Does the Precision of Equity Analysts Matter? Evidence from the Textual Content of Analysts' Reports

Stefano Collina

November 20, 2018

Abstract

I propose that analyst's precision and opinion jointly explain a range of market outcomes, including returns, volume, and volatility, of the publication of an analyst report. I construct a novel measure of precision based on textual analysis of equity analysts' reports. I find that for pessimistic reports, higher precision is associated with a significantly larger negative price reaction. Moreover, the higher precision is associated with higher abnormal turnover, higher volatility, and lower change in uncertainty. However, precision is not significantly or only weakly correlated with market reaction for optimistic reports. I argue that this dichotomy is a result of the well-known optimism bias of equity analysts and of a tendency of analysts to inflate the precision of more optimistic reports. I also show that the relation between precision and price reaction varies depending on the information environment and on textual characteristics of the analyst report.

The latest version of this paper is available through this link

I am particularly indebted to Julien Cujean, Laurent Fresard, Steve Heston, and Nagpurnanand Prabhala for their guidance and advice. I am also grateful to Gurdip Bakshi, Maria Cecilia Bustamante, John Chao, Ruyun Feng, Bo Hu, Mark Loewenstein, Shrihari Santosh, Pablo Slutzky, Liu Yang, Jinming Xue, and all the participants of the University of Maryland, finance department brownbag seminars for the helpful comments.

1 Introduction

Equity sell-side analysts have been widely studied in financial and accounting literature. This is unsurprising since analysts play an important role in producing information for financial markets. The research produced by analysts includes numerical or categorical output such as earnings forecasts and buy or sell recommendations. However, analysts also produce extensive textual output in the form of reports. These reports are read and used by market participants to form their expectations and to set prices. In this paper, I focus on the textual output produced by equity analysts.

The primary features of an analyst report, and of any analyst output, are the tone – or analyst's opinion – and the precision. While opinions of analysts and the tone of reports have been objects of research in several works (e.g., Bradshaw 2011 and Huang et al. 2014), studies that discuss the precision of a single analyst's output are scarce since analysts largely do not provide any information about their precision. But in models on information diffusion in financial markets (e.g., Kim and Verrecchia 1991 and Subramanyam 1996), precision is an important dimension. Precision captures the informativeness of a signal and investors use it to weight an analyst opinion.

This paper studies the precision of equity analyst reports, i.e., the uncertainty and noise of the information contained in an analyst report. I employ textual analysis to construct a measure of precision based on the words present in the text of analysts' reports. This text-based measure can be calculated in a consistent way for any sufficiently-long analyst report and it does not require analysts to provide a numerical measure of precision. I employ the "uncertainty" word list of Loughran and McDonald (2011) to define the measure of precision as the percentage of sentences not containing any of these words. I also use the Loughran and McDonald (2011) dictionaries to construct a measure of tone defined as the difference between the frequency of positive and negative words.

I study the relation between the precision of the information contained in analyst reports and market outcomes, including returns, volume, and volatility, of their publication. Models of information diffusion suggest that investors consider a more precise report to be more informative and thereby place more weight on the analyst report and respond more strongly to its publication. In other words, precision has a multiplicative effect on price reaction, i.e., it amplifies or dampens the investors' response to the publication of a new report.[1] The extent of this relation depends also on the precision of previously held information: A higher precision of priors would mean that a new signal, independently from its precision, is less important for investors and that, as a result, these reports would be weighted less and priors weighted more. All of these would lead to report's precision being less important. A further result of information diffusion models is that precision has a similar multiplicative effect on such other measures of market reaction as volume and volatility.

A key limitation of these arguments is their assumption that investors observe the true, undistorted opinion and precision of the analyst. The well-known optimism bias of analysts means that optimistic opinions are considered of limited informativeness per se and, more generally, that the informativeness of a report depends on its degree of bias in addition to its precision.[2] In other words, a more precise, especially optimistic, report may be not significantly more informative if it is also highly or more biased. Furthermore, investors only observe a noisy and possibly distorted signal about precision. Thus, a more precise report could be not more informative if the signal about precision is distorted upward, i.e., if analysts, consciously or not, inflate the precision of their information and, hence, the signal poorly reflects the true precision. Moreover, the main argument assumes that investors have unlimited attention and are able to fully assess the precision of an analyst report. Instead, investors may not pay attention or may pay extra attention to just some reports or to just some parts of these, resulting in a weaker (or stronger) relation between precision and investors' reaction for them.

