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

January 25, 2022

Abstract. Using firm-level earnings forecast 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) that analysts overreact to guidance surprises; (ii) that the overreaction is stronger to negative guidance surprises but weaker to larger surprises; and (iii) that forecast revisions are neither symmetric in guidance surprises nor monotonic. We organize the facts with a model where analysts are uncertain about the quality of managerial guidance. A structural estimation shows 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, non-monotone, ambiguity aversion

JEL Classification. C53, D83, D84, E31

*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; 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. There is a large literature that examines the standard assumption on expectation formation in macroeconomics, i.e., the full information rational expectation hypothesis (FIRE), using survey data on households, firm managers, financial analysts, and professional forecasters (Coibion and Gorodnichenko 2015). It is a consensus that expectation formation indeed deviates from FIRE. While it is documented that individuals in general overreact to information (Bordalo, Gennaioli, Ma, and Shleifer 2020), there has been a growing interest in the circumstances under which the overreaction is stronger or weaker and whether an underreaction can arise (Afrouzi, Kwon, Landier, Ma, and Thesmar 2021; Angeletos, Huo, and Sastry 2020; Kohlhas and Walther 2021).

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 information whose quality is uncertain. It is commonplace that forecasters receive noisy economic or financial data to make forecasts but they are not certain about the data quality. The way they react in such a scenario has so far received little attention in the expectation formation literature but will be scrutinized in this paper.

Empirically, we explore a setting where financial analysts forecast the earnings of firms, firms release managerial guidance for earnings and then analysts update their earnings forecasts. Forecast revisions are defined to be the differences between analysts' updated forecasts after and their initial forecasts before managerial guidance. Therefore, forecast revisions are constructed to reflect the impact of the guidance. With 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, suggesting that upward revisions can predict earnings being below updated forecasts. This result is consistent with the findings of Bordalo, Gennaioli, Ma, and Shleifer (2020) using macroeconomic survey data.

Second, the overreaction is heterogeneous across analysts, depending on differences between the common managerial guidance and analysts' initial forecasts, i.e., guidance surprises (constructed at the firm-quarter-analyst level). We rank those surprises from the most negative to the most positive and break them into deciles. Estimating the degree of overreaction in each subsample, we find that the overreaction is stronger when the surprises are negative; and the overreaction tends to be weaker when the surprises are larger in size: the degree of the overreaction increases and then decreases in surprises.

Third, we directly explore how forecast revisions respond to guidance surprises with non-parametric estimations. We find that forecast revisions are asymmetric in surprises: the forecasts revisions are stronger when the surprises are negative than when the surprises are of the same magnitude but positive. Further, forecast revisions are not monotonically increasing in surprises either: when the surprises are large enough, the forecast revisions decrease in surprises.

This setting has unique advantages. First, the same managerial guidance delivers different surprises to analysts with different information sets. Therefore, it offers a clean environment for our empirical exploration for heterogeneous overreaction across individuals. Second, managerial guidance, the most important part of information flow in this setting, is observable. For each analyst, the surprise contained in the guidance can be constructed with her initial forecast prior to the guidance. This enables us to measure the characteristics of surprises, such as magnitude and favorability, and explore their impacts.[1] 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.

The new evidence on forecast revisions calls for a theory. In section 3, we build a forecasting model where analysts have access to both private information about the earnings of a firm and managerial guidance for earnings from the firm. The key departure from standard forecasting models is (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, 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.

[1] An ideal testing ground for the relationship between forecast revisions and new information is one in which (i) the information flow acquired by agents is (at least partly) observable, (ii) the new information contained in such an information flow is measurable, and (iii) the agent's response to new information is available. The empirical environment that we consider is fairly close to the ideal.

[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).

We incorporate analysts' aversion towards 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 overweight the states of the world where their expected utility is low and underweight the states where their expected utility is high. Suppose 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 surprises than to unfavorable ones, 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, so as to simultaneously capture both non-monotonicity 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 (max-min), 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 a 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 to our estimations. Our quantitative exercises also reveal that part of the overreaction to information found in the data is driven by analysts' ambiguity aversion. Further, the cross-sectional pattern helps distinguish our theory from alternative hypotheses, such as agency theory and loss aversion, which will be discussed in section 5.5.

Our theory underlies the role of uncertain information quality in organizing new facts regarding expectation formation. The flip side is that once the uncertainty is very low, analysts' forecast revisions are almost linear in guidance surprises. We show in section 5.4 that this auxiliary prediction is empirically supported in a subsample where we only include firms that deliver (ex post) very precise earnings guidance.

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 friction and expectation formation. Coibion and Gorodnichenko (2015) provide a new empirical methodology towards this issue in that 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 forecast and the initial one. Under the full information rational expectations (FIRE) assumption, the coefficient is to be zero. Thus, a statistically significant coefficient suggests a departure from this. Coibion 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 Kohlhas and Broer (2019) with macroeconomic survey data. The same approach is applied to firm earnings forecasts 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 under-reaction 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 only focusing on the average behavior of the whole sample, we deal with differences across groups: analysts who are positively surprised v.s. those who are negatively surprised; analysts who are more surprised v.s. those who are less surprised. Further, we provide a complementary empirical approach of "FR-on-Surprise," which directly explores the relationship between forecast revisions and observable new information, and which can be a useful tool for the literature. It is worthwhile to highlight that we construct a novel and non-trivial environment to study expectation formation, which is useful for other related research in the this area.[4]

[3]Other recent studies also provide evidence on the forecasts of financial market participants, such as Amromin and Sharpe (2014), Barrero (2021), 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.

Afrouzi, Kwon, Landier, Ma, and Thesmar (2021) investigate how expectation biases vary. In an experimental setting, they 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 (i.e., analysts forecasting the earnings of the current quarter), which differs from issues related to forecast horizons and persistence. These two works are complementary for understanding the determinants of overreaction to information.

The building blocks of our model have precedents in the literature of 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 Epstein and Schneider (2008) and Baqaee (2020) characterize the process of expectation formation when agents have an extreme ambiguity averse preference (max-min) 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 allows for uncertainty in information quality. 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 contribution is relevant and necessary. Models that feature a constant ambiguity aversion become common in the recent literature and Baliga, Hanany, and Klibanoff (2013) is one such example.

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 belief (Bordalo, Gennaioli, Ma, and Shleifer 2020), over-confidence (Kohlhas and Broer 2019), cognitive discounting (Gabaix 2020), level-K thinking (García-Schmidt and Woodford 2019, Farhi and Werning 2019), narrow thinking (Lian 2018), autocorrelation averaging (Wang 2020) and loss aversion (Capistrán and Timmermann 2009). The most recent studies combine both, such as over-extrapolation with dispersed information (Angeletos, Huo,

[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 asymmetrical response to inflationary and disinflationary news.

and Sastry 2020).

A common feature of the aforementioned theories is that the response of forecast revision to new information (or surprises) is constant and symmetric: the direction of surprises does not matter; and forecast revisions are monotonically increasing in surprises. Our model differs in both aspects.

2. Evidence

2.1. Data and Sample

In this section, we explore how analysts revise their earnings forecasts upon newly received information. Our goal is to construct a scenario where part of the information flow is observable, measurable and systematically accessible to econometricians. We focus on managerial guidance, which is among very few information sources that satisfy such criteria. 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 covers managerial guidance from more than 6,000 companies within North America that can date back to as early as 1994. Furthermore, the I/B/E/S Guidance data is offered and presented on the same accounting basis as the I/B/E/S Estimates that provide individual analysts' forecasts 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 the entire 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 from 1994 as this is the first year from which the I/B/E/S systematically collects information on managerial guidance.[6] We only keep closed-end managerial guidance, including point and range forecasts, in order 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 new information from firm management, we exclude all managerial guidance bundled with

[6]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's CIG database. Second, the focus of this paper is to understand how analysts update their beliefs, given new information, i.e., the 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 is more likely to be collected for firms with analyst coverage. Fourth, our results are robust to starting the sample period from 1998, after which the coverage bias has been shown to be relatively small.

earnings announcements.[7] We only deal with the 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 keep the latest guidance before the current quarter earnings announcement. However, our results are not sensitive to this specific choice and they are qualitatively the same 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 is also obtained from the I/B/E/S Estimates.

Finally, in this data, the initial analyst forecasts are defined and constructed by individual analyst forecasts that are issued after the prior quarter earnings announcement date and that are the most updated ones before the earnings guidance. The revised analyst forecasts are defined as those issued by the same set of analysts on or right after the earnings guidance date. For analysts who initially offer forecasts but provide no forecast revisions until the earnings announcement, we assume that their revised forecasts remain the same as the initial ones, a practice consistent with prior literature (Feng and McVay 2010; Maslar, Serfling, and Shaikh 2021). Because the EPS estimates can vary substantially across firms, we deflate them by the stock price at the beginning of the quarter using data retrieved from the CRSP. To avoid the small price deflator problem that may distort the distribution, we exclude observations with a stock price of less than one dollar.

The sample construction procedure can be better apprehended with the aid of Figure 1, which delineates the sequence of major events.[8] Our full sample consists of

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

[8] 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

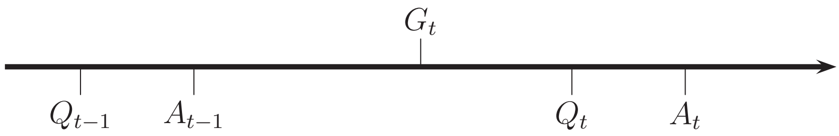


Figure 1. *Timeline. We consider managerial guidance G_t issued between A_{t-1} and A_t. If the guidance for EPS at the quarter t is released at the date of A_{t-1} or within two days after A_{t-1}, then it is bundled. If the guidance is released between Q_t and A_t, it is pre-announcement. If more than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

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

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. Towards 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 the 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% of their respective distributions. We estimate the following equation:

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

announcement date A_{t-1} for quarter $t-1$ would also happen after Q_{t-1}. In this paper, we retrieve an earnings guidance that is issued by firm management at any date between A_{t-1} and A_t. Because an increasing number of firms bundle their earnings projections for quarter t together 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 an earnings guidance G_t, we can accordingly identify the sequence of analysts' earnings forecasts for the same quarter. We define analysts' forecasts which 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 one. 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 as pre-announcement 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 keep 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%			
	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	-0.0952***	-0.0733**	-0.1561***	-0.1545***	-0.0926***	-0.0731***	-0.1536***	-0.1540***
	(0.0146)	(0.0284)	(0.0217)	(0.0469)	(0.0119)	(0.0228)	(0.0171)	(0.0352)
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.	110,895	50,558	46,493	17,606	110,895	50,558	46,493	17,606
Adj R-sq.	0.2429	0.2675	0.3020	0.3412	0.2298	0.2727	0.2842	0.3285

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

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 specifics as well as time-series differences are absorbed and cannot explain our results. The standard errors are clustered on firm and calendar year-quarter to adjust for both inter-temporal and cross-sectional correlations, following Petersen (2009). The result from estimating equation (1) is 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 Kohlhas and Broer (2019). 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 perform robustness checks with different subsamples. The result in column (2) 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 (3) is based on a sample excluding all firm-quarters with multiple guidances. The result in column (4) 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 the scaled 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 full sample and the corresponding subsamples are reported in column (5), (6), (7) and (8). To ensure consistency with the results estimated locally (see sections below), we estimate equation (1) after we trim outliers of the sample and present those results in Table 6 of Appendix I.B. The estimated coefficients in the aforementioned exercises are qualitatively the same 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 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 allows 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 the 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 at quarter t and analyst i's corresponding initial forecast (i.e., F_{0ijt}) for firm j at quarter t before guidance. That is, $\text{Surprise}_{ijt} \equiv G_{jt} - F_{0ijt}$.[9] 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% of their respective distributions (to be consistent with the non-parametric estimations in the next section). We then rank surprises from the most negative to the most positive, break 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) the window j covers decile $j-1$, j, and $j+1$ if $j \neq 1$ or $j \neq 10$; (2) the running decile window 1 covers deciles 1 and 2; and (3) the 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 it 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.

[9] We stress the fact that the constructed surprise variable in managerial guidance is the one that 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.

