managerial guidance is smaller than the median of individual analysts's initial forecasts before guidance, and 0 otherwise. We control for firm and calendar year-quarter fixed effects so that the results cannot be explained by any time-invariant firm characteristics and

[29]We thank one anonymous referee for providing the idea of testing this conjecture.

across-quarter differences.

Column (1) reports the regression results based on the full sample of 15,528 firm-quarter observations, while column (2) presents results of the same specification except that we winsorize the managerial guidance and the difference between guidance and earnings at the 2.5% and 97.5 levels to mitigate potential bias driven by extreme observations. Furthermore, we repeat the same set of exercises (reported in columns (1) to (2)), by excluding all cases where the managerial guidance coincides with the prevailing median analysts' forecast (i.e., conforming cases), and show the respective results in columns (3) to (4).

The coefficient on Negative Guidance is positive and significant in columns (1), implying that the quality of guidance, which is inversely related to the magnitude of differences between guidance and realized earnings, is on average slightly lower on condition that the managerial guidance is negative. We worry that that is driven by outliers, but results in column (2) suggest that it is unlikely. The result, reported in column (3), remains the same, once we exclude conforming cases. The effect becomes insignificant, reported in column (4), if we exclude conforming cases and winsorize at 5%. In any case, the evidence does not favor the conjecture that positive guidance is less reliable. The managerial motives can be complex, often unobservable and unpredictable, which constitutes a source for guidance quality to be unreliable. That is the key motivation for our assumption that guidance quality is uncertain.

"Walk-down to beatable" The literature also documents that managers could have incentives to manage earnings expectation downwards before the earning releases to make it beatable (e.g., Matsumoto, 2002; Cotter et al., 2006; Johnson et al., 2020). Given that one may imagine that more negatively surprised analysts could adjust their forecasts by more to ensure that the firms beat their earnings forecasts. Does this mechanism affect our findings, besides that it contributes to the uncertainty of guidance quality?

To investigate, we rely on Johnson et al. (2020) who construct the expectation management index (EMI) that captures the extent to which firms manage investors' earnings expectations. Specifically, EMI is constructed as a composite score that consists of three broad categories of factors (i.e., "attention," "pressure," and "relevance"), which comprehensively capture firms' incentives to manage expectations.[30] EMI is the first

[30]The first component "attention"refers to the extent to which firms? realized earnings can garner attention from the market. Intuitively, managers are more incentivized to manage market expectations when their earnings performance is covered by more financial analysts and institutional investors. The second component "pressure"captures managerial incentives to soften the negative impact from reporting a break in a string of consecutive earnings increases. The last one, "relevance,"is constructed to capture sensitivity of the market reaction to earnings news. Intuitively, managers are more inclined to manipulate market expectations if their earnings announcement can trigger stronger market reactions.

principal component of these three inputs and a higher value of EMI score indicates stronger incentives for expectation management or driving down earning expectations.

We add EMI as an additional control in our main regressions (reported by Table 1 and 2) and in specifications presented by Figure 2. If such a walk-down-to-beatable mechanism is crucial for our investigations, the estimated coefficients from our regressions should be greatly affected in terms of magnitude and significance. However, we find that all our estimations only change marginally at the best (available upon request), which suggests that our findings are unlikely driven only by managerial strategic guidance.

7. Conclusion

This paper documents a set of cross-sectional facts concerning expectation formation using firm-level earnings forecast and managerial guidance data: the overreaction to information is stronger for unfavorable surprises and weaker for larger surprises, and forecast revisions are asymmetric in surprises and nonmonotonic. We present a model of information uncertainty and smoothed ambiguity aversion to account for these facts. Qualitatively, our model differs from models with extreme ambiguity aversion or those with ambiguity-neutral agents. Quantitatively, we estimate the degree of ambiguity aversion for analysts in this setting and illustrate its role in overreaction to information. Our work adds to the literature that studies expectation formation by documenting new facts and providing new theory.

The empirical setting has instrumental value and will be useful for exploiting other aspects of expectation formation. The empirical strategy in this paper, i.e., the FR-on-surprise approach, can be complementary to the FE-on-FR approach, which is widely used in the literature.

Our results can be also interesting to the literature that studies information production in financial markets. First, sell-side financial analysts constitute a significant share of capital market participants and play an important role by collecting, processing, and conveying relevant information to capital markets. Second, as financial intermediaries, one of the most important functions of financial analysts is to generate forecasts about firms' performance and form expectations. Given the qualitative features of analyst forecasts documented in this paper, it will be interesting to explore how participants in financial markets react to and make use of the information that analysts provide. We leave those topics to future work.

References

Afrouzi, H., S. Kwon, A. Landier, Y. Ma, and D. Thesmar (2022). Overreaction in expectations: Evidence and theory. *Working Paper*.

Amromin, G. and S. A. Sharpe (2014). From the horse's mouth: Economic conditions and investor expectations of risk and return. *Management Science 60*(4), 845–866.

Angeletos, G.-M., Z. Huo, and K. A. Sastry (2020). Imperfect macroeconomic expectations: Evidence and theory. *NBER Macroeconomics Annual 2020, volume 35*.

Angeletos, G.-M. and C. Lian (2016). Incomplete information in macroeconomics: Accommodating frictions in coordination. In *Handbook of macroeconomics*, Volume 2, pp. 1065–1240. Elsevier.

Baginski, S. P. and J. M. Hassell (1990). The market interpretation of management earnings forecasts as a predictor of subsequent financial analyst forecast revision. *Accounting Review*, 175–190.

Baliga, S., E. Hanany, and P. Klibanoff (2013). Polarization and ambiguity. *American Economic Review 103*(7), 3071–83.

Baqaee, D. R. (2020). Asymmetric inflation expectations, downward rigidity of wages, and asymmetric business cycles. *Journal of Monetary Economics 114*, 174–193.

Barber, B. M. and T. Odean (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *The review of financial studies 21*(2), 785–818.

Barrero, J. M. (2022). The micro and macro of managerial beliefs. *Journal of Financial Economics 143*(2), 640–667.

Beyer, A., D. A. Cohen, T. Z. Lys, and B. R. Walther (2010). The financial reporting environment: Review of the recent literature. *Journal of accounting and economics 50*(2-3), 296–343.

Bhattacharya, S. and P. Pfleiderer (1985). Delegated portfolio management. *Journal of Economic Theory 36*(1), 1–25.

Bianchi, V., C. Ilut, and H. Saijo (2022). Diagnostic business cycles. *The Review of Economic Studies (forthcoming)*.

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

Bordalo, P., N. Gennaioli, R. L. Porta, and A. Shleifer (2019). Diagnostic expectations and stock returns. *The Journal of Finance 74*(6), 2839–2874.

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

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

Broer, T. and A. N. Kohlhas (2022, 06). Forecaster (Mis-)Behavior. *The Review of Economics and Statistics*, 1–45.

Capistrán, C. and A. Timmermann (2009). Disagreement and biases in inflation expectations. *Journal of Money, Credit and Banking 41*(2/3), 365–396.

Cerreia-Vioglio, S., F. Maccheroni, and M. Marinacci (2022). Ambiguity aversion and wealth effects. *Journal of Economic Theory 199*, 104898. Symposium Issue on Ambiguity, Robustness, and Model Uncertainty.

Chen, H., Y. K. Lu, and W. Suen (2016). The power of whispers: a theory of rumor, communication, and revolution. *International Economic Review 57*(1), 89–116.

Chen, S., D. Matsumoto, and S. Rajgopal (2011). Is silence golden? an empirical analysis of firms that stop giving quarterly earnings guidance. *Journal of Accounting and Economics 51*(1-2), 134–150.

Chernozhukov, V. and H. Hong (2003). An mcmc approach to classical estimation. *Journal of Econometrics 115*(2), 293–346.

Chuk, E., D. Matsumoto, and G. S. Miller (2013). Assessing methods of identifying management forecasts: Cig vs. researcher collected. *Journal of Accounting and Economics 55*(1), 23–42.

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

Elliott, G., I. Komunjer, and A. Timmermann (2008). Biases in macroeconomic forecasts: Irrationality or asymmetric loss? *Journal of the European Economic Association 6*(1), 122–157.

Elliott, G. and A. Timmermann (2008, March). Economic forecasting. *Journal of Economic Literature 46*(1), 3–56.