Following the above arguments, I hypothesize that the precision of an analyst report is positively related to the magnitude of the price and market reaction to its publication. I also argue that the relation between price reaction and precision as well as the stronger price reaction to more precise reports are not quickly reversed, as non-informational trading theories (e.g., DeLong et al. 1990, Tetlock 2007, and Tetlock 2011) would imply. I also hypothesize that the relation between precision and price and market reaction is weaker for optimistic reports because of the optimism bias of analysts. For instance, Womack (1996) and Malmendier and Shantikumar (2007) show that investors consider optimistic analysts' opinions to have limited informativeness and that they weakly

[1] Intuitively, the multiplicative nature of the precision effect also means that precision plays a limited role if a report is considered largely uninformative per se (e.g., a non-opinionated report), since the response would be null or weak to begin with.

[2] For example, see Lin and McNichols (1998), Michaely and Womack (1999), Jackson (2005), Barber et al. (2007)

react to them. I also hypothesize that this weaker relation for optimistic reports is driven by the existence of a positive correlation between report textual precision and report tone, suggesting that textual precision is inflated for more positive reports and that it does not fully reflect true precision.

I also test whether the relation between precision and price reaction is stronger for reports issued during periods of higher firm uncertainty. Indeed, besides the aforementioned role of the precision of priors, this relation could arise because investors pay more attention to analysts' output when uncertainty is higher (Loh and Stulz 2017). The relation between precision and price reaction should also depend on what part of the report is used to measure the report's tone and precision. Specifically, I argue that a strong tone at the beginning of a report would drive the investors' attention away from the rest of the report, resulting in a weaker role of precision. The last portion of a report, often employed to discuss risks and model assumptions, could better capture the precision perceived by limited-attention investors and better explain their reaction to the publication of a new report.

My sample comprises of analyst reports regarding 290 S&P500 firms published between 2003 and 2015. I study the relation between precision and price reaction, defined as the cumulative abnormal returns between the day the report was issued and one to six trading days afterward. I also study abnormal turnover, realized volatility, and change in the implied volatility of at-the-money options with 30 days to maturity.

I focus on reports published on days when the analyst issued an earnings forecast in order to exclude reports that may be deemed not important by investors and, hence, that may not be read by several investors. For the main analysis, I also exclude any report published in the five days centered around earnings announcements or management guidance issuance, in order to avoid capturing the reaction to major overlapping events. I sort each report into three *Tone* categories: pessimistic (bottom 25th percentile), neutral, and optimistic (top 25th percentile). Furthermore, for each month, I also sort each of these reports into three *Precision* categories: low (bottom sextile), medium, and high precision (top sextile).

I observe that precision peaks on days when earnings or management guidance are announced, while it decreased sharply during the 2008-2009 financial crisis. These results suggest that the measure is related to the availability of clear and high-quality information about a firm. Consistent with the idea that precision also captures firm-level uncertainty, I find that precision is related to idiosyncratic volatility and other proxies such as firm size and number of analysts covering the firm. I also find a positive relation between precision and tone, i.e., optimistic reports are disproportionately more highly precise and disproportionately less lowly precise than pessimistic ones. This relation suggests that analysts inflate the precision of more optimistic reports, possibly to embellish them.

I also find that precision is positively related to the ex-post accuracy of the simultaneously issued earnings forecasts, i.e., more precise reports also appear to be more accurate. However, I find that high precision is weakly related to higher accuracy for more optimistic reports; this is consistent with the idea that analysts inflate their precision for these reports.