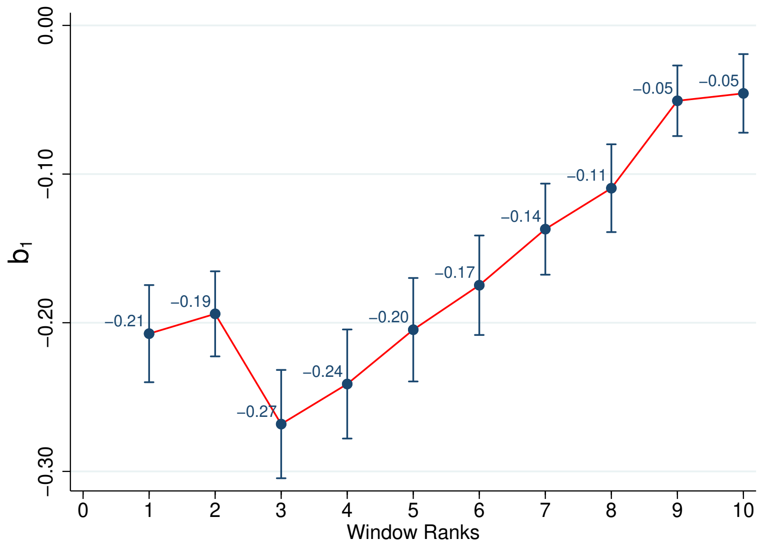


Figure 2. *Heterogeneous Overreaction. The estimated coefficients of the FE-on-FR regressions b_1 and 95% confidence interval for each running decile window are plotted against the window rank. A running decile window j covers decile $j-1$, j, and $j+1$ if $j \neq 1$ or $j \neq 9$; the running decile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

To examine whether our results are robust, we re-do the exercises with a sample where forecast errors, forecast revisions and surprises are trimmed at the 1% and 99% 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).

In summary, one the one hand, we confirm that analysts overreact to information 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. Towards this end, we directly investigate this relationship between forecast revisions and 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,

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

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)	
$\text{Surprise}_i \times \text{Unf}$		0.1130***			0.0846***	
		(0.0231)			(0.0286)	
Large			-0.0055***			-0.0016***
			(0.0005)			(0.0003)
$\text{Surprise}_i \times \text{Large}$			-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). *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

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 result is reported in column (1) of Table 2.[10] The significantly positive coefficient of 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. Towards 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.[11] 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-

[10] 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.

[11] About 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 ones to be conservative. However, our results remain qualitatively unchanged, if we exclude these confirming cases.

side of equation (2) and estimate the following regression:

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

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 result. 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:

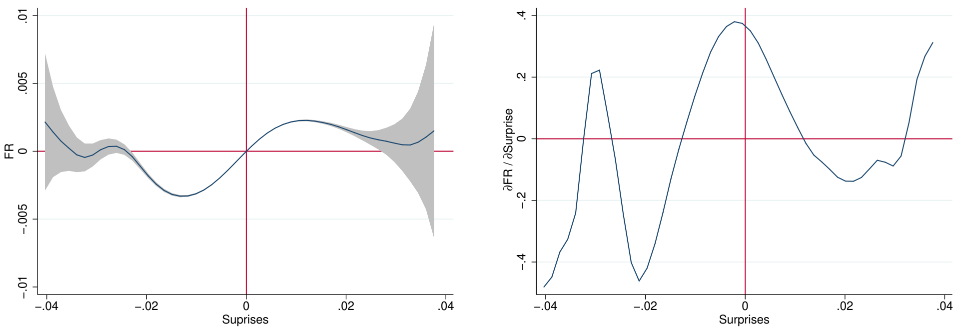
$$\text{FR}_{ijt} = b_0 + b_1\text{Surprise}_{ijt} + b_2\text{Large}_{ijt} + b_3\text{Large}_{ijt} \times \text{Surprise}_{ijt} + \delta_i + \delta_j + \delta_t + \omega_{ijt} \quad (4)$$

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 7, 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 at 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% 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 non-parametric 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 polynomial smooth.



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

Figure 3. *Non-parametric 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 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.*

Because we are interested in the "large" surprises and because we estimate the relationship with local polynomials, the result can be affected and biased by winsorization of the data. To alleviate this concern, we instead trim both forecast revisions and surprises at the 2.5% and 97.5% 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 result is 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 ones 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% of their respective distributions (see Figure 10). For the purpose of comparison, we also present the results by using forecast revisions and surprises which are winsorized at the 2% (1% and 99%) and 5% (2.5% and 97.5%) 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 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 at the tails. After all, the confidence intervals become very wide when the surprises are relatively large in magnitude. However, we find that it is not the case. For the local polynomial estimation using the trimmed data, there are 3060 observations to the left of the trough and 7045 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 to be alleviated.

Another potential issue is that whether offering earnings guidances or not can be strategically chosen by firms, which could affect our estimations. First, it is unlikely that firms make decisions about whether they disclose the earnings guidance every quarter. As discussed in Chen, Matsumoto, and Rajgopal (2011), it is common that firms continue to provide earnings guidances for an extensive period of time and the discontinuation of earnings guidance is typically received unfavorably by the market. Second, we construct a subsample in which we only include earnings forecasts, conditionally on firms (whose earnings are being forecasted) having to release earnings guidances for at least more than 12 consecutive quarters during our sample period.[12] We re-estimate the relationship between forecast revisions and surprises non-parametrically following the procedure described above. The results are presented in Figure 12 in Appendix I.B. They are rather similar to those when using the full sample, and the two key characteristics are even more pronounced. Therefore, the concern of strategic disclosing is inconsequential for our findings.

Our data covers the period of the 2007-2009 financial crisis and it is not unlikely that the participants on financial market behave abnormally during that period, which could potentially 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 one such model where analysts are not certain about the quality of information that they receive.

[12]Based on the initial full sample of management guidance, we select a quarterly management guidance into our sub-sample, 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 assume that the utility function $U(\cdot, \cdot)$ is quadratic in both forecasts and earnings, which is given by:

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

where β is a constant. The utility function satisfies one restriction: analysts' optimal forecast is precisely $F^* = \theta$, conditional on analysts' information being complete (i.e., the earnings θ are known to the analysts). 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 are higher. The utility function can be even more general, but the results are quantitatively and qualitatively similar.[13]

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., $\iota_i \sim N(0, \sigma_z^2)$; and let $\tau_z = 1/\sigma_z^2$. The analyst i makes forecast F_{0i} with heterogeneous information z_{0i}.

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

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

[13] In an earlier version of the paper, we considered the most general quadratic utility functions that satisfy the restriction, which is given by $U(F, \theta) = -(F - \theta)^2 + \alpha\theta^2 + \beta\theta$. We showed that α does not qualitatively matter for the results when its magnitude is small and that in our estimations α is indeed very close to zero. Empirically, we can also indirectly infer that α is zero. All results under this general utility function are available upon request.

where η is normally distributed with mean 0 and variance σ_Y^2, i.e., $\eta \sim N\left(0, \sigma_Y^2\right)$; and 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., $\varepsilon_i \sim N\left(0, \sigma_x^2\right)$; and let $\tau_x = 1/\sigma_x^2$. It can be interpreted as a sufficient statics for all the new information analyst i receives prior to making forecast F_i, except the managerial guidance.[14] 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.

Ambiguity averse preference. 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 the 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 the management has incentives not to release the best possible information at hand, and because even the best possible estimates from the management can be plagued with noise but analysts are not certain about the magnitude of noises.

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.

Further, we assume that analysts dislike uncertainty in the quality of the managerial guidance or they 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

[14] For example, 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}, so that aggregation does not guarantee full revelation. Our qualitative and quantitative results will not be affected.

of the 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 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.[15] The more concave is the function $\phi\left(\cdot\right)$, the stronger is the ambiguity aversion. In other words, it characterizes analysts' taste for ambiguity. In this paper, we consider a function $\phi\left(\cdot\right)$ as in Klibanoff, Marinacci, and Mukerji (2005) 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, it corresponds to the case where analysts are ambiguity neutral or fully Bayesian. When $\lambda\to+\infty$, it corresponds to the case where analysts' aversion to ambiguity is infinite, which is the max-min expected utility model proposed by Gilboa and Schmeidler (1989). Our framework is a generalized version of the standard forecasting problem, in which analysts minimize the mean-squared error of their forecasts of the random variable. In the lens of our framework, the standard case is characterized by the conditions that $\beta=\lambda=0$ and Γ_y is singleton.

3.2. 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 the Bayesian rule:

$$F_{0i}^*=\frac{\tau_z}{\tau_z+\tau_\theta}z_{0i}. \tag{8}$$

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

$$F_i=\int_{\Gamma_y}\left(\frac{\tau_z z_{0i}+\tau_x x_i+\tau_y y}{\tau_\theta+\tau_z+\tau_x+\tau_y}\right)\tilde{p}\left(\tau_y|z_{0i},x_i,y;F_i\right)\mathrm{d}\tau_y, \tag{9}$$

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

$$\tilde{p}\left(\tau_y|z_{0i},x_i,y;F_i\right)\propto\underbrace{\phi'\left(\mathbb{E}_i^{\tau_y}\left[U\left(F_i,\theta\right)\right]\right)}_{\text{Pessimistic Distortion}}\underbrace{p\left(z_{0i},x_i,y|\tau_y\right)p\left(\tau_y\right)}_{\text{Bayesian Kernel}}. \tag{10}$$

[15] 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.

The term with the combined fraction in equation (9) 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 the Bayesian rule.

The distribution of τ_y is updated by following equation (10). When analysts are ambiguity neutral (i.e., $\lambda = 0$), $\phi'(\cdot)$ is constant and the posterior distribution of τ_y simply follows the Bayesian 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 re-weighted 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 the model τ_y as the more likely one, 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 (9) is represented by:

$$F_i = X_i + \kappa\left(X_i, s_i, F_i\right) \cdot s_i, \tag{11}$$

where

$$\kappa\left(X_i, s_i, F_i\right) \equiv \left[\int_{\Gamma_y} \left(\frac{\tau_y}{\tau_\theta + \tau_z + \tau_x + \tau_y}\right) \tilde{p}\left(\tau_y | X_i, s_i; F_i\right) \mathrm{d}\tau_y\right], \tag{12}$$

and the distorted posterior belief is such that

$$\tilde{p}\left(\tau_y | X_i, s_i; F_i\right) \equiv \tilde{p}\left(\tau_y | z_{0i}, \frac{\tau_\theta + \tau_z + \tau_x}{\tau_x}\left(X_i - F_{0i}^*\right) + F_{0i}^*, s_i + X_i; F_i\right). \tag{13}$$

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 the Bayesian 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 1 (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. The 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. The 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 1, the right-hand side of equation (11) always decreases in F_i. The optimal forecast F_i^* is the fixed point of equation (11). The following proposition summarizes the equilibrium existence and the 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 (11) and a unique optimal response $\kappa^*\left(s_i\right) \equiv \kappa\left(X_i, s_i, F_i^*\right)$ associated with it.*

Two special cases are discussed as follows. First, if the quality of managerial guidance is certain and known to analysts, the belief $\tilde{p}$ is degenerate (rational expectation) and so is the response to surprises κ^{RE}. The Bayesian rule dictates:

$$\kappa^{\mathrm{RE}} = \frac{\tau_Y}{\left(\tau_\theta + \tau_z + \tau_x + \tau_Y\right)}. \tag{14}$$

Second, if analysts are ambiguity neutral, there is no dependence of analyst i's posterior belief $\tilde{p}$ on F_i. Therefore, the Bayesian 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 the 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 guidances. 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 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 says 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 improve analyst i's expected utility, she would behave with more caution (due to the ambiguity averse preference) 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 ones of the same magnitude, that is, $\kappa^* \left(s_i^-\right) > \kappa^* \left(s_i^+\right)$.