Epstein, L. G. and M. Schneider (2008). Ambiguity, information quality, and asset pricing. *The Journal of Finance 63*(1), 197–228.

Farhi, E. and I. Werning (2019). Monetary policy, bounded rationality, and incomplete markets. *American Economic Review 109*(11), 3887–3928.

Farmer, L., E. Nakamura, and J. Steinsson (2021). Learning about the long run. *Working Paper*.

Feng, M. and S. McVay (2010). Analysts' incentives to overweight management guidance when revising their short-term earnings forecasts. *The Accounting Review 85*(5), 1617–1646.

Frost, C. A. (1997). Disclosure policy choices of uk firms receiving modified audit reports. *Journal of Accounting and Economics 23*(2), 163–187.

Gabaix, X. (2020). A behavioral new keynesian model. *American Economic Review 110*(8), 2271–2327.

García-Schmidt, M. and M. Woodford (2019, January). Are low interest rates deflationary? a paradox of perfect-foresight analysis. *American Economic Review 109*(1), 86–120.

Gentzkow, M. and J. M. Shapiro (2006). Media bias and reputation. *Journal of political Economy 114*(2), 280–316.

Gilboa, I. and D. Schmeidler (1989). Maxmin expected utility with non-unique prior. *Journal of mathematical economics 18*(2), 141–153.

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

Hassell, J. M., R. H. Jennings, and D. J. Lasser (1988). Management earnings forecasts: Their usefulness as a source of firm-specific information to security analysts. *Journal of Financial Research 11*(4), 303–319.

Ilut, C. (2012). Ambiguity aversion: Implications for the uncovered interest rate parity puzzle. *American Economic Journal: Macroeconomics 4*(3), 33–65.

Ilut, C. and H. Saijo (2021). Learning, confidence, and business cycles. *Journal of Monetary Economics 117*, 354–376.

Ilut, C. and M. Schneider (2014). Ambiguous business cycles. *American Economic Review 104*(8), 2368–2399.

Kahneman, D. and A. Tversky (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology 3*(3), 430–454.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005). A smooth model of decision making under ambiguity. *Econometrica 73*(6), 1849–1892.

Kohlhas, A. and A. Walther (2021). Asymmetric attention. *American Economic Review (Forthcoming)*.

Lian, C. (2020, 12). A Theory of Narrow Thinking. *The Review of Economic Studies*.

Ma, Y., T. Ropele, D. Sraer, and D. Thesmar (2020). A quantitative analysis of distortions in managerial forecasts. *Working Paper*.

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

Mankiw, N. G. and R. Reis (2002). Sticky information versus sticky prices: a proposal to replace the new keynesian phillips curve. *The Quarterly Journal of Economics 117*(4), 1295–1328.

Maslar, D. A., M. Serfling, and S. Shaikh (2021). Economic downturns and the informativeness of management earnings forecasts. *Journal of Accounting Research 59*(4), 1481–1520.

Morris, S. and H. S. Shin (2002). Social value of public information. *American Economic Review 92*(5), 1521–1534.

Newman, P. and R. Sansing (1993). Disclosure policies with multiple users. *Journal of Accounting research 31*(1), 92–112.

Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *The Review of financial studies 22*(1), 435–480.

Rogers, J. L. and P. C. Stocken (2005). Credibility of management forecasts. *The Accounting Review 80*(4), 1233–1260.

Rogers, J. L. and A. Van Buskirk (2013). Bundled forecasts in empirical accounting research. *Journal of Accounting and Economics 55*(1), 43–65.

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

Skinner, D. J. (1994). Why firms voluntarily disclose bad news. *Journal of accounting research 32*(1), 38–60.

Solomon, D. H. (2012). Selective publicity and stock prices. *The Journal of Finance 67*(2), 599–638.

Wald, A. (1950). Statistical decision functions. *Wiley*.

Wang, C. (2020). Under- and over-reaction in yield curve expectations. *Working paper*.

Woodford, M. (2003). Imperfect common knowledge and the effects of monetary policy. Technical report, National Bureau of Economic Research.

Appendix I: Data and Robustness Tests

A. Summary of Statistics

Table 7. Summary of Statistics

	(1) N	(2) mean	(3) sd	(4) p25	(5) p50	(6) p75
Initial forecasts	110,895	0.0120	0.0129	0.0070	0.0123	0.0180
Revised forecasts	110,895	0.0104	0.0149	0.0057	0.0113	0.0173
Forecast revision	110,895	-0.0016	0.0055	-0.0017	0.0000	0.0000
Forecast errors	110,895	-0.0000	0.0047	0.0000	0.0003	0.0011
Unfavorable	110,895	0.6256	0.4840	0.0000	1.0000	1.0000
Large	110,895	0.0848	0.2785	0.0000	0.0000	0.0000
Surprise	110,895	-0.0040	0.0171	-0.0062	-0.0012	0.0003
Managerial guidance	16,241	0.0067	0.0293	0.0027	0.0089	0.0160
Earnings	16,241	0.0089	0.0197	0.0044	0.0112	0.0177

Table 8. Forecast Error on Forecast Revision: Samples

	Outcome Variable: Forecast Error FE_i					
	Winsorization at the 1% and 99%			Winsorization at the 2.5% and 97.5%		
	Excl Pre-anc	Excl Multiple	Excl Both	Excl Pre-anc	Excl Multiple	Excl Both
	(1)	(2)	(3)	(4)	(5)	(6)
FR_i	-0.0733**	-0.1561***	-0.1545***	-0.0731***	-0.1536***	-0.1540***
	(0.0284)	(0.0217)	(0.0469)	(0.0228)	(0.0171)	(0.0352)
Quarter FEs	YES	YES	YES	YES	YES	YES
Analyst FEs	YES	YES	YES	YES	YES	YES
Firm FEs	YES	YES	YES	YES	YES	YES
Obs.	50,558	46,493	17,606	50,558	46,493	17,606
Adj R-sq.	0.2675	0.3020	0.3412	0.2727	0.2842	0.3285

The standard errors are clustered on firm and calendar year-quarter following Petersen (2009).*** $p<0.01$, ** $p<0.05$, * $p<0.1$

B. Robustness

Overreaction: Subsamples. The result in column (1) of Table 8 is based on a sample excluding all firm-quarters with pre-announcement guidance, which is defined as the guidance issued between firms' fiscal quarter-end and the earnings announcement date for the quarter. The result in column (2) of Table 8 is based on a sample excluding all firm-quarters with multiple guidances. The result in column (3) of Table 8 is based on a sample excluding all firm-quarters with either pre-announcement guidance or multiple guidances. To ensure that our results are not driven by outliers, we winsorize FE_{ijt} and FR_{ijt} at the 2.5% and 97.5% of their respective distributions and re-estimate equation (1). The results for the corresponding subsamples are reported in column (4), (5) and (6).

Table 9. Forecast Error on Forecast Revision: Trimming Outliers

	Outcome Variable: Forecast Error FE_i							
	Trimmed at 1% and 99%				Trimmed at 2.5% and 97.5%			
	Full	Excl Pre-anc	Excl Multiple	Excl Both	Full	Excl Pre-anc	Excl Multiple	Excl Both
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
FR_i	-0.1024***	-0.0942***	-0.1627***	-0.1774***	-0.0854***	-0.0819***	-0.1492***	-0.1568***
	(0.0105)	(0.0208)	(0.0137)	(0.0287)	(0.0082)	(0.0137)	(0.0107)	(0.0186)
Quarter FEs	YES	YES	YES	YES	YES	YES	YES	YES
Analyst FEs	YES	YES	YES	YES	YES	YES	YES	YES
Firm FEs	YES	YES	YES	YES	YES	YES	YES	YES
Obs.	106,614	48,950	43,756	16,738	100,308	46,363	40,148	15,484
Adj R-sq.	0.2250	0.2762	0.2817	0.3336	0.2110	0.2748	0.2654	0.3139

The standard errors are clustered on firm and calendar year-quarter following Petersen (2009).*** p<0.01, ** p<0.05, * p<0.1

Overreaction: Trimming Outliers. In the main text, we estimate equation (1) with winsorized data to mitigate the influence of outlier observations. In this Appendix, we re-estimate equation (1) with trimmed data and examine the robustness of our results reported in the main text. The corresponding results are summarized in Table 9. All results are robust, thus suggesting that our results are not sensitive to the way in which we handle outliers.