I find an economically and statistically significant relation between price reaction and precision for pessimistic reports. Indeed, high precision is associated with a larger negative price reaction of between 40 and 95 bps than low precision, depending on time horizon and empirical specification. On the other hand, this relation, as well as the overall price reaction, is largely insignificant for optimistic reports. This finding is consistent with the idea that investors, at least partially, see through the optimism bias and precision inflation of optimistic reports and, hence, discount them accordingly. These results, especially for longer horizons, are robust to alternative $Precision$ sorting, to using stricter definitions of included dates, and to controlling for report topics. I also find that the magnitude of the difference in price reaction is larger for the longer horizons and that the price reaction to highly precise pessimistic reports is largely unchanged, even at horizons of one trading week or longer. The absence of a significant reversal is consistent with the idea that the larger price reaction is due to a higher informativeness and not only due to overreaction driven by sentiment or behavioral-biases-induced trading (e.g., Tetlock 2007 and Tetlock 2011).

I also find a relation similar to the one between precision and price reaction for other measures of market reaction. Indeed, higher precision is associated with a significantly higher abnormal turnover and with higher realized volatility, suggesting that the publication of higher quality reports is associated with increased market activity. On a different note, I find that greater precision is associated with a smaller change in implied volatility, suggesting that more precise reports are associated with a reduction or, at least, no increase in uncertainty. As observed for price reaction, all these relations appear to be especially significant for pessimistic reports.

I also find that the relation between price reaction and precision is stronger when firm uncertainty is higher. Indeed, the difference in price reaction between high and low precision pessimistic reports is around 1 percent higher when idiosyncratic volatility is above the median. Consistent with theoretical models, this finding can be explained by the fact that a highly precise signal is particularly valuable when uncertainty about a firm is higher, i.e., when its relative precision is also high. Furthermore, this result is also consistent with the fact that investors pay more attention to analysts during periods of higher uncertainty (Loh and Stulz 2017) and hence they are able to better discern the degree of precision of the reports.

Consistent with the hypothesis that investors' attention could affect the relation between informativeness and precision, I observe that the portion of an analyst report used to measure tone and precision matters. The relation between precision and price reaction is weaker if the tone is calculated based on the first 30 sentences. This result suggests that investors pay less attention to report precision when the beginning of a report displays a strong opinion. This is especially true for pessimistic reports since strong negative opinions could be interpreted by investors as a good-enough signal (Joos and Piotroski 2016). Instead, especially at longer horizons, I find that the relation is stronger if precision is measured based on the last 30 sentences of the reports. Indeed, for long reports, the last pages often contain information that is of direct use for assessing precision (e.g., discussions about risk or modeling choices); limited-attention investors may focus specifically on these parts.

1.1 Literature and Contributions

My main contribution is to the literature about the informativeness of analyst reports and, more generally, to that of professional financial forecasters. Consistently with both the literature that claims that analysts do not produce informative output (e.g., Altınkılıc et al. 2009, 2013) and research that argues that analysts produce valuable content (e.g., Bradley et al. 2014 and Li et al. 2015), I find that a sizable number of reports are largely uninformative, but also that some reports, especially highly precise pessimistic reports, contain useful information. Indeed, I observe that the publication of these latter reports is associated with a significant and persistent larger price reaction and, more generally, market reaction (abnormal turnover and volatility). In particular, I show that informativeness varies not only along a measure of analyst opinion (e.g., Womack 1996), but also along a measure of analyst precision. These results suggest that the research studying the informativeness of analysts' output should take into account both the analyst's opinion and the analyst's precision. These findings also suggest that analysts' precision is a variable of interest for researchers using analysts' output to answer broader questions, such as the literature that uses analysts' disappearance as a shock to information asymmetry (e.g., Kelly and Ljungqvist 2012 and Derrien and Kecskec 2013), since analysts' precision is related to the value of this output for investors.