4.2. Non-monotonicity

Then, we show that the model also features a non-monotonic relationship between forecast revisions and surprises. Two key take-away messages are as follows. First, the non-monotonicity does not rely on ambiguity aversion, but instead on ambiguity (uncertainty) in quality. Second, in fact, the non-monotonicity disappears when the degree of aversion towards ambiguity becomes extreme. Proposition 3 formalizes the former and proposition 4 characterizes the latter. To simultaneously capture both non-monotonicity and asymmetry, neither the ambiguity neutral preference 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.*

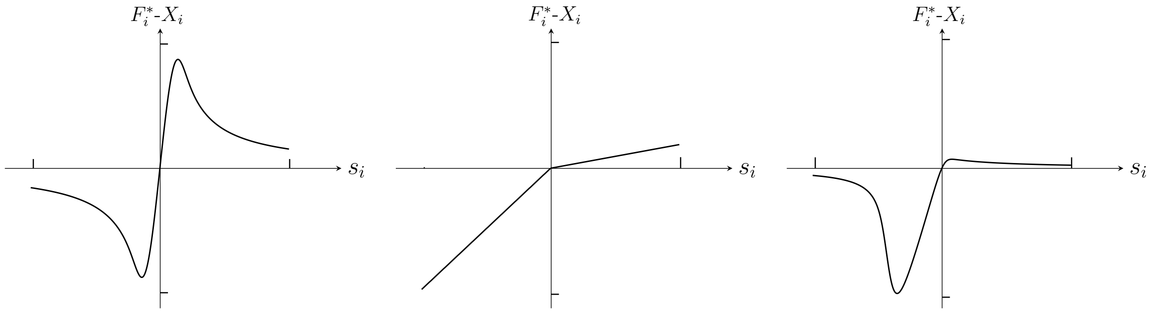
Given that the quality of guidance is uncertain, analyst i updates her belief through two mechanisms. First, for any given quality τ_y, analyst i updates her belief about the earnings upon receiving the guidance. This mechanism dictates that positive (negative) surprises raise (suppress) forecasts. Second, 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. Or, they tend to believe that large surprises are of low quality. Crudely, that 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 case: the extreme ambiguity aversion ($\lambda \to +\infty$) or the max-min preference.

Proposition 4. *If analysts have the max-min preference ($\lambda \to +\infty$), the optimal forecast revision $F_i^* - X_i$ is increasing in surprise s_i.*

For ease of explanation, consider $X_i = 0$. First, as predicted by proposition 3, the Bayesian mechanism dictates that analyst i believes that larger (smaller) surprises are of lower (higher) quality. Second, the ambiguity aversion mechanism dictates that analyst i believes that negative surprises are of higher (lower) quality than positive ones of the same magnitude, if $\beta > 0$ ($\beta < 0$). In particular, given that the ambiguity aversion is extreme, analysts believe that the quality of the negative news is of 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 max-min preferences ($\lambda \to +\infty$). 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 non-monotonicity are present.*

highest possible value and that of the positive news is of the lowest possible value, if $\beta > 0$; and vice versa. In other words, the impact of extreme ambiguity aversion dominates the Bayesian mechanism. Therefore, forecast revisions always increase in surprises, despite the sign of β. 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 ones 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.

In sum, the contrast of the two polar cases reveals (i) that ambiguity in guidance quality gives rise to non-monotonicity in surprises and (ii) that aversion towards 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

Does the preference of ambiguity aversion contribute to analysts' overreaction to information in our model and its asymmetrical and non-monotonic 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 to the negative.

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 20) and the analyst's response to surprises κ^{RE} is characterized by Bayes' rule (see equation 14). 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 $p(s_i)$. That is,*

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

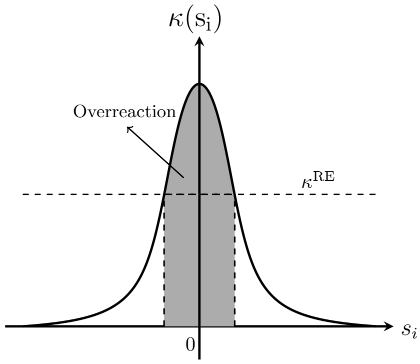
where $\hat{\mathbb{E}}$ is an expectation operator under the adjusted belief $\hat{p}(s_i)$,

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

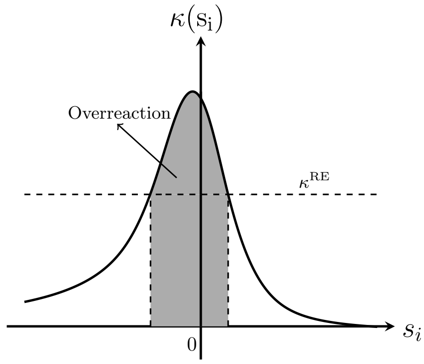
The term $\Omega(s_i)$ is an adjusted term for the transformed belief. If the average response of analysts (i.e., $\hat{E}\left[\kappa\left(s_i\right)\right]$) is higher than the rational expectation benchmark κ^{RE}, analysts appear to be overreacting to information (i.e., $\hat{b}_1 < 0$); if it is 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 $\hat{E}\left[\kappa\left(s_i\right)\right] = \kappa$, 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 (Kohlhas and Broer 2019), 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 $\hat{E}[\kappa]$ degenerates to precisely κ^{RE}. It is consistent with the prediction that there is no over- or underreaction in a rational expectations model, i.e., $\hat{b}_1 = 0$.

Observe that $\hat{E}$ 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



(a) $\lambda = 0$



(b) $\lambda > 0$

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

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 $\kappa^{\mathrm{RE}} < \hat{E}\left[\kappa\left(s_i\right)\right]$ 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 the 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 rest is 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 underreact to information, without noticing that some analysts overreact and others underreact to information.

Figure 5(b) illustrates the case when the ambiguity averse preference is 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 towards the negative and analysts react even more strongly to new information when it contains unfavorable news than when it contains favorable news. That is consistent with our finding in section 2.3 that analysts' overreaction to new information is stronger when the managerial guidance is negative.

Taken together, this pattern emerging from our model says 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, non-monotonicity 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. Our strategy is that 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 to information in sections 2.2 and 2.3.

5.1. Connect 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., $S_i \equiv y - F_{0i}$. 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 observable surprises to the econometrician 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 a 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 which is consistent with our empirical findings. We relegate the relevant characterization and discussion to Appendix III.

Further, our model features a finite degree of ambiguity aversion, i.e., λ is between $(0, +\infty)$. 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 non-parametrically from the data (section 2.4). Various quantitative exercises, using this estimated model, will reveal the roles of key mechanisms, such as prior beliefs and ambiguity averse preferences.

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 preference of analysts, i.e., ambiguity aversion λ, and analysts' attitude towards 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 the 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 the 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 the 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 (non-parametrically 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 non-parametric regression as the empirical data (see section 2.4). We then obtain the vector m from the estimated relationship

Table 3. Estimated Model Parameters

	Mean	90% HPDI	95% HPDI
λ	295.0	$(145.2, 394.1)$	$(123.9, 394.3)$
β	1.015	$(0.575, 1.517)$	$(0.608, 1.873)$
U	0.684	$(0.649, 0.722)$	$(0.641, 0.728)$
L	0.055	$(0.025, 0.087)$	$(0.017, 0.089)$
$100\sigma_x$	2.106	$(0.963, 3.204)$	$(0.915, 3.411)$
$100\sigma_z$	0.262	$(0.215, 0.300)$	$(0.214, 0.305)$
$100\sigma_Y$	0.400	$(0.376, 0.424)$	$(0.373, 0.424)$

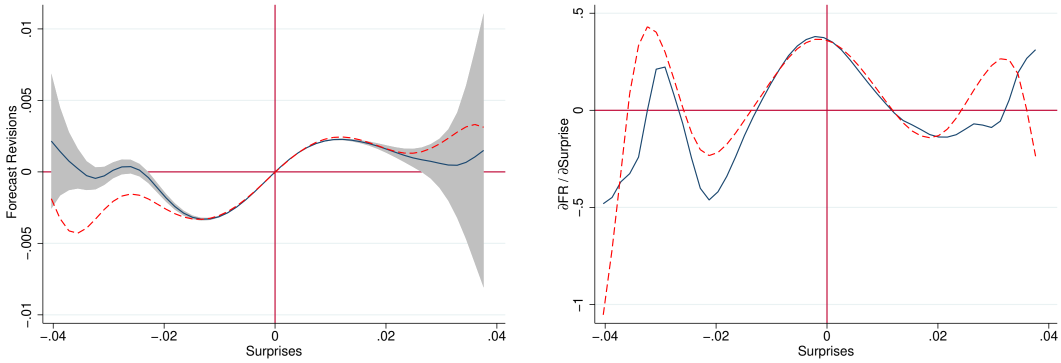
between forecast revisions and surprises observable to the econometrician. The distance that we construct is:

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

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 for 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 to 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 estimate $1/\sigma_\theta^2 + 1/\sigma_z^2$ directly in the estimation and back out σ_Y by requiring that the unconditional volatility of surprises matches its empirical counterpart, an internal consistent condition for our estimation strategy.

The estimated parameters are reported in Table 3 together with the 90% and 95% high posterior density interval (HPDI), respectively. 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. It 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 date of managerial guidance release). Based on this set of parameters, the



(a) $\lambda = 295$, Non-parametric estimation

(b) $\lambda = 295$, Derivative

Figure 6. *The revision-surprise relationship non-parametrically estimated with simulation data. We simulate the model with the set of parameters reported in Table 3 and, in particular, $\lambda = 295$. 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.*

response of forecast revisions to surprises under noisy rational expectation (i.e., κ^{RE}) is 0.32. The upper bound for the subjective belief of managerial guidance precision is 0.684 and the lower bound is 0.055. The support is large enough to allow sufficient ambiguity and encompass κ^{RE}.

The parameter β is positive, 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.[16] The degree of ambiguity aversion is the key and its value is estimated to be $\lambda = 295$.

Using this set of estimated parameters, we simulate the model and estimate the revision-surprise relationship with the simulation data non-parametrically. In Figure 6(a), we display the relationship, together with its empirical counterparts (previously shown in Figure 3(a)). In Figure 6(b), we illustrate its implied derivative with respect

[16]There are multiple channels through which financial analysts would benefit from a better earnings performance of the firms that they cover and therefore view positive surprises in managerial guidance as favorable ones. First, stronger earnings performances can be rewarding to financial analysts who make earnings forecasts and recommendations for the underlying stocks. 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 which, 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. Second, weaker earnings performances can be costly to financial analysts. The firms' management may be hostile towards negative sell-type recommendations, despite the earnings performance. Implicitly, when the earnings of firms underperform, analysts' negative recommendations may damage their relationship with firms' management, through which they gain access to the firm-related information (e.g.,Lim, Richardson, and Roberts 2004).

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

		(1)	(2)	(3)	(4)
	Coff.	Data	$\lambda = 295.0$	$\lambda = 0$	$\lambda \to \infty$
(A) $\text{FE}_i = b_0 + b_1 \text{FR}_i + \omega_i$	b_1	-0.0950^{***}	-0.0906^{***}	-0.0729^{***}	-0.3853^{***}
			$(-0.1034, -0.0794)$	$(-0.0848, -0.0618)$	$(-0.3943, -0.3771)$
(B) $\text{FR}_i = b_0 + b_1 \text{Surp}_i + \omega_i$	b_1	0.2441^{***}	0.3029^{***}	0.2992^{***}	0.3720^{***}
			$(0.3017, 0.3041)$	$(0.2981, 0.3004)$	$(0.3695, 0.3784)$
(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^{***}	0.2164^{***}	0.2585^{***}	0.0598^{***}
			$(0.2137, 0.2192)$	$(0.2560, 0.2613)$	$(0.0587, 0.0608)$
	b_3	0.0846^{***}	0.0922^{***}	0.000	0.6243^{***}
			$(0.0887, 0.0962)$	$(-0.0038, 0.0038)$	$(0.6233, 0.6254)$
(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^{***}	0.3605^{***}	0.3573^{***}	0.3720^{***}
			$(0.3597, 0.3613)$	$(0.3565, 0.3580)$	$(0.3699, 0.3741)$
	b_3	-0.2674^{***}	-0.0719^{***}	-0.0725^{***}	0.0000
			$(-0.0733, -0.0702)$	$(-0.0740, -0.0710)$	$(-0.0024, 0.0027)$

*** 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.

to surprises. Our model can successfully capture both features of non-monotonicity and asymmetry.