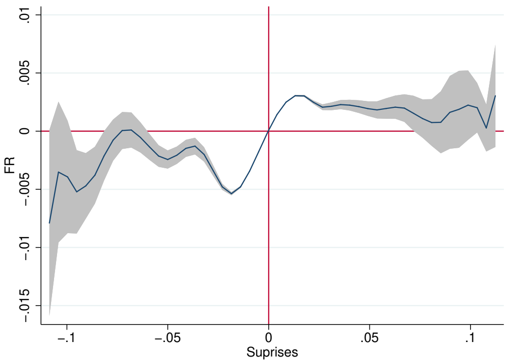
Table 10. Robustness: Definition of Large Surprises

	Outcome Variable: Forecast Revision FR_i			
	Winsorization at 1% and 99%		Winsorization at 2.5% and 97.5%	
	(1)	(2)	(3)	(4)
Surprise_i	0.4311***	0.3971***	0.4575***	0.4193***
	(0.0188)	(0.0184)	(0.0162)	(0.0169)
Large	-0.0060***	-0.0046***	-0.0020***	-0.0019***
	(0.0006)	(0.0007)	(0.0003)	(0.0003)
$\text{Surprise}_i \times \text{Large}$	-0.3502***	-0.3203***	-0.2852***	-0.2655***
	(0.0194)	(0.0182)	(0.0167)	(0.0175)
Constant	0.0001	-0.0001	0.0001	-0.0000
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Quarter FEs	YES	YES	YES	YES
Analyst FEs	YES	YES	YES	YES
Firm FEs	YES	YES	YES	YES
Obs	110,895	110,895	110,895	110,895
Adj R-sq.	0.4819	0.4811	0.5019	0.5032

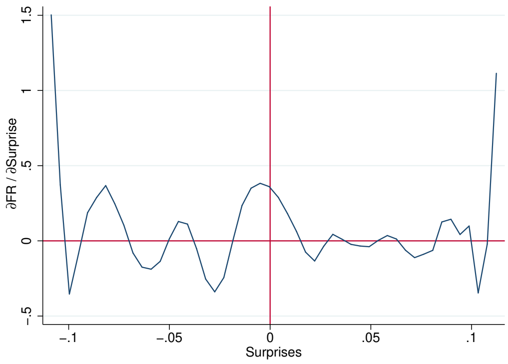
The standard errors are clustered on firm and calendar year-quarter following (Petersen 2009).
*** p<0.01, ** p<0.05, * p<0.1

Definition of large surprises. In the main text, we define a surprise to be large for analyst i, if Surprise_{ijt} is larger than the mean value of the variable Surprise_{ijt} by one standard deviation. However, this definition appears to be arbitrary. In this section, we present the results with alternative definitions. We define a surprise to be large for analyst i, if Surprise_{ijt} is larger than the mean value of the variable Surprise_{ijt} by 1.5 or 2 standard deviations. We estimate equation (4) and report the former in column (1) and the latter in column (2) of Table 10. We repeat the same exercise using the sample in which all variables are winsorized at the 2.5% and 97.5% of their respective distributions. Correspondingly, they are reported in columns (3) and (4) of Table 10, respectively. All estimations present very similar results to those in the main text. The definition of surprises being large does not drive the results reported in the main text.

Trimming at 2%.



(a) Trimming, Non-parametric estimation

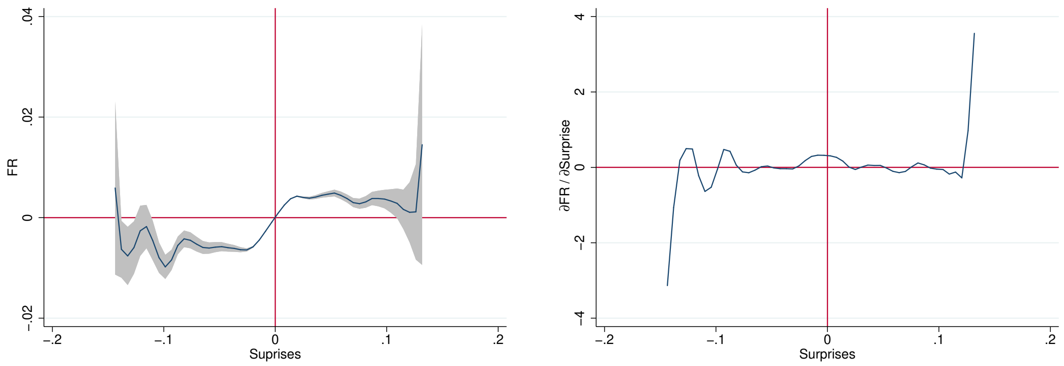


(b) Trimming, Derivative: marginal effect

Figure 10. *Non-parametric estimation, 2% trimming (1%, 99%). Panel (a) illustrates the relationship between forecast revisions and surprises in managerial guidances (both trimmed at 2%) that is non-parametrically estimated using the Epanechnikov kernel and the third degree of the polynomial smooth. It is decreasing, increasing and decreasing and asymmetric around the origin. The shaded areas stand for the 95% confidence intervals for the respective estimations. Panel (b) illustrates its derivatives with respect to surprises. The derivatives are negative when the surprises are large enough and positive when they are small. Forecast revisions respond more strongly to negative surprises than to positive ones of the same magnitude.*

Trimming. In the main text, we estimate the relationship between forecast revisions and surprises using the standard tool of local polynomial regression. In this section, we re-estimate the same relationship by trimming forecast revisions and surprises at the 1% and 99% of their respective distributions and by controlling for quarter, firm and analyst fixed effects. We still rely on the Epanechnikov kernel and the third degree of the polynomial smooth, as in the main text. The result is presented in Figure 10. The pattern is very similar to the one reported in the main text, while the difference is that, as expected, the tails of the estimated relationship are longer with larger confidence intervals. For the trimmed version, there are 1,947 observations to the left of the trough of the estimated relationship and 4,083 observations to the right of the peak.

Winsorization at 2%.

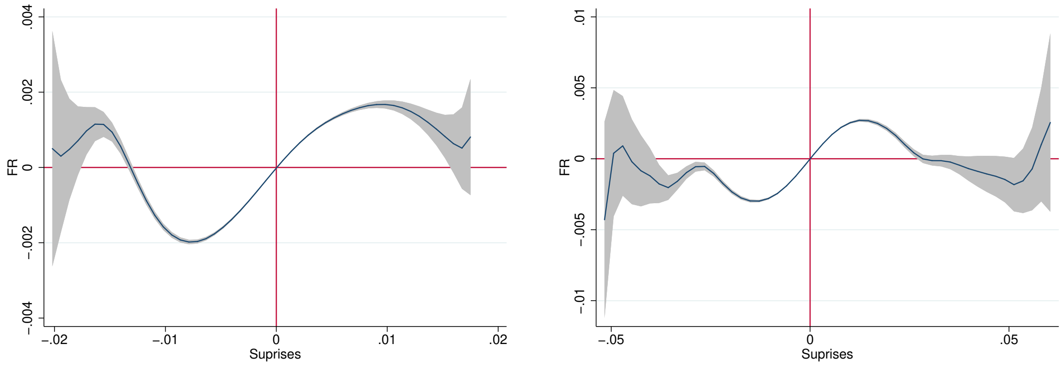


(a) Winsorization, Non-parametric estimation (b) Winsorization, Derivative: marginal effect

Figure 11. *Non-parametric estimation, 2% winsorization and trimming (1%, 99%). Panel (a) illustrates the relationship between forecast revisions and surprises in managerial guidances (both winsorized at 2%) that is non-parametrically estimated using the Epanechnikov kernel and the third degree of the polynomial smooth. It is decreasing, increasing and decreasing and asymmetric around the origin. Panel (b) illustrates its derivatives with respect to surprises. The derivatives are negative when the surprises are large enough and positive when they are small, in the regions where the response of forecast revisions to surprises is significantly different from zero.*

Winsorization. For comparison, we re-estimate the same relationship by winsorizing forecast revisions and surprises at the 1% and 99% of their respective distributions and by controlling for quarter, firm and analyst fixed effects. We still rely on the Epanechnikov kernel and the third degree of the polynomial smooth, as in the main text. The result is presented in Figure 11. The pattern is very similar to the one reported in the main text, while the difference is that, as expected, the tails of the estimated relationship are longer with larger confidences intervals.