I also bring new evidence to the literature about the characteristics of analysts' output. While earnings forecasts and recommendations have been largely analyzed in existing literature, there is limited research about analysts' reports. We have very limited knowledge of the characteristics of analysts' precision. Similarly to what observed by Joos and Piotroski (2016) about the "alternative scenarios" provided by some brokers, I find that precision is related to a series of firm characteristics related to uncertainty as well as to ex-post earnings forecast accuracy. However, I also find that (textual) precision is correlated with optimism, suggesting that not are only analysts' opinions distorted upward (e.g., McNichols and O'Brien 1997), but also that precision is characterized by a similar distortion for more optimistic opinions. A better knowledge of analyst reports is especially important given the industry trend toward a business model where investors have to purchase these reports.

This study also contributes to the growing literature about the textual content of analysts' reports. Current literature has focused on the tone of analysts' reports (Asquith et al. 2005 and Huang et al. 2014), their novelty with respect to conference calls (Huang et al. 2017), and the use of misleading wording (Bellstam 2017). My work adds results concerning the degree of precision or uncertainty that stems from the text of analysts' reports. This measure captures a more general definition of uncertainty than measures of "assertiveness" (Huang et al. 2014) and "weaseling" (Bellstam 2017). Moreover, I provide further evidence concerning the existence of a significant relation between the textual content of analysts' reports and different market variables, such as turnover, realized volatility, and implied volatility, besides the commonly used price reaction. Finally, differently from part of the existing research, I show the importance of identifying what part of analysts' reports is employed to calculate textual measures. Indeed, my results vary depending on whether the full text or only some specific parts are employed, consistent also with the idea that investors' limited attention plays a role in the relation between textual content and investors' reaction.

Finally, this paper contributes to the literature about uncertainty in financial and accounting documents. Existing research has studied uncertainty in documents produced by management such as 10-Ks (Loughran and McDonald 2011), IPO-related filings (Loughran and McDonald 2013), earnings announcements (Demers and Vega 2014), and earnings conference calls (Huang et al. 2017). I show that textual analysis and, specifically, the dictionaries of Loughran and McDonalds (2011) can be used to measure uncertainty in analysts' reports. My results also suggest that the same methodology can be extended to other information produced by media or professional forecasters such as credit analysts.

The rest of the paper is structured as follows. In Section 2, I present my main hypotheses. Section 3 focuses on the construction of my data, while Section 4 discusses the construction of the textual measures as well as their characteristics. Section 5 discusses the main results concerning the relation between precision and price reaction and Section 6 discusses the relation with other market variables. In Section 7, I study the price reaction precision relation varies and in Section 8, I present some robustness results. Finally, I conclude the paper in Section 9.

2 Hypotheses Development

While much of the existing literature has focused on the point estimates produced by equity analysts. However, precision also matters (e.g., Kim and Verrecchia 1991, Subramanyam 1996). Bayesian investors use precision to "weight" a signal they receive about a stochastic payoff, such as firm fundamentals. Investors place more weight on signals that are more precise and, hence, more informative. A consequence is that the magnitude of the price reaction to a signal about firm fundamentals, such as an analyst report, is increasing in its precision. Precision has a multiplicative effect on the price reaction to the publication of a report: High precision amplifies this reaction and low precision dampens it.[3] Intuitively, this multiplicative effect also means that precision play a limited role when the signal is perceived to be not informative per se, such as a non-opinionated report, since the price reaction is null or weak to begin with.

Hypothesis 1: The magnitude of the price reaction to an analyst report publication is increasing in its precision.

[3] In other words, report's precision "modulates" the relation between tone and investors' response.

Much of the literature about the relation between analysts output and returns (e.g., Huang et al. 2014) focuses on very short-term market reaction measures. I also examine longer horizons in order to distinguish between informational trading and sentiment or behavioral-driven trading (Tetlock 2007, 2011, Heston and Sinha 2017). Indeed, theories suggest that substantial price reaction can be generated in the short term by sentiment trading (e.g., DeLong et al. 1990 or Campbell et al. 1993), by behavioral-biases-induced trading (e.g., overconfidence, Barberis 2018) or by salience (e.g., Merton 1987 with Duffie 2010), these reactions, however, are reversed at longer horizons. For instance, existing literature has observed that, particularly less sophisticated, investors overreact to bad news (Tetlock 2007) or to stale information (Tetlock 2011); in these scenarios, returns are fully reversed or almost reversed within a trading week.