The estimated value of λ is not too high or too low, indicating that neither max-min preferences ($\lambda \to +\infty$) nor ambiguity neutral preferences ($\lambda = 0$) would be realistic for analysts in this setting. To illustrate this, we simulate the model for $N = 1000$ times and for each simulation there are 101,086 observations of analysts (i.e., the size is consistent with our empirical data). For each simulation, we estimate equations (B), (C) and (D) in Table 4, which are counterparts of equations (2), (3) and (4) in section 2.4.[17] For comparison, we repeat the aforementioned exercises twice with the same set of parameters, except allowing λ to be 0 and $+\infty$, respectively. Then, we estimate equations (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 $\lambda = 295$, $\lambda = 0$ 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 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 by using linear regressions.

Column (3) presents the results when $\lambda = 0$. It is evident that the revision-surprise

[17]For this purpose, we do not need to use the surprises uncovered in the data to simulate the model. Instead, we re-sample them for each simulation.

relationship is symmetric, i.e., the interaction term in equation (C) is zero and insignificant. But 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. But 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.

Our results from the simulated data are consistent with our model predictions in section 4. In sum, to capture both qualitative features of non-monotonicity and asymmetry simultaneously, 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). Our first step is to examine the average extent of overreaction implied by the estimated model. To investigate this, we first estimate the regression specified in equation (A) (in Table 4) with our simulation data ($\lambda = 295$). 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. In regression (A), we observe that the estimated coefficient for b_1 is negative and significantly different from zero, indicating that analysts do indeed overreact to information in our model. The magnitude is larger than that in the data, which is expected given that it is not targeted in the estimation.

Can our estimated model also predict the pattern of cross-sectional varying overreaction? To investigate this, we utilize the simulated data and construct the observable surprises 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 break 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) the running decile window 1 covers deciles 1 and 2; and (3) the running decile window 10 covers deciles 9 and 10. For each subsample, we re-estimate the 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 which is also fairly close to the pattern in the empirical data (Figure 2).

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 (10). It

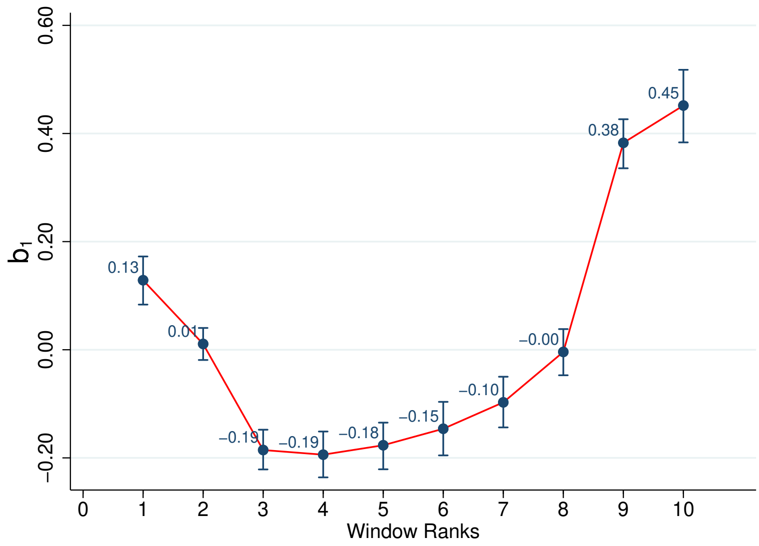


Figure 7. *Overreaction by surprise deciles with simulated data. Using simulated data, we report the estimated coefficients of the FE-on-FR regressions* b_1 *for each running decile window and we plot them against the window rank. A running decile window* j *covers decile* $j-1$, j, *and* $j+1$ *if*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

has two key components: Bayesian kernel and pessimistic distortion, both of which contribute to the observed overreaction.

To isolate the impact of ambiguity averse preferences in this setting, we re-estimate equation (A) with the simulation data by setting $\lambda = 0$. The regression results are reported in column (3) of Table 4. We observe that the magnitude of overreaction, the estimated coefficient b_1 in regression (A), is smaller.

Intuitively, when the ambiguity aversion is stronger, we should predict that the magnitude of overreaction is much larger. To see that this is indeed the case, we re-estimate equation (A) with the simulation data by setting λ to infinity. The regression results are reported in column (4) of Table 4. The corresponding estimated coefficients in question (i.e., b_1 in equation (A)) are indeed larger in magnitude than those in our estimated model with a finite degree of ambiguity aversion.

The comparison of results with and without ambiguity aversion reveals mechanisms that drive the overreaction to information in this model. First, the prior belief $p(\tau_y)$ also plays a role in the observed overreaction: the estimated prior allocates a sufficiently large density to the right of objective quality τ_Y. Second, the pessimistic distortion (or ambiguity aversion) also contributes to the observed overreaction: analysts overreact to information because they treat unfavorable surprises as higher quality news than they actually are.

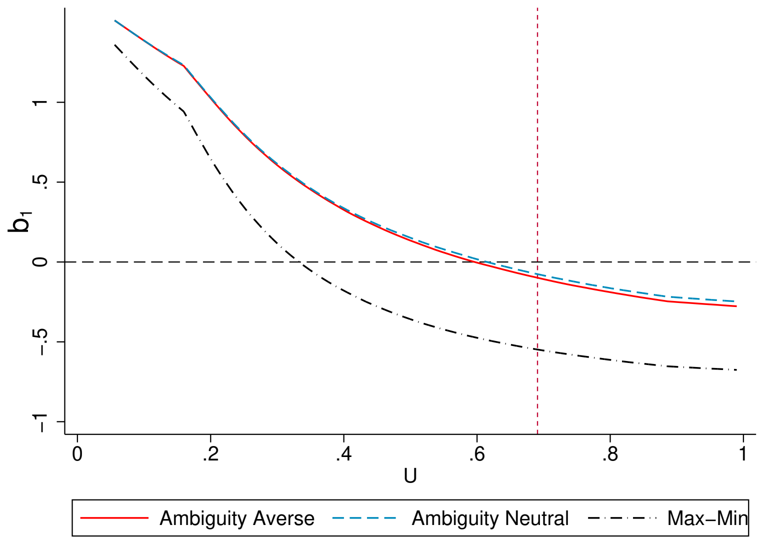


Figure 8. *Ambiguity aversion and prior beliefs. The solid line illustrates the relationship between estimated coefficients b_1 and the upper bound of prior belief, i.e., U. When U is large enough or the prior belief is that the managerial guidance tends to be very lousy, analysts underreact on average. When U is small enough, analysts overreact. The dashed line corresponds to the case where $\lambda = 0$ and the broken line to the case where $\lambda = +\infty$. 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.*

To further illustrate this point, we simulate the model by varying the prior distribution of τ_y, or, equivalently, the upper bound of the prior belief U (defined in section 5.2) from a value slightly larger than the lower bound L to 1, while we keep the 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 managerial guidance.

We re-estimate the 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 for b_1 are positive; and when U is large, they are negative. This is intuitive: when analysts' prior belief is such that the guidance is, on average, more precise than it actually is, they overreact; when the prior belief is such that the guidance is on average very lousy in quality, they underreact.

We repeat the same exercise by using simulation data which sets $\lambda = 0$ and $\lambda = \infty$ and plot the counterparts in Figure 8 with the dashed and broken lines, respectively. The contrast between the cases of $\lambda = 0$, $\lambda = 295$, and $\lambda = \infty$ shows that the ambiguity averse preference indeed contributes to the overreaction. For the same prior belief (represented by U here), analysts tend to overreact more when they are more ambiguity averse. This result from the simulation is consistent with our model predictions. This mechanism is relatively new to the literature concerning whether agents over- or underreact to information.

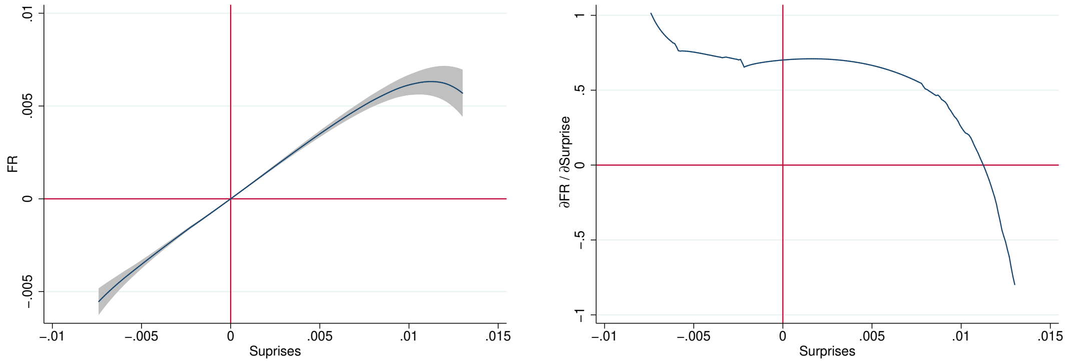
5.4. Auxiliary Predictions

In this paper, we provide a theory about how the expectation is formed when forecasters are not completely certain about 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 doubts 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.

To test this auxiliary prediction with our data, there is a conceptual hurdle: the perceived uncertainty in guidance quality is not observable and therefore not measurable. To work around it, we proxy 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 which only includes firms that deliver a very precise earnings guidance whose uncertainty in the 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-quarters 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 with a presence of less 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 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.



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

Figure 9. *Non-parametric estimation, using a subsample with the top 5% firms in terms of guidance precision. Panel (a) illustrates the relationship between forecast revisions and surprises in managerial guidances that is non-parametrically estimated using the Epanechnikov kernel and the third degree of the polynomial smooth. The shaded areas stand for the 95% confidence intervals for the respective estimations. Panel (b) illustrates the derivative of forecast revisions with respect to surprises. It is almost flat around zero, indicating a linear relationship between forecast revisions and surprises.*

Using this subsample, we re-estimate the relationship between forecast revisions and guidance surprises non-parametrically 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 the derivative estimated using the full sample (shown in Figure 3(b)). On the one hand, this exercise strengthens our confidence in our model mechanisms. On the other hand, it also helps us stress that uncertainty in information quality does have an impact on how analysts update their beliefs.

5.5. Alternative Hypotheses: Agency Issues and Loss Aversion

In this paper, we provide a simple unified framework to account for new facts regarding how analysts update their forecasts or expectations. It is important that our estimated model can generate the skewed U-shaped pattern of overreaction that is consistent with the data, which helps differentiate our theory from other potential explanations. We discuss two alternative hypotheses in this section.

While being informative about expectation formation and being new to the literature, this empirical setting can potentially be contaminated by agency issues between analysts and the managerial teams. Concerns are likely to arise from both sides. First, managers may be reluctant to issue bad news in general, and to issue negative management guidance in particular. For example, Kothari, Shu, and Wysocki (2009) and

Ge and Lennox (2011) show that management tends to withhold and delay the disclosure of bad news to investors. Therefore, it is likely that a negative guidance can be seen as a stronger and more creditable piece of information about firms' earnings, which may trigger greater reactions from analysts. Second, analysts could also have incentives to adjust their forecasts such that their forecasts are beatable by the reported earnings, i.e., "a pattern of the walk-down to beatable" documented by Richardson, Teoh, and Wysocki (2004) and Cotter, Tuna, and Wysocki (2006). 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. While it is acknowledged that both mechanisms may contribute to the asymmetric pattern found in this paper, they do not offer any predictions consistent with the findings that larger managerial surprises can lead to smaller forecast revisions of analysts. If those mechanisms were the only underlying forces, then we would not expect to observe the non-monotonicity in the relationship between forecast revisions and surprises.

Another plausible conjecture is that models with loss averse analysts may also likely generate the empirical pattern documented in section 2.4, given that those analysts also behave in a pessimistic way. To investigate this possibility, we follow Capistrán and Timmermann (2009) and specify the loss function of analysts to be:

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

where F_i stands for the forecast of analyst i, the fundamental $\theta \sim N\left(0, 1/\tau_\theta\right)$, and the parameter ϕ is a constant which captures asymmetries in the loss function. If $\phi > 0$, analysts dislike the negative forecast error $\theta - F_i < 0$ more than the positive forecast error $\theta - F_i > 0$. If ϕ goes to zero, the loss function is reduced to the standard mean-square error function.