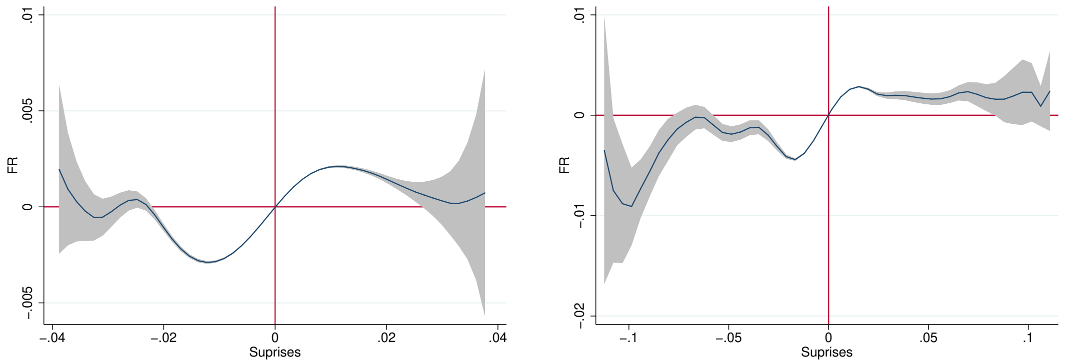
Subsample of Firms where Guidances were Issued at Least 12 Consecutive Guidances.



(a) 5% Triming, Non-parametric estimation (b) 2% Trimming, Non-parametric estimation

Figure 12. *Non-parametric estimation, for the subsample that only includes earnings forecasts on condition that firms release earnings guidances more than 12 consecutive quarters during our sample period. Panel (a) illustrates the relationship between forecast revisions and surprises in managerial guidances, both of which are trimmed at 2.5% and 97.5%. It is non-parametrically estimated using the Epanechnikov kernel and the third degree of the polynomial smooth. It is decreasing, increasing and decreasing and asymmetric around the origin. Panel (b) illustrates the relationship between forecast revisions and surprises in managerial guidances, both of which are trimmed at 1% and 99%, by using the same procedure. The pattern is rather similar.*

Subsample that Excludes the Period of the Financial Crisis.



(a) 5% Trimming, Non-parametric estimation (b) 2% Trimming, Non-parametric estimation

Figure 13. *Non-parametric estimation, for the subsample that excludes observations during the financial crisis. Panel (a) illustrates the relationship between forecast revisions and surprises in managerial guidances, both of which are trimmed at 2.5% and 97.5%. It is non-parametrically estimated using the Epanechnikov kernel and the third degree of the polynomial smooth. It is decreasing, increasing and decreasing and asymmetric around the origin. Panel (b) illustrates the relationship between forecast revisions and surprises in managerial guidances, both of which are trimmed at 1% and 99%, using the same procedure. The pattern is rather similar.*

Heterogeneous Overreaction with Trimming 2% Outlier observations.

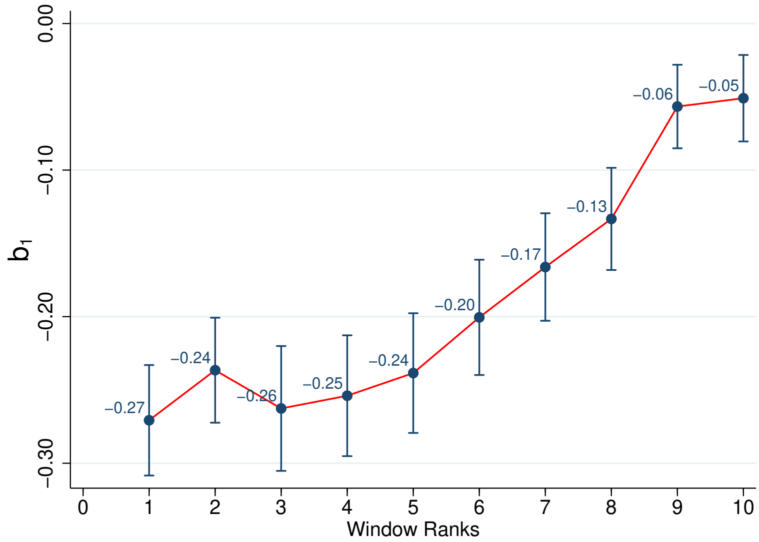


Figure 14. *Heterogeneous Overreaction, Trimming 2% Outliers. The estimated coefficients of the FE-on-FR regressions b_1 and 95% confidence interval for each running decile window is plotted against the window rank. A running decile window j covers decile $j-1$, j, and $j+1$ if $j \neq 1$ or $j \neq 9$; the running decile window 1 covers deciles 1 and 2 and the running decile window 10 covers deciles 9 and 10.*

Heterogeneous Overreaction for Each Decile of Surprises.

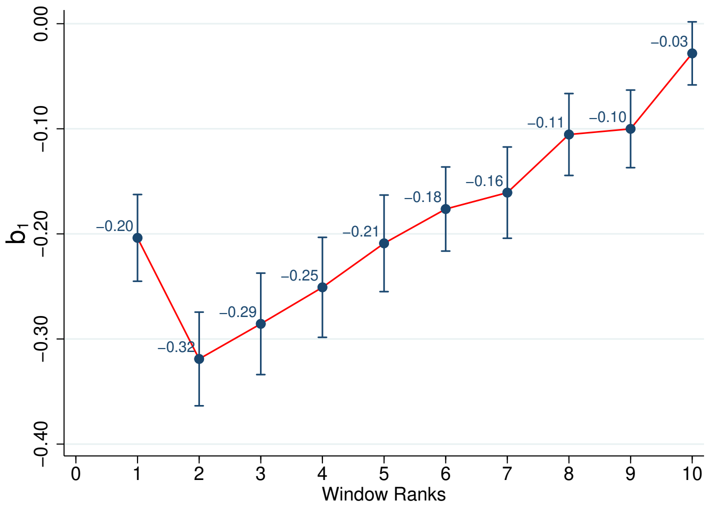
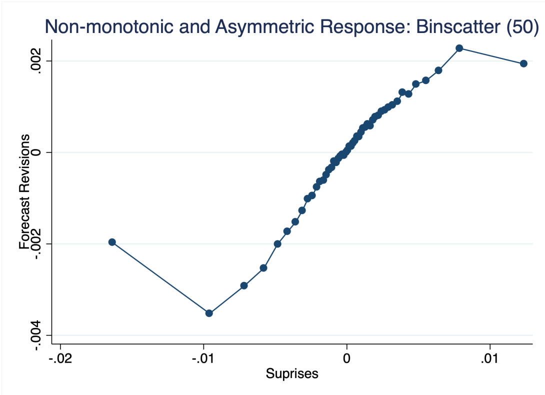
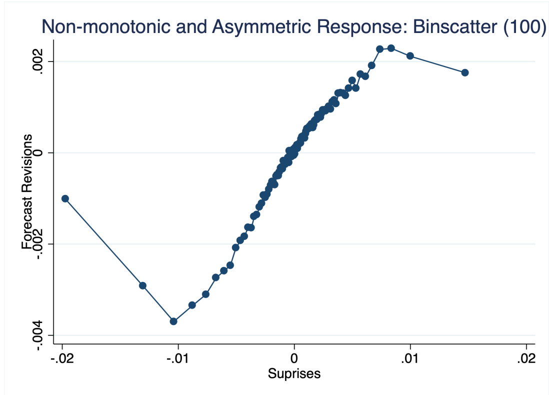


Figure 15. *Heterogeneous Overreaction. The estimated coefficients of the FE-on-FR regressions b_1 and 95% confidence interval for each decile of surprises, without using running windows.*

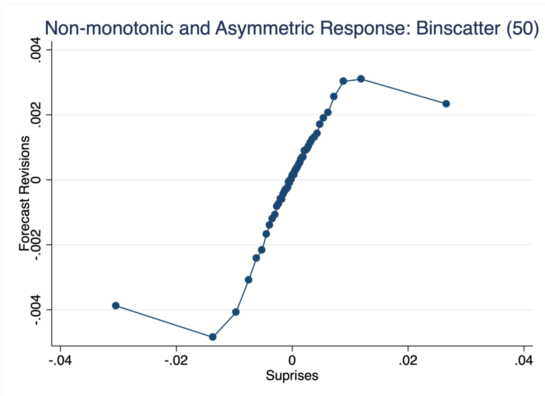
Binscatter Plot.