Consistent with this argument, if more precise reports are actually more informative, then their effect on prices should not be very short lived. Similarly, the relation between precision and price reaction should not disappear within a few trading days horizon.

Hypothesis 1a: The relation between price reaction and precision is not reversed within a trading week.

Theoretical models suggest that precision is related not only to price reaction but also to other measures of market reaction. For instance, Kim and Verrecchia (1991) argue that precision is positively related to both trading volume[4] and volatility (as in variance of price change); the higher the quality of the information released, the stronger the reaction of traders. In other words, both turnover and volatility should, at least temporarily, increase after more precise information is released. However, it is worth pointing out that volume cannot be employed to rule out sentiment or behavioral motives since theory suggests that they also produce substantial volume (e.g., Tetlock 2007).

Hypothesis 1b: Turnover and stock volatility reaction to an analyst report publication are increasing in its precision.

One of the main assumptions behind the previous hypotheses is that the reports reflect the actual expectations of the analysts, i.e., there is no bias in their opinions. This assumption means that, for instance, optimistic reports actually reflect optimistic information. However, existing literature about equity analysts has highlighted the existence of an optimism bias of analysts (e.g., McNichols

[4]The main assumption is that there is disagreement about precision before the information is released

and O'Brien 1997, Michaely and Womack 1999) that has partially persisted even after regulations were implemented in the early 2000s (e.g., Jackson 2005, Barber et al. 2007). Furthermore, literature has highlighted the fact that investors are able to see through the bias and, accordingly, heavily discount and react less strongly to optimistic opinions (Womack 1996, Malmendier and Shantikumar 2007, Huang et al. 2014).

The existence of this bias means that optimistic reports may actually reflect only neutral or weakly positive information. For instance, Malmendier and Shantikumar (2007) found that sophisticated investors treat positive stock recommendations as neutral. In this case, the analyst opinion itself would be considered by investors not particularly informative and precision, given the multiplicative nature of its effect, would play a limited role.

Another main assumption is that investors know the true precision of the signal. However, investors only observe a noisy signal about signal's precision, for instance, from the text of the report. My results highlight the existence of a positive correlation between the report's textual precision and the report's tone: Optimistic and, to a lesser extent, neutral reports appear to be more precise than pessimistic ones. This correlation suggests that analysts, consciously or as a result of "overprecision" (Moore and Healy 2008), inflate the precision of their reports when they are more optimistic. Consequently, more (textually) precise optimistic reports are not necessarily more precise and more informative. Investors who can see through this precision inflation will discount the observed precision and will not react more strongly.

Hypothesis 2: The relation between price reaction and report precision is weaker for optimistic reports.

The previous hypothesis concerning volume and volatility can be generally extended to account for this dichotomy between optimistic and pessimistic reports. In particular, if the signal about precision that investors receive is of poor quality due to precision inflation, then the relation between precision and volume or volatility should be similar to the relation between precisions and price reaction, i.e., less strong. However, it is also worth mentioning that, even if the relation with price reaction is weaker, this is not necessarily true for volume; an increase in volume can also be spurred by disagreement about the interpretation of a signal (e.g., Kandel and Person 1995, Kim and Verrecchia 1994, 1997).

Hypothesis 2a: The relation between volume or volatility and report precision is weaker for *optimistic reports.*

Investors' reactions to a signal depend not only on the precision of the signal but also on the precision of their priors. When uncertainty is low and investors have prior information that is highly precise, signals are generally not valuable and investors put more weight on their priors. Thus, investors largely ignore new signals and signal's precision, as well as its multiplicative effect, matter less. In other words, the magnitude of the relation between price reaction and precision is increasing in uncertainty[5].