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

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

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

Can such a systematic bias in the analysts' behavior lead to asymmetry in expectation formation? Assume that analyst i makes an initial forecast F_{0i}^* about the fundamental θ after receiving private information z_i, where $z_i \sim N\left(\theta, 1/\tau_z\right)$. Her initial optimal forecast is:

$$F_{0i}^{*}=\frac{\tau_{z}}{\tau_{\theta}+\tau_{z}}z_{i}-\frac{1}{2}\phi\frac{1}{\tau_{\theta}+\tau_{z}}.$$

After making forecast F_{0i}^{*}, each analyst receives a private signal x_i, where $x_i \sim N(\theta, 1/\tau_x)$ and managerial guidance y, where $y \sim N(\theta, 1/\tau_Y)$. The updated optimal forecast is:

$$F_{i1}^{*}=\mathbb{E}_{i}[\theta]-\frac{1}{2}\phi\mathrm{Var}_{i}[\theta]=\frac{\tau_{z}z_{i}+\tau_{x}x_{i}+\tau_{Y}y}{\tau_{\theta}+\tau_{z}+\tau_{x}+\tau_{Y}}-\frac{1}{2}\phi\frac{1}{\tau_{\theta}+\tau_{z}+\tau_{x}+\tau_{Y}}.$$

Therefore, the forecast revision is such that

$$\mathrm{FR}_{i}\equiv F_{i1}^{*}-F_{i0}^{*}=\frac{\tau_{x}}{\tau_{\theta}+\tau_{z}+\tau_{x}+\tau_{Y}}\left(x_{i}-F_{i0}^{*}\right)+\frac{\tau_{y}}{\tau_{\theta}+\tau_{z}+\tau_{x}+\tau_{Y}}S_{i}. \tag{15}$$

where $S_i \equiv y - F_{i0}^{*}$ (i.e., the definition of surprises to the econometrician used in section 2 and 5). Intuitively, $x_i - F_{i0}^{*}$ and $y - F_{i0}^{*}$ are independent. Despite a systematic bias existing in forecasts, forecast revisions are still linear in surprises.

6. Conclusion

This paper documents a set of 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 ones; and forecast revisions are asymmetric in surprises and non-monotonic. 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 values 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. It is also important to study our empirical setting itself. First, sell-side financial analysts constitute a significant part of capital market participants and play an important role by collecting, processing, and conveying relevant information to the 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.

Our results can be also interesting to the literature which studies information production in financial markets. Given the qualitative features of analyst forecasts documented in this paper, it will be interesting to explore how participants in the financial markets react to and make use of the information that analysts provide. We leave those to future work.

References

Afrouzi, H., S. Y. Kwon, A. Landier, Y. Ma, and D. Thesmar (2021). Overreaction in expectations: Evidence and theory. *Available at SSRN*.

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. and Z. Huo (2021). Myopia and anchoring. *American Economic Review 111*(4), 1166–1200.

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. (2021). The micro and macro of managerial beliefs. *Journal of Financial Economics*.

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.

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.

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.

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

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.

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

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.

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.

Cotter, J., I. Tuna, and P. D. Wysocki (2006). Expectations management and beatable targets: How do analysts react to explicit earnings guidance? *Contemporary Accounting Research 23*(3), 593–624.

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.

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.

Ge, R. and C. Lennox (2011). Do acquirers disclose good news or withhold bad news when they finance their acquisitions using equity? *Review of Accounting Studies 16*(1), 183–217.

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. *The 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.

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

Kohlhas, A. and T. Broer (2019). Forecaster (mis-) behavior. *Working paper*.

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

Kothari, S. P., S. Shu, and P. D. Wysocki (2009). Do managers withhold bad news? *Journal of Accounting research 47*(1), 241–276.

Lian, C. (2018). A theory of narrow thinking. *The Review of Economic Studies*.

Lim, J., V. J. Richardson, and T. L. Roberts (2004). Information technology investment and firm performance: a meta-analysis. In *37th Annual Hawaii International Conference on System Sciences, 2004.*, pp. 10.

Ma, Y., T. Ropele, D. Sraer, and D. Thesmar (2020). A quantitative analysis of distortions in managerial forecasts. National Bureau of Economic Research.

Mackowiak, B. and M. Wiederholt (2009). 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*.

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.

Richardson, S., S. H. Teoh, and P. D. Wysocki (2004). The walk-down to beatable analyst forecasts: The role of equity issuance and insider trading incentives. *Contemporary accounting research 21*(4), 885–924.

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.

Van Nieuwerburgh, S. and L. Veldkamp (2006). Learning asymmetries in real business cycles. *Journal of Monetary Economics 53*(4), 753–772.

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 5. 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 6. 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

B. Robustness

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 6. All results are robust, thus suggesting that our results are not sensitive to the way in which we handle outliers.

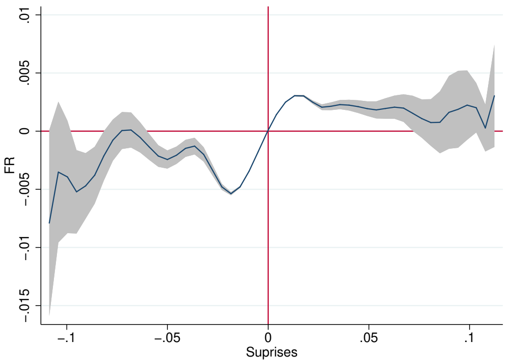
Table 7. Robustness: Definition of Large Surprises

	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

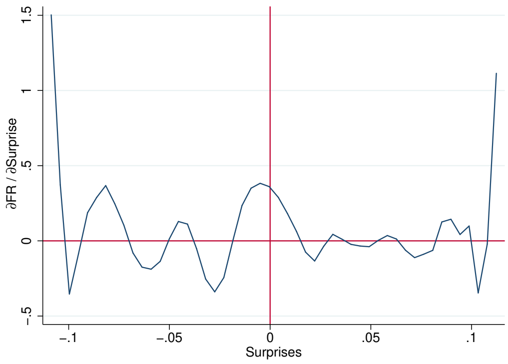
The standard errors are clustered on firm and calendar year-quarter following (Petersen 2009).
*** p<0.01, ** p<0.05, * p<0.1

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 7. 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 7, 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

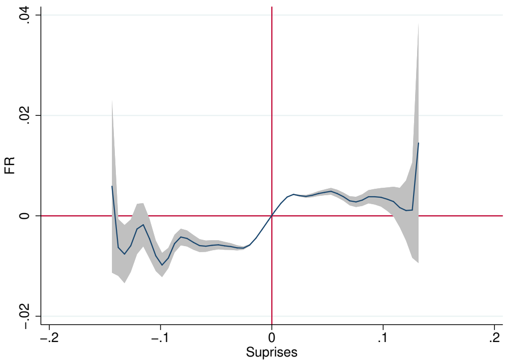


(b) Trimming, Derivative: marginal effect

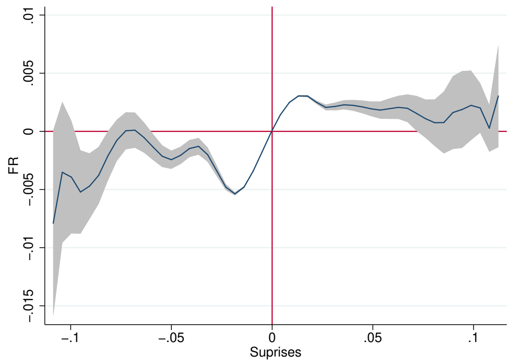
Figure 10. *Non-parametric estimation, 2% trimming (1%, 99%). Panel (a) illustrates the relationship between forecast revisions and surprises in managerial guidances (both trimmed at 2%) that is non-parametrically estimated using the Epanechnikov kernel and the third degree of the polynomial smooth. 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.



(a) Winsorization, Non-parametric estimation

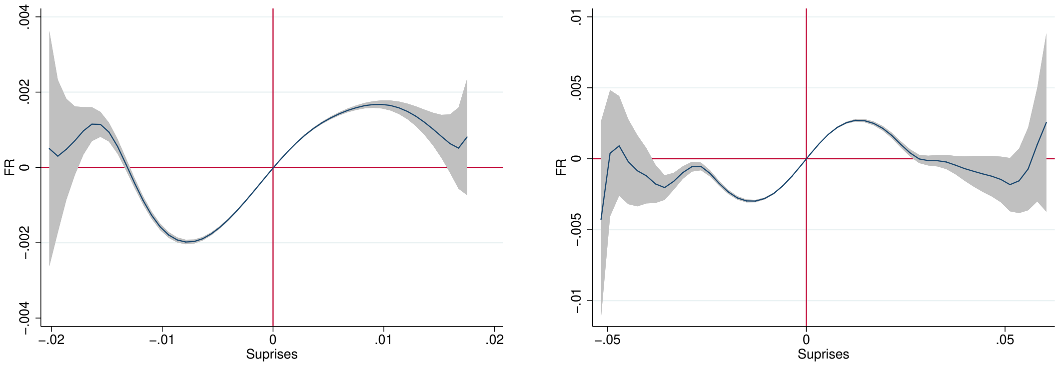


(b) Trimming, Non-parametric estimation

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 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. The pattern is rather similar.*

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 in both cases 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.

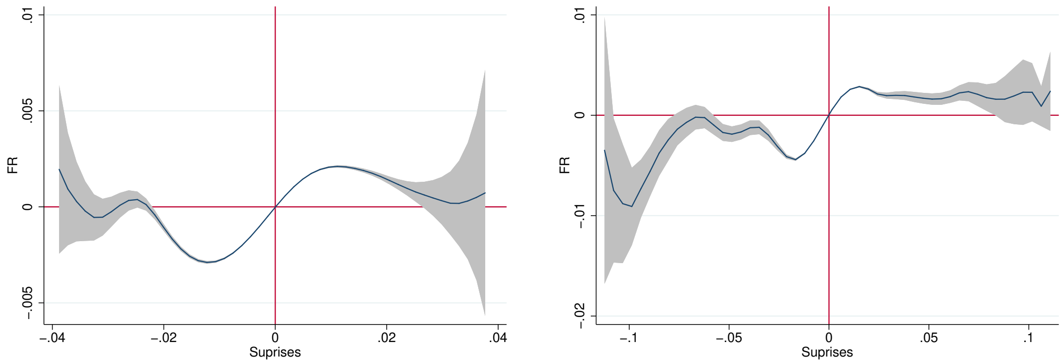
Subsample of Firms that Issued at Least 12 Consecutive Guidances.



(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.