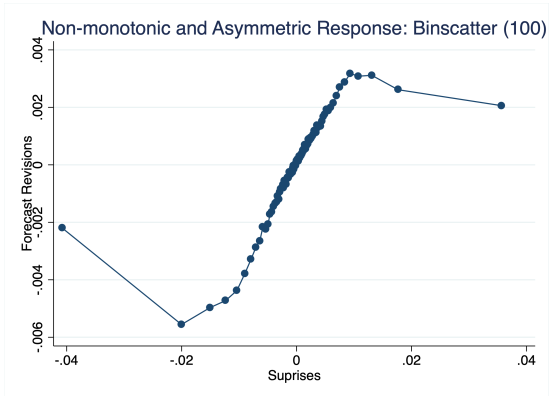
(a) Binscatter (50 bins), trimmed at 5%



(b) Binscatter (100 bins), trimmed at 5%



(c) Binscatter (50 bins), trimmed at 2%



(d) Binscatter (100 bins), trimmed at 2%

Figure 16. *Binscatter Plot, 2% and 5% trimming. Panel (a) illustrates the binscatter plot relationship between forecast revisions and surprises in managerial guidances (both trimmed at 5%), with 50 bins. Panel (b) presents the binscatter plot with the same data and 100 bins. Panel (c) illustrates the binscatter plot relationship between forecast revisions and surprises in managerial guidances (both trimmed at 2%), with 50 bins. Panel (d) presents the binscatter plot with the same data and 100 bins.*

Appendix II: Proofs

Proof of Lemma 1. Forecast revisions being linear in guidance surprises directly follows (9). Zero correlation between forecast errors and forecast revisions follows the fact that noisy information expectation features rationality, in which case forecast errors are uncorrelated with any observables in the information set including forecast revisions. ■

Derivation of Equation (13)-(15). Denote $\delta \equiv \frac{\tau_y}{\tau_\theta+\tau_z+\tau_x+\tau_y}$. Then, it can be shown that

$$\begin{aligned}
&\tilde{p}\left(\tau_y|X_i,s_i;F_i\right)\\
\equiv&\tilde{p}\left(\tau_y|z_{0i},\frac{\tau_\theta+\tau_z+\tau_x}{\tau_x}\left(X_i-F_{0i}\right)+F_{0i},s_i+X_i;F_i\right)\\
=&\tilde{p}\left(\tau_y|z_{0i},x_i,y\right)\\
\propto&\underbrace{\exp\left(-\lambda\left\{-F_i^2+\left(2F_i+\beta\right)\left(X_i+\delta s_i\right)-\left[\left(X_i+\delta s_i\right)^2+\frac{1-\delta}{\tau_\theta+\tau_z+\tau_x}\right]\right\}\right)}_{\phi'\left(\mathbb{E}_i^{\tau_y}\left[U(F_i,\theta)\right]\right)}\\
&\times\underbrace{p\left(F_{0i}\right)p\left(X_i-F_{0i}\right)p\left(s_i|\tau_y\right)}_{=p\left(z_{0i},x_i,y|\tau_y\right)}p\left(\tau_y\right)\\
\propto&\exp\left(-\lambda\left[\left(2F_i+\beta\right)\delta s_i-\left(2X_i\delta s_i+\delta^2 s_i^2-\frac{\delta}{\tau_\theta+\tau_z+\tau_x}\right)\right]\right)p\left(s_i|\tau_y\right)p\left(\tau_y\right)
\end{aligned}$$

where the third line uses the fact that F_{0i}, X_i-F_{0i}, and s_i are independent with only the distribution of s_i affected by τ_y; and the last line drops all terms that are not a function of τ_y. Then, the optimality condition (11) can be compactly written as

$$F_i = X_i + \kappa\left(X_i,s_i,F_i\right)\cdot s_i$$

where

$$\kappa\left(X_i,s_i,F_i\right)=\int_{\Gamma_y}\left(\frac{\tau_y}{\tau_\theta+\tau_z+\tau_x+\tau_y}\right)\tilde{p}\left(\tau_y|X_i,s_i;F_i\right)\mathrm{d}\tau_y$$

■

Proof of Lemma 2. The log-likelihood ratio can be specifically written by:

$$\log\left(L\left(\tau_y\right)\right)=-\lambda s_i\left[2\left(F_i'-F_i\right)\left(\frac{\tau_y}{\tau_\theta+\tau_z+\tau_x+\tau_y}\right)\right]+\text{constant}.$$

Given the fact that $\tau_y/\left(\tau_\theta+\tau_z+\tau_x+\tau_y\right)$ increases in τ_y and that $F_i'-F_i>0$, $L(\tau_y)$

decreases in τ_y, if and only if $s_i > 0$; and $L(\tau_y)$ increases in τ_y, if and only if $s_i < 0$. The lemma is shown. ■

Proof of Proposition 1. The optimality condition (11) is equivalent to (13):

$$F_i = X_i + \underbrace{\left[\int_{\Gamma_y} \left(\frac{\tau_y}{\tau_\theta + \tau_z + \tau_x + \tau_y}\right) \tilde{p}\left(\tau_y | z_{0i}, x_i, y; F_i\right) d\tau_y\right]}_{\kappa} \cdot s_i \tag{21}$$

To obtain the second equality, we use the definition of X_i and s_i and the definition of $\tilde{p}\left(\tau_y | z_{0i}, x_i, y; F_i\right)$ specified in the main text.

We first demonstrate that the right-hand side of (21) is decreasing in F_i. Towards this end, we show

$$\begin{aligned}
\frac{1}{2}\frac{\partial \kappa}{\partial F_i} s_i = & \left\{ \int_{\Gamma_y} \left(\frac{\tau_y}{\tau_\theta + \tau_z + \tau_y + \tau_x}\right) \frac{\phi''\left(\mathbb{E}_i^{\tau_y}\left[U\left(F_i, \theta\right)\right]\right)}{\phi'\left(\mathbb{E}_i^{\tau_y}\left[U\left(F_i, \theta\right)\right]\right)} \frac{\partial \mathbb{E}_i^{\tau_y}\left[U\left(F_i, \theta\right)\right]}{\partial F_i} \tilde{p}\left(\tau_y | z_{0i}, x_i, y; F_i\right) \mathrm{d}\tau_y \right. \\
& \left. - \kappa \left[\int_{\Gamma_y} \frac{\phi''\left(\mathbb{E}_i^{\tau_y}\left[U\left(F_i, \theta\right)\right]\right)}{\phi'\left(\mathbb{E}_i^{\tau_y}\left[U\left(F_i, \theta\right)\right]\right)} \frac{\partial \mathbb{E}_i^{\tau_y}\left[U\left(F_i, \theta\right)\right]}{\partial F_i} \tilde{p}\left(\tau_y | z_{0i}, x_i, y; F_i\right) \mathrm{d}\tau_y \right] \right\} s_i, \\
= & \int_{\Gamma_y} \frac{\phi''\left(\mathbb{E}_i^{\tau_y}\left[U\left(F_i, \theta\right)\right]\right)}{\phi'\left(\mathbb{E}_i^{\tau_y}\left[U\left(F_i, \theta\right)\right]\right)} \left(\frac{\partial \mathbb{E}_i^{\tau_y}\left[U\left(F_i, \theta\right)\right]}{\partial F_i}\right)^2 \tilde{p}\left(\tau_y | z_{0i}, x_i, y; F_i\right) \mathrm{d}\tau_y, \\
& < 0.
\end{aligned}$$

The first equality is obtained by using the definition of κ and the expression of $\partial \tilde{p} / \partial F_i$. That is,

$$\begin{aligned}
& \frac{\partial \tilde{p}\left(\tau_y | z_{0i}, x_i, y; F_i\right)}{\partial F_i} \\
= & \frac{\phi''\left(\mathbb{E}_i^{\tau_y}\left[U\left(F_i, \theta\right)\right]\right)}{\phi'\left(\mathbb{E}_i^{\tau_y}\left[U\left(F_i, \theta\right)\right]\right)} \frac{\partial \mathbb{E}_i^{\tau_y}\left[U\left(F_i, \theta\right)\right]}{\partial F_i} \tilde{p}\left(\tau_y | z_{0i}, x_i, y; F_i\right) \\
& - \tilde{p}\left(\tau_y | z_{0i}, x_i, y; F_i\right) \left[\int_{\Gamma_y} \frac{\phi''\left(\mathbb{E}_i^{\tau_y}\left[U\left(F_i, \theta\right)\right]\right)}{\phi'\left(\mathbb{E}_i^{\tau_y}\left[U\left(F_i, \theta\right)\right]\right)} \frac{\partial \mathbb{E}_i^{\tau_y}\left[U\left(F_i, \theta\right)\right]}{\partial F_i} \tilde{p}\left(\tau_y | z_{0i}, x_i, y; F_i\right) \mathrm{d}\tau_y \right].
\end{aligned}$$

To get the second equality, we use the expression of $\partial \mathbb{E}_i^{\tau_y}\left[U\left(F_i, \theta\right)\right] / \partial F_i$. That is,

$$\frac{\partial \mathbb{E}_i^{\tau_y}\left[U\left(F_i, \theta\right)\right]}{\partial F_i} = \left(\frac{\tau_y}{\tau_\theta + \tau_z + \tau_y + \tau_x} - \kappa\right) s_i.$$

The third inequality holds given $\phi'(\cdot) > 0$ and $\phi''(\cdot) < 0$.