The relation between precision and price reaction may also be affected by investors' attention. A now quite extensive line of accounting and finance research has highlighted the role that inattention plays in explaining different empirical facts (Barberis 2018). Since analysts largely do not provide information about their precision, investors have to read and analyze their reports to understand how precise they are. An inattentive investor may not be able to fully capture the degree of a report's precision and, hence, not react according to it. For instance, research has highlighted that investors pay more attention to financial information (Ouimet and Tate 2017) and, specifically, to analysts' output (Loh and Stulz 2017) in periods of higher uncertainty.

Hypothesis 3b: The relation between price reaction and report precision is stronger when firm uncertainty is higher.

The fact that investors' attention plays a role in the relation between price reaction and report content suggests that stylistic characteristics of the reports may affect the price reaction to differentially precise reports. Indeed, recent papers have shown that text characteristics (e.g., Huang et al. 2014, Zhou 2018) affect investors' reaction to new information. In particular, part of this research has highlighted the importance of focusing on some specific parts of a text such as the title or headline (e.g., Huang et al. 2014, Umar 2017) since investors may focus their attention only on these.

Some reports produced by analysts are significantly long; investors with limited-attention may focus only on specific parts of these. For instance, the first page or two usually represents a summary of the analyst's opinion. Limited-attentive investors may focus specifically on this part of the report to gauge the analyst's opinion. If the tone is strong enough, the investors may partially disregard

[5]For instance, in Subramanyam (1996) the sensitivity of price reaction to precision is decreasing in the precision of the payoff of the risky asset

its precision. This phenomenon is particularly important for pessimistic opinions, since a strong negative opinion could already be considered highly informative (Joos et al. 2016).

Hypothesis 4a: The relation between price reaction and report precision is weaker if the analyst displays a strong opinion, especially a pessimistic opinion, at the beginning of the report.

The ending part of reports often contains conclusions, discussions about modeling choices, risks involved with the forecast, and general firm characteristics. Limited-attentive investors may focus specifically on this last part of the report to gauge the report's precision, since this is the part where analysts are more prone to discuss the uncertainty surrounding their analysis. In other words, the last part of the report may better capture the report's degree of precision and, possibly, the precision perceived by limited-attentive investors.

Hypothesis 4b: The relation between price reaction and report precision is stronger if precision is measured based on the ending of the report.

Here it is important to mention that attention may also play a role against this hypothesis. Indeed, inattentive and less sophisticated investors may just focus on the first part and disregard the ending of these longer reports. In this scenario, the relation may be actually weaker.

3 Data

I start with all non-financial firms that are in the S&P500 at the end of 2015 or were in the S&P500 for at least four years between 2003 and 2015. The choice of the starting date is two-fold. First, I will base my full analysis on a post-RegFD and post-Global Research Settlement period. Second, the quality of textual data in older report is lower since many are just scannerized copies and, hence, it is hard to extract the text. I obtain equity analysts reports from Thomson One Investext database. This is not a comprehensive database since not all brokers are included in all the periods, but it covers a significant part of the reports universe[6]. I also exclude firms with less than 100 reports. Firms appear in Thomson One database with name, ticker and CUSIP. I use the CUSIP to match reports to other data since it is the most reliable identifier. However, Thomson One does not store the CUSIP for all the firms. To be conservative, I exclude firms for which Thomson One does not report the CUSIP[7]. This selection leaves me with 290 firms for which

[6]My assumption is that my reports sample is representative of the whole universe. At the best of my knowledge, there is no reason why it should not be.

[7]Matching data via names is feasible for some firms, but not always reliable.

data is ample and that is easy to match with other sources.

I match the reports to I/B/E/S brokers and analyst names[8] as well as further data about the recommendation and forecasts they issued during the sample period. It is possible to match approximately 95% of the reports available via Investext. I use this matched database to construct my final sample. In particular, I focus on reports published on days when an analyst issued an annual earnings forecast[9]. The idea is to exclude reports that could be considered not particularly important and that are probably read by a limited number of investors. I further exclude reports that are either too short for a reliable textual analysis or for which it is not feasible to extract the textual content, resulting in a final sample of approximately 115,000 reports. Before computing the textual measures, I clear all these reports from disclaimer sections, numerical tables, and other parts that do not contain relevant content (e.g. contact information).