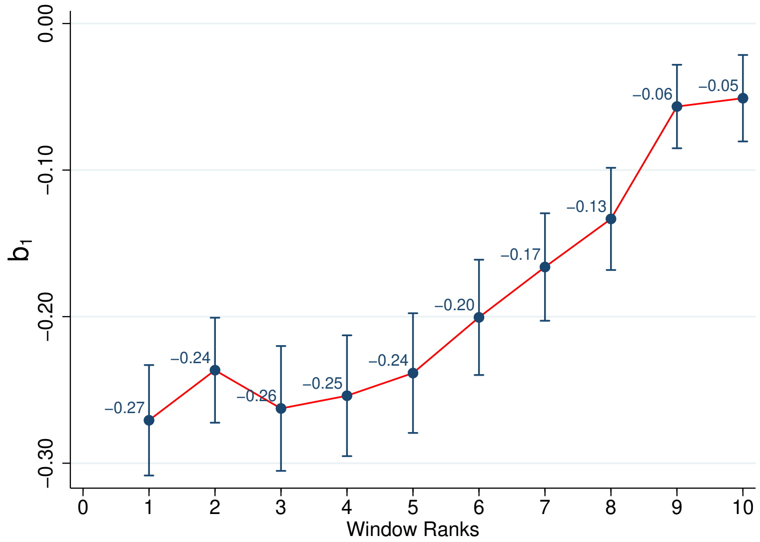


Figure 14. *Heterogeneous Overreaction, Trimming 2% Outliers. The estimated coefficients of the FE-on-FR regressions* b_1 *and 95% confidence interval for each running decile window is plotted against the window rank. A running decile window j covers decile* $j-1$*, j, and* $j+1$ *if* $j \neq 1$ *or*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

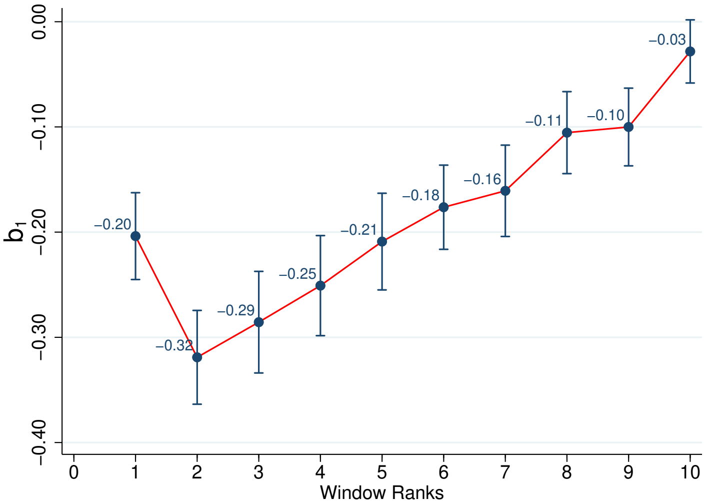
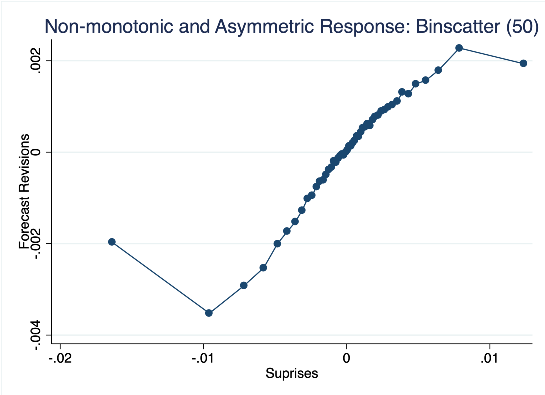
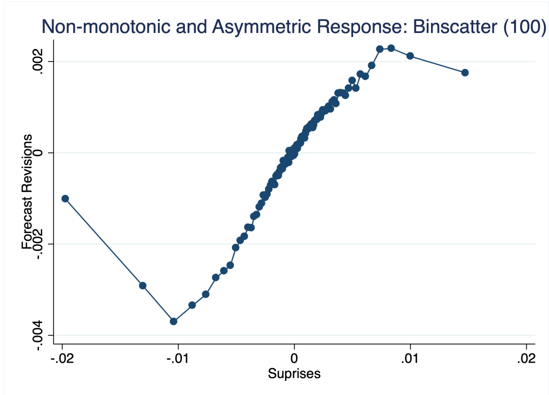


Figure 15. *Heterogeneous Overreaction. The estimated coefficients of the FE-on-FR regressions b_1 and 95% confidence interval for each decile of surprises.*

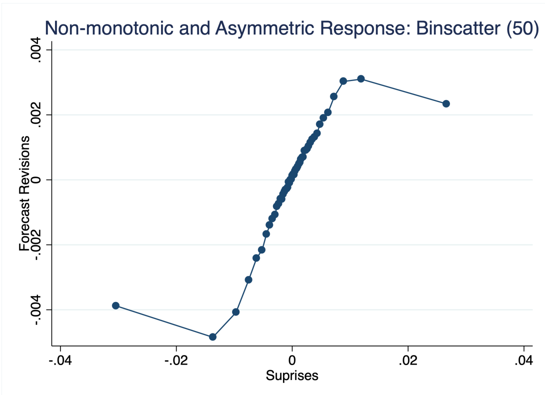
Binscatter Plot.



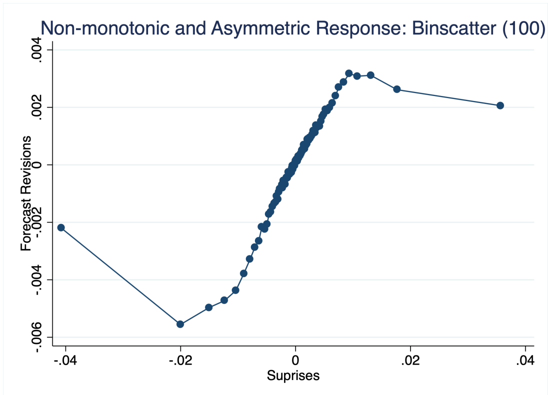
(a) Binscatter (50 bins), trimmed at 5%



(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

Derivation of Equation (11)-(13)**.** Denote $\delta \equiv \frac{\tau_y}{\tau_\theta+\tau_z+\tau_x+\tau_y}$. Then, it can be shown that

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

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 (9) 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 1. 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 (9) is equivalent to (11):

$$F_i = X_i + \underbrace{\left[\int_{\Gamma_y} \left(\frac{\tau_y}{\tau_\theta + \tau_z + \tau_x + \tau_y}\right) \tilde{p}\left(\tau_y | z_{0i}, x_i, y; F_i\right) d\tau_y\right]}_{\kappa} \cdot s_i \tag{16}$$

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 (16) 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 (11) 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 (11) 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:

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

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:

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

By the definition of κ, this implies:

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

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 $\kappa^*(s_i^-) = \kappa^*(s_i^+) = \kappa'$. For any pair of (s_i^-, s_i^+), we have:

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

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 $\to \infty$, 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 $\beta \to \infty$. 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 preference 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 preference. 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.

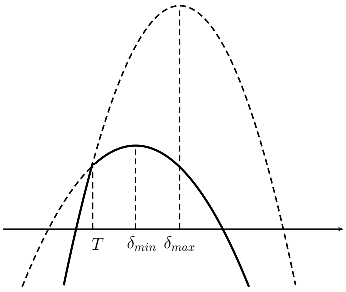
  Figure 17(b) graphically illustrates the value function under the worst case scenario when $\delta_{\max} < T(s_i)$. Therefore, it must be the case that the optimal $\kappa^* = \delta_{\max}$ when $\delta_{\max} < T(s_i)$.

- 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.

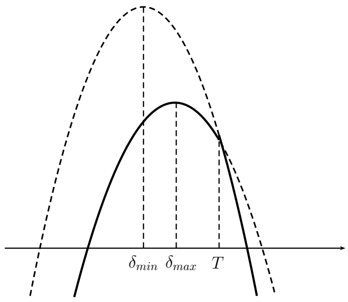
  Figure 17(c) graphically illustrates the value function under the worst case scenario when $\delta_{\min} < T(s_i) < \delta_{\max}$. Therefore, it must be the case that the optimal $\kappa^* = T(s_i)$ when $\delta_{\min} < T(s_i) < \delta_{\max}$.

To summarize, we have the following optimal forecasting rule under the maxmin preferences when $\alpha < 1$:

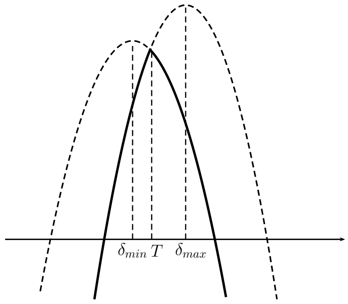
$$\kappa^* = \begin{cases} \delta_{\min} & \text{if } \delta_{\min} > T(s_i) \\ \delta_{\max} & \text{if } \delta_{\max} < T(s_i) \\ T(s_i) & \text{otherwise} \end{cases} \tag{17}$$



(a) $\delta_{\min} > T(s_i)$



(b) $\delta_{\max} < T(s_i)$



(c) $\delta_{\min} < T(s_i) < \delta_{\max}$

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

Or equivalently,

$$F^* - X_i = \begin{cases} \delta_{\min} s_i & \text{if } \delta_{\min} > T(s_i) \\ \delta_{\max} s_i & \text{if } \delta_{\max} < T(s_i) \\ T(s_i)\, s_i & \text{otherwise} \end{cases} \tag{18}$$

Note that $T(s_i)\, s_i$ is always increasing in s_i. Therefore, following the continuity of $F_i^* - X_i$ w.r.t 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

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

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

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

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 $\kappa(s_i)$ 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

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

■

Table 8. Forecast Revision and Surprises in Managerial Guidance: Cubic Regressions

	Outcome Variable: Forecast Revision FR_i							
	Winsorization at 1% and 99%				Winsorization at 2.5% and 97.5%			
	Sample:	Sample: Exclude			Sample:	Sample: Exclude		
	Full	Pre-anc	Multiple	Both	Full	Pre-anc	Multiple	Both
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Surprise_i	0.3443***	0.3670***	0.3994***	0.4431***	0.4232***	0.4553***	0.4656***	0.5273***
	(0.0196)	(0.0324)	(0.0191)	(0.0348)	(0.0169)	(0.0249)	(0.0176)	(0.0275)
Surprise_i^2	-0.6659***	-0.8034***	-0.7156***	-0.8246***	-0.9007***	-1.2585***	-0.7315***	-0.9236**
	(0.0764)	(0.1766)	(0.0773)	(0.1727)	(0.1593)	(0.2615)	(0.1923)	(0.4005)
Surprise_i^3	-35.3214***	-40.4462***	-40.0359***	-48.7169***	-212.4218***	-239.1811***	-216.9838***	-253.5881***
	(2.2350)	(4.4495)	(2.3357)	(4.2898)	(14.5469)	(23.3964)	(17.3704)	(32.1236)
Constant	-0.0004***	-0.0005***	-0.0005***	-0.0006***	-0.0002***	-0.0002***	-0.0003***	-0.0002**
	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
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.	110,895	50,558	46,493	17,606	110,895	50,558	46,493	17,606
Adj. R^2	0.4789	0.5427	0.5197	0.6044	0.4984	0.5602	0.5244	0.5987

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

Appendix III: Cubic Approximations and Evidence

Evidence In this appendix, we consider the cubic regression, a non-linear diagnostic test, which is particularly useful for us in order to discern whether forecast revisions are less responsive to larger surprises and whether forecast revisions respond more strongly to negative surprises than to positive ones. To illustrate the diagnostic power of cubic specification, we derive the following:

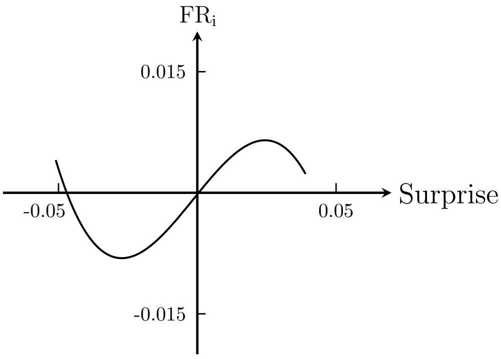
$$\frac{\partial \mathbb{E}\left[\text{Forecast Revision}|\text{Surprise}\right]}{\partial \text{Surprise}} = b_1 + 2b_2 \times \text{Surprise} + 3b_3 \times \text{Surprise}^2. \tag{19}$$

It is evident that a non-zero b_2, the coefficient on the quadratic term, says that the sign of surprise or favorability of managerial guidance matters for how forecast revisions respond to surprises in guidance. A negative b_2 says that forecast revisions should be stronger when the surprises are negative. If the coefficient on the linear term b_1 is positive, it implies that forecast revisions should be positively correlated with surprises, when the magnitude of surprises is small enough. If the coefficient on the cubic term b_3 is negative and large, it suggests that the relationship can be decreasing when the magnitude of surprises is large enough.