We then notice that κ is bounded between 0 and 1. Therefore, the right-hand side of equation (13) goes to ∞, when F_i approaches $-\infty$; and it goes to $-\infty$ when F_i approaches ∞. Both existence and uniqueness are implied.

Next we show that the optimal response κ^* only depends on s_i. Observe that

$$\begin{aligned}\tilde{p}\left(\tau_y|X_i,s_i;F_i\right) &= \tilde{p}\left(\tau_y|s_i;\kappa\right)\\ &\propto \exp\left(-\lambda\left[\beta\delta s_i+2\kappa\delta s_i^2-\left(\delta^2 s_i^2-\frac{\delta}{\tau_\theta+\tau_z+\tau_x}\right)\right]\right)p\left(s_i|\tau_y\right)p\left(\tau_y\right).\end{aligned}$$

To derive the first equality, we use the equation (13) to replace F_i, and therefore X_i drops out. Therefore, κ^* is the fixed point of the following condition:

$$\kappa^* = \int_{\Gamma_y}\left(\frac{\tau_y}{\tau_\theta+\tau_z+\tau_x+\tau_y}\right)\tilde{p}\left(\tau_y|s_i;\kappa^*\right)\mathrm{d}\tau_y$$

Therefore, it is the case that κ^* is only a function of s_i. ■

Proof of Proposition 2. By using the definition F_i^*, the difference in the expected utilities is explicitly given by:

$$\begin{aligned}&\mathbb{E}^{\tau_y}\left[U\left(F^*\left(X_i,s_i^+\right),\theta\right)\right]-\mathbb{E}^{\tau_y}\left[U\left(F^*\left(X_i,s_i^-\right),\theta\right)\right]\\ =\ &2\beta\delta s_i^+ + \left[\left(\kappa^*\left(s_i^-\right)-\delta\right)^2-\left(\kappa^*\left(s_i^+\right)-\delta\right)^2\right]\left(s_i^+\right)^2.\end{aligned}$$

where $\delta \equiv \tau_y/\left(\tau_\theta+\tau_z+\tau_x+\tau_y\right)$.

Let $T(\beta) \equiv \kappa^*\left(s_i^-\right)-\kappa^*\left(s_i^+\right)$.

Claim 1: If $\beta = 0$, then $T(\beta) = 0$.

We guess and verify that it holds that $\kappa^*\left(s_i^-\right) = \kappa^*\left(s_i^+\right)$. If this is true, we establish that $\mathbb{E}_i^{\tau_y}\left[U\left(F_i,\theta\right)\right]$ is symmetric in s_i: for any τ_y and any pair of $\left(s_i^-,s_i^+\right)$, we have:

$$\mathbb{E}^{\tau_y}\left[U\left(F^*\left(X_i,s_i^+\right),\theta\right)\right] = \mathbb{E}^{\tau_y}\left[U\left(F^*\left(X_i,s_i^-\right),\theta\right)\right]$$

In other words, for any τ_y, we have:

$$\phi'\left(\mathbb{E}_i^{\tau_y}\left[U\left(F^*\left(X_i,s_i^+\right),\theta\right)\right]\right) = \phi'\left(\mathbb{E}_i^{\tau_y}\left[U\left(F^*\left(X_i,s_i^-\right),\theta\right)\right]\right).$$

By the definition of κ, this implies:

$$\kappa^*\left(s_i^-\right) = \kappa^*\left(s_i^+\right).$$

which implies that $\beta = 0$ is a solution to $T(\beta) = 0$. Further, according to Proposition

1, both $\kappa^*\left(s_i^-\right)$ and $\kappa^*\left(s_i^+\right)$ are unique.

Claim 2: If $\beta \neq 0$, $T(\beta) \neq 0$.

Suppose towards a contradiction that there exists some $\beta' > 0$, such that $T(\beta') = 0$. This implies that $\kappa^*\left(s_i^-\right) = \kappa^*\left(s_i^+\right) = \kappa'$. For any pair of $\left(s_i^-, s_i^+\right)$, we have:

$$\frac{\partial \log\left(\frac{\tilde{p}\left(\tau_y|X_i,s_i^-;X_i+\kappa' s_i^-\right)}{\tilde{p}\left(\tau_y|X_i,s_i^+;X_i+\kappa' s_i^+\right)}\right)}{\partial \tau_y} = \lambda\left(\frac{\partial \mathbb{E}_i^{\tau_y}\left[U\left(F^*\left(X_i,s_i^+;\kappa'\right),\theta\right)\right]}{\partial \tau_y} - \frac{\partial \mathbb{E}_i^{\tau_y}\left[U\left(F^*\left(X_i,s_i^-;\kappa'\right),\theta\right)\right]}{\partial \tau_y}\right) > 0.$$

The last inequality is obtained by using the fact that

$$\mathbb{E}_i^{\tau_y}\left[U\left(F^*\left(X_i,s_i^+\right),\theta\right)\right] - \mathbb{E}_i^{\tau_y}\left[U\left(F^*\left(X_i,s_i^-\right),\theta\right)\right] = 2\beta'\delta\left(s_i^+ - s_i^-\right) > 0.$$

In other words, $\tilde{p}\left(\tau_y|X_i,s_i^-;X_i+\kappa' s_i^-\right)$ first-order stochastically dominates $\tilde{p}\left(\tau_y|X_i,s_i^+;X_i+\kappa' s_i^+\right)$. By the definition of κ, this implies:

$$\kappa^*\left(s_i^-\right) > \kappa^*\left(s_i^+\right).$$

A contradiction. Similarly, suppose towards a contradiction that there exists some $\beta' < 0$ such that $T(\beta') = 0$. It implies that $\kappa^*\left(s_i^+\right) > \kappa^*\left(s_i^-\right)$. A contradiction. The claim is shown.

Claim 3: If β goes to ∞, $T(\beta) > 0$.

When β goes to $\rightarrow \infty$, both $\kappa^*\left(s_i^-\right)$ and $\kappa^*\left(s_i^+\right)$ are bounded. Therefore,

$$\mathbb{E}_i^{\tau_y}\left[U\left(F^*\left(X_i,s_i^+\right),\theta\right)\right] - \mathbb{E}_i^{\tau_y}\left[U\left(F^*\left(X_i,s_i^-\right),\theta\right)\right] \rightarrow 2\beta\delta\left(s_i^+ - s_i^-\right) > 0.$$

$\tilde{p}\left(\tau_y|X_i,s_i^-;X_i+\kappa' s_i^-\right)$ first-order stochastically dominates $\tilde{p}\left(\tau_y|X_i,s_i^+;X_i+\kappa' s_i^+\right)$, given $\beta \rightarrow \infty$. Therefore, by the definition of κ, it implies that

$$\kappa^*\left(s_i^-\right) > \kappa^*\left(s_i^+\right).$$

That is, $T(\beta) > 0$. The claim is shown.