I obtain S&P500 components as well as stock returns data from CRSP. Balance sheet data is obtained from Compustat. Information about earnings announcements, management guidance announcements, and other data about analysts and brokers is obtained from I/B/E/S. Analysts' awards are obtained from Euromoney magazine issues. The economic uncertainty index is obtained from policyuncertainty.com.

4 Textual Measures

4.1 Precision

It is possible to obtain firm-level measures of uncertainty based on the forecasts issued by multiple analysts (e.g., dispersion of analysts' forecasts, Diether et al. 2002, Johnson 2004, and Zhang 2006). However, it is difficult to measure the precision of a specific forecast or opinion. Analysts largely provide only point forecasts or categorical recommendations and not the confidence interval of these. An exception is the "alternative scenarios" framework for earnings or target prices, usually a "bull" and a "bear" scenario, that some brokers (such as Morgan Stanley and Barclays) have started to provide in some of their reports. The major problems are that these measures are available only for few brokers and that their actual implementation varies even within a broker.

[8]See Appendix B for more details about sample construction

[9]The results are largely similar if quarter forecasts are used (around 90% are issued contemporaneously with annual forecasts), but sample is smaller. Results are also largely similar if also recommendation-only reports or also reports issued within two days from the forecast issuance are included

Furthermore, even if the range of scenarios appears to be correlated with firm uncertainty (Joos and Piotroski 2016), these scenarios appear to be often used to display how strongly optimistic the analyst is and are characterized by the same optimism bias of point estimates (Joos and Piotroski 2017), i.e., high certainty about positive outcomes tends to mainly reflect a larger bias.

In this paper, I propose a novel text-based measure of precision of analysts' reports. The benefit of using a text-based methodology is that it can be employed to consistently measure precision. It also does not rely on analyst providing some numerical measure of precision. No assumption is required regarding investors paying attention to, being familiar with, and being able to interpret specific parts of these reports, as it is the case with measure like "alternative scenarios". To the best of my knowledge, this is the first paper that proposes a measure of report-level analysts' precision with such characteristics.

It is here worth mentioning two papers that introduced text-based measures related to precision and how these differ from the measure suggested in this paper. The first is the measure of "assertiveness" suggested by Huang et al. (2014). While it is reasonable to assume some correlation between assertiveness and precision, this measure is mainly geared at identifying reports using a strong language (using words like "lowest", "never", and "strongly") and, hence, reports that want to convey a particularly strong optimistic or pessimistic opinion.[10] The second is the measure of analyst "weaseling" introduced by Bellstam (2017), i.e., a measure of how extensively an analyst uses elusive and possibly misleading language. While there is some overlapping between what can be classified as weaseling and what can be classified as uncertainty (Zerva et al. 2017), weaseling could, for instance, be used to "soften" the language and, more importantly, lack of precision is expressed not just via weaseling.

To construct my measure of precision, I rely on a "bag-of-words" methodology.[11] This methodology relies on the creation of a "dictionary" of words that convey some particular message and on the construction of a measure based on the frequency of these words. To measure precision, I rely on the financial dictionaries proposed by Loughran and McDonald (2011). Specifically, I use their "uncertainty" dictionary, i.e., a set of words that are usually associated with different aspects of uncertainty. For instance, this dictionary has been used to study uncertainty in 10-Ks (Loughran

[10]Xiao and and Zang (2017) finds a strong relation between assertiveness and strong opinions.

[11]For examples of papers using a similar methodology, see Tetlock (2007), Loughran and McDonald (2011), and Huang et al. (2014)

and McDonald 2011), IPO-related filings (Loughran and McDonald 2013), earnings announcements (Demers and Vega 2014), and earnings conference calls (Huang et al. 2017).[12]

I identify sentences that contain any of these uncertainty words and define precision as one minus the frequency of these sentences.[13]

$PrecisionC = 1 - \frac{UncertainSentences}{TotalSentences}$

Appendix C contains some examples of paragraphs containing multiple sentences with uncertainty words. It is interesting to notice that some reports have sections specifically aimed at discussing uncertainty factors potentially affecting their estimates and valuations.