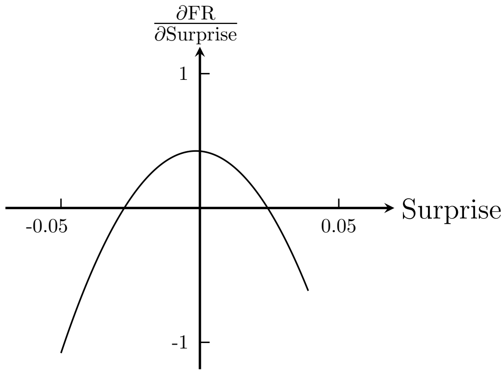
Therefore, we estimate the following cubic regression,

$$\text{FR}_{ijt} = b_0 + b_1 \text{Surprise}_{ijt} + b_2 \text{Surprise}_{ijt}^2 + b_3 \text{Surprise}_{ijt}^3 + \delta_i + \delta_j + \delta_t + \omega_{ijt} \tag{20}$$

Column (1) of Table 8 shows the regression results. The estimated coefficients b_0



(a) Cubic Regression



(b) Derivative

Figure 18. *Cubic regression, 5% winsorization (2.5%, 97.5%). Panel (a) illustrates the relationship between forecast revisions and surprises in managerial guidances, that is estimated with cubic regression. It is non-monotonic and asymmetric. Panel (b) illustrates its derivative with respect to surprises.*

and b_2 are negative and significantly different from zero. The estimated coefficient b_1 is significantly positive but b_2 is significantly negative and large in magnitude. Similarly to the previous section, we winsorize both forecast revisions and surprises at the 2.5% and 97.5% of their respective distributions and re-estimate equation (20) with the full sample. The result is shown in column (5) of Table 8. The magnitude of the key coefficients is different but their signs are all consistent with those reported in column (1). Figure 18(a) illustrates the cubic relationship with the estimated coefficients reported in column (5) of Table 8 and Figure 18(b) illustrates the corresponding derivatives associated with the relationship.

Further, we show that the results from the cubic regression are robust to various subsamples. The results in columns (2), (3) and (4) of Table 8 are based on a sample excluding all firm-quarters with pre-announcement guidance, a sample excluding all firm-quarters with multiple guidances and a sample excluding all firm-quarters with either pre-announcement guidance or multiple guidances, respectively. All of them are consistent with those of the full sample and the estimated coefficients are qualitatively the same and only different in magnitude. The results in columns (6), (7) and (8) of Table 8 are counterparts by using the sample in which all variables are winsorized at the 2.5% and 97.5% of their respective distributions.

Cubic Approximations In section 5.1, we raise the question that the econometrician cannot have access to private information available to analysts and can only construct observable surprises with guidance and initial forecast, i.e., $S_i \equiv y - F_{0i}$. To demonstrate that our model can be used to organize the facts that we uncovered in the data (that is, the relationship between forecast revisions and surprises), we provide an explicit approximation for the relationship between forecast revisions and observable surprises from the econometrician's point of view. To address the problem of unavailability of private information, we derive a theoretical relationship from our model, which connects forecast revisions with managerial guidance surprises observable to the econometrician (i.e., $S_i \equiv y - F_{0i}$) that have counterparts in the data.

The key idea is to re-orthogonalize the information set and construct a stochastic variable e_i to represent the *unpredictable* part of private information given managerial guidance y and initial forecasts F_{0i}, i.e.,

$$e_i \equiv x_i - \underbrace{\left[F_{0i} + \left(\frac{\tau_Y}{\tau_\theta + \tau_z + \tau_x} \right) \underbrace{(y - F_{0i})}_{S_i} \right]}_{\mathbb{E}[\theta|z_i, y]}$$

Observe that e_i is orthogonal to F_{0i} and S_i. Therefore, $\{F_{0i}, S_i, e_i\}$ forms an alternative basis for the space of signals. The distribution of e_i is still normal with mean 0 and variance σ_e^2 being determined by τ_θ, τ_z, τ_x, and τ_Y.[18] For the econometrician, e_i is not observable either, just like the private information x_i. However, the independence of e_i from observable surprises S_i is useful for constructing the relationship between FR_i^* and S_i. Given this re-orthogonalization, we derive

$$s_i = \left(1 - \nu_x \nu_y\right) S_i - \nu_x e_i \qquad\qquad X_i = F_{0i} + \nu_x e_i + \nu_x \nu_y S_i \tag{21}$$

where $\nu_x = \frac{\tau_x}{\tau_\theta + \tau_z + \tau_x}$ and $\nu_y = \frac{\tau_Y}{\tau_\theta + \tau_z + \tau_Y}$.

Proposition 6 (Observable Relationship)**.** *The conditional expectation of optimal forecast revisions* FR_i^* *can be approximated by a cubic polynomial in observable managerial guidance surprises* S_i*, that is:*

$$\mathbb{E}\left[\mathrm{FR}_i^* | S_i\right] \approx c_0 + c_1 \cdot S_i + c_2 \cdot S_i^2 + c_3 \cdot S_i^3, \tag{22}$$

where c_0*,* c_1*,* c_2 *and* c_3 *are constant coefficients, which are determined by preference, utility and distribution parameters. Approximately, the error term* $\xi_i \equiv \mathrm{FR}_i^* - \mathbb{E}\left[\mathrm{FR}_i^* | S_i\right]$ *is on average zero, conditional on* S_i*, that is:* $\mathbb{E}\left[\xi_i | S_i\right] = 0$.

The proof can be found at the end of this appendix. It takes two steps to obtain equation (22). First, we approximate $\kappa^*(X_i, s_i)$ characterized in Proposition 1 with a quadratic Taylor expansion around the origin $(X_i = 0, s_i = 0)$. That is the lowest possible degree of polynomials that can capture asymmetry and non-linearity. Therefore, FR_i^* can be approximated by bivariate cubic polynomials in (X_i, s_i). Second, we

[18] Note that $\mathbb{E}\left[e_i\right] = \mathbb{E}\left[e_i^3\right] = 0$ and $\mathbb{E}\left[e_i^2\right] = \left(\frac{\tau_x}{\tau_\theta + \tau_z + \tau_x}\right)^2 \left(\frac{1}{\tau_\theta + \tau_z + \tau_Y} + \frac{1}{\tau_x}\right)$ is constant.

rewrite FR_i^* with cubic polynomials of (e_i, S_i). That is, we replace (X_i, s_i) with (e_i, S_i), using equation (21). The conditional expectation $\mathbb{E}[\text{FR}_i^*|S_i]$ represents the average forecast revisions of analysts that receive the same observable surprise S_i but idiosyncratic unpredictable private information e_i. As shown in the proof of Proposition 6, the orthogonality between e_i and F_{0i} and between e_i and S_i guarantees that the expectation $\mathbb{E}[\text{FR}_i^*|S_i]$ is approximately cubic polynomial in S_i and that the error term ξ_i is conditionally zero.

Further, our approach offers an analytical characterization of c_0, c_1, c_2 and c_3 and therefore predictions about their signs under various parameter choices.

Proposition 7 (Predictions for Signs of Coefficients)**.**
(i) The coefficients c_0 and c_2 govern the asymmetry of forecast revisions:

$$c_0 = -\lambda\beta\widetilde{\mathbb{V}}\left(\frac{\tau_y}{\tau_\theta + \tau_z + \tau_x + \tau_y}\right)\sigma_e^2;$$
$$c_2 = -\lambda\beta\widetilde{\mathbb{V}}\left(\frac{\tau_y}{\tau_\theta + \tau_z + \tau_x + \tau_y}\right)\left(1 - \nu_x\nu_y\right)^2.$$

(ii) The coefficients c_1 and c_3 govern the non-monotonicity of forecast revisions. If λ is not too large,

$$c_3 < 0;$$

if λ is very large,

$$c_3 \to 0;$$

and if σ_e^2 is not too large,

$$c_1 \approx \gamma + (1-\gamma)\,\widetilde{\mathbb{E}}\left(\frac{\tau_y}{\tau_\theta + \tau_z + \tau_x + \tau_y}\right) > 0,$$

where both expectation $\widetilde{\mathbb{E}}(\cdot)$ and variance $\widetilde{\mathbb{V}}(\cdot)$ are evaluated with the distorted subjective posterior belief $\tilde{p}(\tau_y) \equiv \tilde{p}(\tau_y|X_i = s_i = 0) \propto \sqrt{\delta}\exp\left(-\frac{\lambda\delta}{\tau_\theta+\tau_z+\tau_x}\right)p(\tau_y)$.

The proof can be found at the end of this appendix. The exact analytical expressions of c_1 and c_3 are collected in the proof of this proposition. This proposition explicitly characterizes a relationship between forecast revisions and observable managerial guidance surprises by circumventing the unavailability of unobservable private information. Such an approach is informative, because it helps us explicitly map a combination of model parameters to empirical findings uncovered in the data, in particular results from the cubic regression (20), reported in column (1) of Table 8.

Consider a special case that helps us illustrate the interpretation of those coefficients: analysts are ambiguity neutral ($\lambda = 0$). The forecasting problem is only dictated by the Bayesian rule. As implied by Proposition 7, forecast revisions should be symmetric, since $\lambda = 0$. That is,

$$c_0^B = c_2^B = 0.$$

Given the presence of ambiguity, analysts use the information acquired to update their beliefs about the quality distribution of managerial guidances. Therefore, the forecast revision should be non-monotonic in S_i. That is,[19]

$$c_1^B = \nu_x \nu_y + \left(1 - \nu_x \nu_y\right) \left[\mathbb{E}\left(\frac{\tau_y}{\tau_\theta + \tau_z + \tau_x + \tau_y} \right) - \frac{3}{2} \left(\tau_\theta + \tau_z + \tau_x \right) \mathbb{V}\left(\frac{\tau_y}{\tau_\theta + \tau_z + \tau_x + \tau_y} \right) \sigma_e^2 \right];$$
$$c_3^B = -\left(1 - \nu_x \nu_y\right)^3 \frac{\left(1 + \tau_x\right)}{2} \mathbb{V}\left(\frac{\tau_y}{\tau_\theta + \tau_z + \tau_x + \tau_y} \right) < 0.$$

When σ_e^2 is small, c_1^B is positive, indicating that forecast revisions should be in the same direction of surprises, when the surprises are not large enough. The coefficient for the cubic term c_3^B is negative, implying that forecast revisions decrease in surprises, when they are large enough.

Consider the general case that analysts are ambiguity averse ($\lambda > 0$). If $\beta > 0$, then Proposition 7 (item i) predicts that both c_0 and c_2 should be negative simultaneously. This prediction is consistent with the empirical estimates of c_0 and c_2 (reported in column (1), rows (4) and (2) of Table 8, respectively), which are negative and significantly different from zero. Further, we observe that the magnitude of the estimate of c_0 is rather small (though significant), as compared to that of c_2. It suggests that the variance of e_i or σ_e^2 is very small, because Proposition 7 (item i) implies that $\sigma_e^2 < c_0 / c_2$.

Proposition 7 (item ii) has predictions about the sign of c_1 and c_3 under some conditions: when σ_e^2 is small, c_1 is positive; when the degree of ambiguity aversion is not extreme, c_3 is negative; and when the degree of ambiguity aversion goes to infinity, c_3 converges to zero. To understand why the size and sign of c_3 are informative about analysts' aversion preferences, recall equation (19) in this section. The size of the coefficient for the cubic term relative to that for the quadratic term governs how asymmetric the relation is. If the coefficient for the cubic term is rather small, the coefficient for the quadratic term dominates and the relation is very asymmetric. In contrast, if the coefficient for the cubic term is large relative to that for the quadratic term, the relation will be closer to symmetric.

The corresponding empirical estimates are reported in column (1), rows (1) and (3)

[19]Given $\lambda = 0$, the subjective posterior belief $\tilde{p}(\tau_y)$ is reduced to be an undistorted Bayesian belief $p(\tau_y)$ and therefore, both the expectation and the variance are undistorted.

of Table 8, respectively. The estimated coefficient is significantly positive for the linear term and significantly negative for the cubic term. The former is consistent with the auxiliary prediction that σ_e^2 is indeed small and the latter indicates that the degree of ambiguity aversion is not extremely high and analysts are unlikely to possess maxmin preferences.

In sum, this section demonstrates that our model can be transformed to inform the relationship between forecast revisions and surprises observable to the econometrician and that the model can deliver qualitative features (asymmetry and non-monotonicity) that are consistent with empirical facts from reduced-form and parametric regressions.