Claims 1 and 2 imply that $T(\beta)$ crosses zero once and only at $\beta = 0$. Combined with Claim 3, it further implies that $\beta T(\beta) \geq 0$, where the equality holds only when $\beta = 0$. The proposition is shown. ■

Proof of Proposition 3 . If forecasters are ambiguity neutral, the optimal forecasts are such that

$$F_i^* = X_i + \left[\int_{\Gamma_y} \delta p\left(\tau_y|s_i\right) \mathrm{d}s_i\right] s_i$$

where $\delta \equiv \tau_y/(\tau_\theta + \tau_z + \tau_x + \tau_y)$ and the posterior belief $p\left(\tau_y|s_i\right)$ is given by

$$p\left(\tau_y|s_i\right) \propto \sqrt{\delta}\exp\left(-\frac{1}{2}\left(\tau_\theta + \tau_z + \tau_x\right) s_i^2 \delta\right) p\left(\tau_y\right),$$

Taking the derivative of F_i^* w.r.t s_i leads to

$$\frac{\partial F_i^*}{\partial s_i} = \int_{\Gamma_y} \delta p\left(\tau_y|s_i\right) ds_i - \left(\tau_\theta + \tau_z + \tau_x\right) \mathbb{V}\left(\delta|s_i\right) s_i^2$$

where $\mathbb{V}\left(\delta|s_i\right)$ denotes the conditional volatility of δ under probability density $p\left(\tau_y|s_i\right)$.

It is then straightforward to show that:

$$\lim_{|s_i|\to 0} \frac{\partial F_i^*}{\partial s_i} = \lim_{|s_i|\to 0} \int_{\Gamma_y} \delta p\left(\tau_y|s_i\right) \mathrm{d}s_i > 0$$

Furthermore, when $|s_i| \to +\infty$, $p\left(\tau_y|s_i\right)$ converges to $p_\infty\left(\tau_y\right)$ and is given by:

$$p_\infty\left(\tau_y\right) \propto \sqrt{\delta} p\left(\tau_y\right)$$

Then it must be the case that $\lim_{|s_i|\to+\infty} \mathbb{V}\left(\delta|s_i\right) s_i^2 \to +\infty$. Further using the fact that $\int_{\Gamma_y} \delta p\left(\tau_y|s_i\right) \mathrm{d}s_i$ is bounded above by $\delta_{\max}$, it is straightforward to demonstrate that

$$\lim_{|s_i|\to+\infty} \frac{\partial F_i^*}{\partial s_i} \to -\infty.$$

Finally, the symmetry of $F_i^* - X_i$ around the origin directly follows from the fact that $\int_{\Gamma_y} \delta p\left(\tau_y|s_i\right) \mathrm{d}s_i$ is symmetric, since $p\left(\tau_y|s_i\right) = p\left(\tau_y|-s_i\right)$ for $\forall s_i \in \mathbb{R}$.

■

Proof of Proposition 4 . The objective function (6) under the maxmin criterion becomes:

$$\max_{F\in\mathbb{R}} \min_{\tau_y\in\Gamma_y} \mathbb{E}\left[-\left(F-\theta\right)^2 + \beta\theta|z_i, x_i, y; \tau_y\right]$$

where Γ_y is the full support for τ_y. Let the upper bound be $\tau_y^{\max}$ and the lower bound be $\tau_y^{\min}$. For ease of notation, denote the subjective relative precision of guidance to be

$$\delta \equiv \frac{\tau_y}{\tau_\theta + \tau_z + \tau_x + \tau_y}.$$

and accordingly, it is bounded by

$$\delta_{\min} \equiv \frac{\tau_y^{\min}}{\tau_\theta + \tau_z + \tau_x + \tau_y^{\min}} \quad \text{and} \quad \delta_{\max} \equiv \frac{\tau_y^{\max}}{\tau_\theta + \tau_z + \tau_x + \tau_y^{\max}}.$$

To prove the proposition, we first characterize the optimal forecasting rule under the maxmin criterion. Then, we proceed to prove that $F_i^* - X_i$ is non-decreasing in s_i.

First of all, it can be shown that

$$\bar{\theta}_{\tau_y} = X_i + \delta s_i \qquad \mathbb{E}_i \left[\theta^2 | z_i, x_i, y; \tau_y\right] = (X_i + \delta s_i)^2 + (1 - \delta) \left(\frac{1}{\tau_\theta + \tau_z + \tau_x}\right)$$

Then, the problem can be transformed into

$$\max_{\kappa \in \mathbb{R}} \min_{\delta \in \Delta} V(\kappa, \delta)$$

where $\Delta \equiv [\delta_{\min}, \delta_{\max}]$ and the value function $V(\kappa, \delta)$ is given by

$$V(\kappa, \delta) \equiv -(X_i + \kappa s_i)^2 + [2(X_i + \kappa s_i) + \beta](X_i + \delta s_i) - \left[(X_i + \delta s_i)^2 + (1 - \delta)\frac{1}{\tau_\theta + \tau_z + \tau_x}\right]$$

where we have used the fact that $F = X_i + \kappa s_i$. Notice that $V(\kappa, \delta)$ is quadratic in κ and δ. Also note that $V(\kappa, \delta)$ is concave in δ. Therefore, we have that for any $\kappa \in \mathbb{R}$:

$$\operatorname*{argmin}_{\delta \in \Delta} V(\kappa, \delta) \in \{\delta_{\min}, \delta_{\max}\}$$

Notice that

$$\begin{aligned}
&V(\kappa, \delta_{\max}) - V(\kappa, \delta_{\min}) \\
&= (2\kappa s_i + \beta) s_i (\delta_{\max} - \delta_{\min}) + \frac{1}{\tau_\theta + \tau_z + \tau_x}(\delta_{\max} - \delta_{\min}) - s_i^2 \left(\delta_{\max}^2 - \delta_{\min}^2\right)
\end{aligned}$$

It can then be shown that

$$\begin{aligned}
&V(\kappa, \delta_{\max}) - V(\kappa, \delta_{\min}) > 0 \\
\Leftrightarrow &\kappa > T(s_i) \equiv \frac{(\delta_{\max} + \delta_{\min})}{2} - \frac{\beta s_i + \frac{1}{\tau_\theta + \tau_z + \tau_x}}{2 s_i^2}
\end{aligned}$$

In what follows, we characterize the optimal forecasting rule for three exclusive cases:

- If $\delta_{\min} > T(s_i)$, it can be shown that
  - when $\kappa \in (-\infty, T(s_i)]$, $\min_{\delta\in\Delta} V(\kappa,\delta) = V(\kappa,\delta_{\max})$. Hence, $\min_{\delta\in\Delta} V(\kappa,\delta)$ is increasing in κ.
  - when $\kappa > T(s_i)$, $\min_{\delta\in\Delta} V(\kappa,\delta) = V(\kappa,\delta_{\min})$. Hence, $\min_{\delta\in\Delta} V(\kappa,\delta)$ is first increasing in κ and then decreasing in κ. It achieves its maximum at $\kappa = \delta_{\min}$.

  Figure 17(a) graphically illustrates the value function under the worst case scenario when $\delta_{\max} < T(s_i)$. Therefore, it must be the case that the optimal $\kappa^* = \delta_{\min}$ when $\delta_{\min} > T(s_i)$.

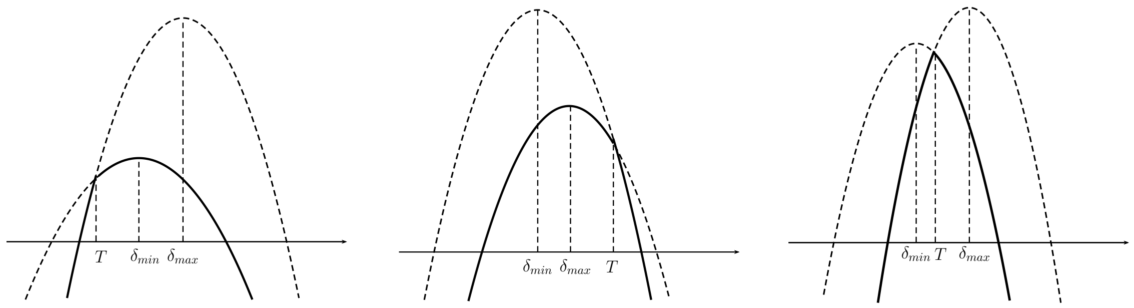
- If $\delta_{\max} < T(s_i)$, it can be shown that
  - when $\kappa \in (-\infty, T(s_i)]$, $\min_{\delta\in\Delta} V(\kappa,\delta) = V(\kappa,\delta_{\max})$. Hence, $\min_{\delta\in\Delta} V(\kappa,\delta)$ is first increasing in κ and then decreasing in κ. It achieves its maximum at $\kappa = \delta_{\max}$.
  - when $\kappa \in [T(s_i), +\infty)$, $\min_{\delta\in\Delta} V(\kappa,\delta) = V(\kappa,\delta_{\min})$. Hence, $\min_{\delta\in\Delta} V(\kappa,\delta)$ is decreasing in κ.