Proof of Proposition 6 . The first step is to approximate κ^* with a quadratic Taylor expansion around the point $s_i = 0$ and derive the following bivariate cubic polynomials in (X_i, s_i):

$$\begin{aligned} F_i^* &= X_i + \kappa^*(s_i) s_i \\ &= X_i + \left(\bar{\kappa} + \bar{\kappa}_s s_i + \frac{1}{2}\bar{\kappa}_{ss} s_i^2\right) s_i + \mathbb{O}\left(s_i^4\right). \end{aligned}$$

where $\bar{\kappa}$, $\bar{\kappa}_s$, and $\bar{\kappa}_{ss}$, are constant coefficients and $\mathbb{O}(\cdot)$ is the sum of other higher order terms.

In what follows, we derive analytical expressions of $\{\bar{\kappa}, \bar{\kappa}_s, \bar{\kappa}_{ss}\}$. Let $\delta \equiv \tau_y/(\tau_\theta + \tau_z + \tau_x + \tau_y)$ and the equilibrium distorted belief $\tilde{p}\left(\tau_y|s_i;\kappa^*\right) \equiv \tilde{p}\left(\tau_y|X_i, s_i; F_i^*\right)$ is such that

$$\tilde{p}\left(\tau_y|s_i;\kappa^*\right) \propto \sqrt{\delta}\exp\left(H^*\left(\delta, s_i;\kappa^*\right)\right) p\left(\tau_y\right),$$

where $p\left(\tau_y\right)$ is analysts' common prior belief over $\tau_y \in \Gamma_y$ and the function $H^*\left(\delta, s_i;\kappa^*\right)$ is quadratic in (δ, s_i):

$$H^*\left(\delta, X_i, s_i;\kappa^*\right) = -\lambda\left[\beta s_i\delta - \left(\delta^2 - 2\kappa^*\delta\right)s_i^2 + \left(\frac{1}{\tau_\theta + \tau_z + \tau_x}\right)\delta\right] - \frac{1}{2}\left(\tau_\theta + \tau_z + \tau_x\right)s_i^2\delta.$$

Then, the equilibrium κ^* can be compactly written as

$$\kappa^*\left(s_i\right) = \int_{\Gamma_y}\delta\tilde{p}\left(\tau_y|s_i;\kappa^*\right)d\tau_y$$

First, we obtain $\bar{\kappa}$ by evaluating $\kappa^*\left(s_i\right)$ at $s_i = 0$:

$$\bar{\kappa} \equiv \kappa^*\left(s_i\right)|_{s_i=0} = \int_{\Gamma_y}\delta\tilde{p}\left(\tau_y\right)\mathrm{d}\tau_y = \widetilde{\mathbb{E}}\left(\delta\right) > 0,$$

where

$$\tilde{p}\left(\tau_y\right) \equiv \tilde{p}\left(\tau_y | s_i = 0\right) \propto \sqrt{\delta} \exp\left(-\frac{\lambda\delta}{\tau_\theta + \tau_z + \tau_x}\right) p\left(\tau_y\right).$$

Second, $\partial\kappa^*/\partial s_i$ is calculated as follows:

$$\begin{aligned}
\frac{\partial\kappa^*}{\partial s_i} &= \int_{\Gamma_y} \delta \frac{\tilde{p}\left(\tau_y | s_i; \kappa^*\right)}{\partial s_i} \mathrm{d}\tau_y \\
&= \int_{\Gamma_y} \left\{-\lambda\left[\beta\delta - 2\left(\kappa^* - \delta\right)^2 s_i\right] - \left(\tau_\theta + \tau_z + \tau_x\right) s_i \delta\right\} \delta \tilde{p}\left(\tau_y | s_i; \kappa^*\right) \mathrm{d}\tau_y \\
&+ \kappa^* \int_{\Gamma_y} \left\{\lambda\left[\beta\delta - 2\left(\kappa^* - \delta\right)^2 s_i\right] + \left(\tau_\theta + \tau_z + \tau_x\right) s_i \delta\right\} \tilde{p}\left(\tau_y | s_i; \kappa^*\right) \mathrm{d}\tau_y \\
&- 2\lambda s_i^2 \frac{\partial\kappa^*}{\partial s_i} \int_{\Gamma_y} \left(\delta - \kappa^*\right)^2 \tilde{p}\left(\tau_y | s_i; \kappa^*\right) \mathrm{d}\tau_y.
\end{aligned}$$

Evaluating this at $s_i = 0$, we have:

$$\bar{\kappa}_s \equiv \left.\frac{\partial\kappa^*}{\partial s_i}\right|_{s_i=0} = -\lambda\beta \int_{\Gamma_y} \left(\delta - \kappa^*\right)^2 \tilde{p}\left(\tau_y\right) \mathrm{d}\tau_y \equiv -\lambda\beta \widetilde{\mathbb{V}}\left(\delta\right) < 0.$$

Third, three second-order derivative is calculated as follows:

$$\begin{aligned}
\frac{\partial^2\kappa^*}{\partial s_i^2} &= \int_{\Gamma_y} \delta \frac{\partial^2 H^*}{\partial s_i^2} \tilde{p}\left(\tau_y | s_i; \kappa^*\right) \mathrm{d}\tau_y + \int_{\Gamma_y} \delta \left(\frac{\partial H^*}{\partial s_i}\right)^2 \tilde{p}\left(\tau_y | s_i; \kappa^*\right) \mathrm{d}\tau_y \\
&- 2 \int_{\Gamma_y} \delta \frac{\partial H^*}{\partial s_i} \tilde{p}\left(\tau_y | s_i; \kappa^*\right) \mathrm{d}\tau_y \int_{\Gamma_y} \frac{\partial H^*}{\partial s_i} \tilde{p}\left(\tau_y | s_i; \kappa^*\right) \mathrm{d}\tau_y + 2\kappa^* \left(\int_{\Gamma_y} \frac{\partial H^*}{\partial s_i} \tilde{p}\left(\tau_y | s_i; \kappa^*\right) \mathrm{d}\tau_y\right)^2 \\
&- \kappa^* \int_{\Gamma_y} \frac{\partial^2 H^*}{\partial s_i^2} \tilde{p}\left(\tau_y | s_i; \kappa^*\right) \mathrm{d}\tau_y - \kappa^* \int_{\Gamma_y} \left(\frac{\partial H^*}{\partial s_i}\right)^2 \tilde{p}\left(\tau_y | s_i; \kappa^*\right) \mathrm{d}\tau_y \\
&= \widetilde{Cov}\left(\delta, \frac{\partial^2 H^*}{\partial s_i^2}\right) + \widetilde{Cov}\left(\delta, \left(\frac{\partial H^*}{\partial s_i}\right)^2\right) - 2\left(\int_{\Gamma_y} \frac{\partial H^*}{\partial s_i} \tilde{p}\left(\tau_y | s_i; \kappa^*\right) \mathrm{d}\tau_y\right) \widetilde{Cov}\left(\delta, \frac{\partial H^*}{\partial s_i}\right).
\end{aligned}$$

where $\widetilde{Cov}\left(\cdot, \cdot\right)$ denotes the covariance under the equilibrium distorted posterior belief $\tilde{p}\left(\tau_y | s_i; \kappa^*\right)$. Evaluating it at $s_i = 0$, we have:

$$\bar{\kappa}_{ss} \equiv \left.\frac{\partial^2\kappa^*}{\partial s_i^2}\right|_{X_i = s_i = 0} = \left[\lambda^2\beta^2 + 2\lambda\right] \widetilde{Cov}\left(\delta, \delta^2\right) - \left[2\lambda^2\beta^2\bar{\kappa} + 4\lambda\bar{\kappa} + \left(\tau_\theta + \tau_z + \tau_x\right)\right] \widetilde{\mathbb{V}}\left(\delta\right).$$

Therefore, FR_i^* can be written as

$$FR_i^* \equiv F_i^* - F_{0i} = X_i - F_{0i} + \left(\bar{\kappa} + \bar{\kappa}_s s_i + \frac{1}{2}\bar{\kappa}_{ss} s_i^2\right) s_i + \mathbb{O}\left(s_i^4\right).$$

Replace X_i and s_i using (21). We derive a two-variable cubic polynomial in (e_i, S_i)

$$FR_i^* = c_{i0} + c_{i1}S_i + c_{i2}S_i + c_{i3}S_i + \mathbb{O}\left(S_i^4, e_i^4\right)$$

where

$$\begin{aligned}
c_{i0} &= (1-\bar{\kappa})\,\nu_x e_i + \bar{\kappa}_s \nu_x^2 e_i^2 - \frac{1}{2}\bar{\kappa}_{ss}\nu_x^3 e_i^3 \\
c_{i1} &= \left[\nu_x\nu_y + \bar{\kappa}\left(1-\nu_x\nu_y\right)\right] - 2\bar{\kappa}_s\left(1-\nu_x\nu_y\right)\nu_x e_i + \frac{3}{2}\bar{\kappa}_{ss}\left(1-\nu_x\nu_y\right)\nu_x^2 e_i^2 \\
c_{i2} &= \bar{\kappa}_s\left(1-\nu_x\nu_y\right)^2 + \frac{3}{2}\bar{\kappa}_{ss}\left(1-\nu_x\nu_y\right)^2\nu_x e_i \\
c_{i3} &= \frac{1}{2}\bar{\kappa}_{ss}\left(1-\nu_x\nu_y\right)^3
\end{aligned}$$

Given that we know the property that $\mathbb{E}\left[e_i|S_i\right] = \mathbb{E}\left[e_i\right] = 0$, $\mathbb{E}\left[e_i^2|S_i\right] = \mathbb{E}\left[e_i^2\right] = \sigma_e^2$, and $\mathbb{E}\left[e_i^3|S_i\right] = \mathbb{E}\left[e_i^3\right] = 0$, we derive the following:

$$\begin{aligned}
c_0 &= \mathbb{E}\left[c_{i0}|S_i\right] = \bar{\kappa}_s\sigma_e^2 \\
c_1 &= \mathbb{E}\left[c_{i1}|S_i\right] = \left[\nu_x\nu_y + \bar{\kappa}\left(1-\nu_x\nu_y\right)\right] + \frac{3}{2}\bar{\kappa}_{ss}\left(1-\nu_x\nu_y\right)\nu_x^2\sigma_e^2 \\
c_2 &= \mathbb{E}\left[c_{i2}|S_i\right] = \bar{\kappa}_s\left(1-\nu_x\nu_y\right)^2 \\
c_3 &= \mathbb{E}\left[c_{i3}|S_i\right] = \frac{1}{2}\bar{\kappa}_{ss}\left(1-\nu_x\nu_y\right)^3.
\end{aligned}$$

Finally, given that e_i and y are uncorrelated, $\mathbb{E}\left[\xi_i|y\right] = 0$, where

$$\xi_i \equiv \left(c_{i0}-c_0\right) + \left(c_{i1}-c_1\right)y + \left(c_{i2}-c_2\right)y^2 + \left(c_{i3}-c_3\right)y^3.$$

■

Proof of Proposition 7. Part (i) of Proposition 7 directly follows the characterization in the proof of Proposition 6.

To prove part (ii), first of all, it is straightforward to show the following:

$$\begin{aligned}
\lim_{\lambda\to 0} c_3 &= -\frac{1}{2}\left(1-\nu_x\nu_y\right)^3\left(\tau_\theta+\tau_z+\tau_x\right)\lim_{\lambda\to 0}\widetilde{\mathbb{V}}\left(\delta\right) < 0 \\
\lim_{\sigma_e\to 0} c_1 &= \nu_x\nu_y + \bar{\kappa}\left(1-\nu_x\nu_y\right) > 0
\end{aligned}$$

In addition, it can be shown that when $\lambda \to +\infty$, $\tilde{p}\left(\tau_y\right)$ converges to a degenerate distribution, which implies

$$\lim_{\lambda\to+\infty}\widetilde{\mathbb{V}}\left(\delta\right) = \lim_{\lambda\to+\infty}\widetilde{Cov}\left(\delta,\delta^2\right) = 0 \Rightarrow \lim_{\lambda\to+\infty}\bar{\kappa}_{ss} = 0$$

Therefore, we have:

$$\lim_{\lambda \to +\infty} c_3 = -\frac{1}{2} \left(1 - \nu_x \nu_y\right)^3 \lim_{\lambda \to +\infty} \bar{\kappa}_{ss} = 0$$

■