  Figure 17(b) graphically illustrates the value function under the worst case scenario when $\delta_{\max} < T(s_i)$. Therefore, it must be the case that the optimal $\kappa^* = \delta_{\max}$ when $\delta_{\max} < T(s_i)$.

- If $\delta_{\min} < T(s_i) < \delta_{\max}$, it is then straightforward to show the following:
  - when $\kappa \in (-\infty, T(s_i)]$, $\min_{\delta\in\Delta} V(F,\delta) = V(F,\delta_{\max})$. Hence, $\min_{\delta\in\Delta} V(F,\delta)$ is increasing in κ.
  - when $\kappa \in [T(s_i), +\infty)$, $\min_{\delta\in\Delta} V(F,\delta) = V(F,\delta_{\min})$. Hence, $\min_{\delta\in\Delta} V(F,\delta)$ is decreasing in κ.

  Figure 17(c) graphically illustrates the value function under the worst case scenario when $\delta_{\min} < T(s_i) < \delta_{\max}$. Therefore, it must be the case that the optimal $\kappa^* = T(s_i)$ when $\delta_{\min} < T(s_i) < \delta_{\max}$.

To summarize, we have the following optimal forecasting rule under the maxmin



(a) $\delta_{\text{min}} > T(s_i)$ (b) $\delta_{\text{max}} < T(s_i)$ (c) $\delta_{\text{min}} < T(s_i) < \delta_{\text{max}}$

Figure 17. *The value function under the worst case scenario:* $\min_{\tau_y \in \Gamma_y} V(\kappa, \delta)$.

criterion:

$$\kappa^* = \begin{cases} \delta_{\text{min}} & \text{if } \delta_{\text{min}} > T(s_i) \\ \delta_{\text{max}} & \text{if } \delta_{\text{max}} < T(s_i) \\ T(s_i) & \text{otherwise} \end{cases} \tag{22}$$

Or equivalently,

$$F^* - X_i = \begin{cases} \delta_{\text{min}} s_i & \text{if } \delta_{\text{min}} > T(s_i) \\ \delta_{\text{max}} s_i & \text{if } \delta_{\text{max}} < T(s_i) \\ T(s_i) s_i & \text{otherwise} \end{cases} \tag{23}$$

Note that $T(s_i) s_i$ is always increasing in s_i. Therefore, given the continuity of $F_i^* - X_i$ with respect to s_i, it must be the case that $F_i^* - X_i$ is non-decreasing in s_i.

■

Proof of Proposition 5 . It can be shown that

$$\begin{aligned} \text{Cov}(\text{FE}_i, \text{FR}_i) =& \text{Cov}\left(\kappa^{\text{RE}} s_i, \kappa(s_i) s_i\right) - \text{Var}(\kappa(s_i) s_i) \\ =& \text{Cov}\left(\kappa^{\text{RE}} s_i - \kappa(s_i) s_i, \kappa(s_i) s_i\right) \\ =& \mathbb{E}\left[\left(\kappa^{\text{RE}} s_i - \kappa(s_i) s_i\right) \kappa(s_i) s_i\right] - \mathbb{E}\left[\kappa^{\text{RE}} s_i - \kappa(s_i) s_i\right] \mathbb{E}[\kappa(s_i) s_i] \\ =& \kappa^{\text{RE}} \mathbb{E}\left[\kappa(s_i) s_i^2\right] - \mathbb{E}\left[\kappa^2(s_i) s_i^2\right] \end{aligned}$$

where the third equality uses the fact that $\mathbb{E}[\kappa(s_i) s_i] = 0$ under symmetry of $\kappa(s_i)$ when $\lambda = 0$.

It is then straightforward to show that

$$\operatorname{sgn}\left\{\hat{b}_1\right\} = \operatorname{sgn}\left\{\operatorname{Cov}\left(\mathrm{FE}_i, \mathrm{FR}_i\right)\right\} = \operatorname{sgn}\left\{\kappa^{\mathrm{RE}} - \frac{\mathbb{E}\left[\kappa^2\left(s_i\right)s_i^2\right]}{\mathbb{E}\left[\kappa\left(s_i\right)s_i^2\right]}\right\}$$

Notice that $\frac{\mathbb{E}\left[\kappa^2(s_i)s_i^2\right]}{\mathbb{E}\left[\kappa(s_i)s_i^2\right]}$ is nothing more than an average of $\kappa\left(s_i\right)$ over some adjusted beliefs of s_i:

$$\frac{\mathbb{E}\left[\kappa^2\left(s_i\right)s_i^2\right]}{\mathbb{E}\left[\kappa\left(s_i\right)s_i^2\right]} = \hat{\mathbb{E}}\left[\kappa\left(s_i\right)\right] \equiv \int_{\mathbb{R}} \kappa\left(s_i\right)\hat{p}\left(s_i\right) ds_i$$

where

$$\hat{p}\left(s_i\right) \propto \Omega\left(s_i\right)p\left(s_i\right) \qquad\qquad \Omega\left(s_i\right) \equiv \frac{\kappa\left(s_i\right)s_i^2}{\mathbb{E}\left[\kappa\left(s_i\right)s_i^2\right]}$$

■

Proof of Proposition 6. The optimal forecast F_i^L is globally monotone in the signal is a sufficient condition for the global monotonicity. According to (19), the optimal forecasts can be expressed as

$$\mathrm{F}_i^L = \mathbb{E}\left[\theta|\mathcal{I}_i\right] + \frac{1-2\alpha}{\alpha}\int_{-\infty}^{\mathrm{F}_i^L}\left(\theta - \mathrm{F}_i^L\right) f\left(\theta|\mathcal{I}_i\right) \mathrm{d}\theta \tag{24}$$

Assume that $\mathcal{I}_i = \{x_i\}$ with $x_i \sim N\left(\theta, \sigma_x^2\right)$, where the fundamental $\theta \sim N\left(0, \sigma_\theta^2\right)$.

In what follows, we prove that $\frac{\mathrm{dF}_i^L}{\mathrm{d}x_i} > 0$. Taking total derivative w.r.t x_i on both sides of (24) and re-arranging leads to

$$\begin{aligned}\frac{\mathrm{dF}_i^L}{\mathrm{d}x_i} =& \frac{\mathrm{d}\mathbb{E}\left[\theta|x_i\right]}{\mathrm{d}x_i} - \frac{1-2\alpha}{\alpha}\mathrm{P}\left(\theta < \mathrm{F}_i^L \middle| x_i\right)\frac{\mathrm{dF}_i^L}{\mathrm{d}x_i} \\ &+ \frac{1-2\alpha}{\alpha}\left(\frac{1}{\sigma_\theta^2} + \frac{1}{\sigma_x^2}\right)\int_{-\infty}^{\mathrm{F}_i^L}\left(\theta - \mathrm{F}_i^L\right)\left(\theta - \mathbb{E}\left[\theta|x_i\right]\right) f\left(\theta|x_i\right)\mathrm{d}\theta\end{aligned}$$

where $\mathrm{P}\left(\theta < \mathrm{F}_i^L \middle| x_i\right)$ denotes the conditional probability of negative forecast error. Using the fact that $\mathrm{F}_i^L < \mathbb{E}\left[\theta|x_i\right]$, we can prove that

$$\frac{1-2\alpha}{\alpha}\left(\frac{1}{\sigma_\theta^2} + \frac{1}{\sigma_x^2}\right)\int_{-\infty}^{\mathrm{F}_i^L}\left(\theta - \mathrm{F}_i^L\right)\left(\theta - \mathbb{E}\left[\theta|x_i\right]\right) f\left(\theta|x_i\right)\mathrm{d}\theta > 0$$

Combined with the fact that $\frac{\mathrm{d}\mathbb{E}[\theta|x_i]}{\mathrm{d}x_i} > 0$, it is straight-forward to see that

$$\frac{\mathrm{dF}_i^L}{\mathrm{d}x_i} > 0$$

■