4.2 Tone

To construct a measure of tone of the report, I choose a text-based measure. Specifically, I use the Loughran and McDonald (2011) financial dictionaries to identify positive and negative words. I define tone as the difference between the frequency of positive and negative words. Similar results are obtained by using the differences in frequencies relative to the total number of positive and negative words.

$ToneC = \frac{PosWords - NegWords}{TotalWords}$

I classify the reports in three bins based on this measure of tone. I classify reports as pessimistic if in the bottom 25th percentile, as optimistic if in the top 25th percentile and as neutral if within the 25h and 75th percentiles. The choice of the thresholds assures that reports classified as pessimistic or optimistic really reflect a significantly negative or positive view about a firm[14]. I obtain similar results if the threshold for optimistic reports is chosen in order to mirror the tone threshold for pessimistic reports.

Summary statistics about the two measures are reported in Table 1 (Panel A). Interesting to notice that the positive median suggest that *ToneC* tends to be positive more frequently than negative, even if my sample includes the period during and around the 2008-2009 financial crisis. This result is consistent with the idea that analysts are disproportionately optimistic.

[12] Demers and Vega (2014) is methodologically the closest to my paper. Indeed, the authors use textual measures of optimism and uncertainty to study whether they explain the price reaction to earnings announcements.

[13] I use sentences since it is the primary language unit to express an opinion (Huang et al. 2014) and, given the low frequency of these uncertainty words, it produces a less noisy measure. The results are similar if a simple words frequency is used, especially if a stricter constraint on minimum number of words is applied.

[14] Cumulative abnormal returns are non-significant for pessimistic-leaning reports and significant only for very short horizons for optimistic-leaning reports

4.3 Determinants of Precision

I examine how the measure of precision ($PrecisionC$) varies across time and cross-sectionally.

I examine whether precision varies along the business cycle and across the earnings reporting cycle. It is reasonable to expect that precision is higher when macro uncertainty is low as well as when management has recently released valuable information such as earnings or guidance. In these time periods, analysts have access to more information about a firm and, hence, their reports should be more precise.

Interesting to point out that the precision measure appears to have an upward trend, i.e., analysts' reports precision has increased over time. This is unimportant for the main results since my precision classification is made within month.

Figure 1 shows the average precision around days when earnings or management guidance are announced. Consistently with the previous argument, it is possible to observe that precision peaks on those dates and, then, starts to decrease in the following days. The results are even more striking if looking at the frequency of high or low precision, i.e., the frequency of reports whose precision is in the top or bottom sextile. Figure 1 also reports these results and it shows that while high precision reports frequency peak on those dates, low precision reports frequency reaches its minimum. To summarize, the results suggest that, indeed, analysts' reports appears to be more precise right after analysts obtained information about a firm.

Figure 2 shows the average value of the detrended precision measure. It shows that precision dipped between late 2008 and early 2009, i.e., at the height of the global financial crisis. This is consistent with the argument that higher macro uncertainty should lead to lower precision. Interestingly, the dip is mainly driven by neutral and, particularly, optimistic reports. This is not surprising since it is reasonable to assume that uncertainty was particularly high for any non-negative forecast.

I next examine the cross-sectional determinants of precision. I run a series of regressions of my measure on month fixed effects as well as industry, analyst, and firm fixed effects. The results are reported in Table 2. First it is interesting to notice that both time, industry, and firm fixed effects explain only a small fraction of the variation in precision. The analyst fixed effect appears to explain a significant part of the variation in precision, around 30%. This result suggests that some analysts are more precise than others due to some characteristics or some writing style.

Given these results, I investigate what firm and, especially, analysts characteristics are related to precision. Table 3 presents a series of results for regressions of the precision measure on different firm, analyst, and opinion characteristics. Different sets of fixed effects are included as well as a control for days close to earnings/guidance announcements. First, the results suggest that there is a positive relation between number of analysts covering a firm and precision. This could be explained by the fact that firms covered by several analysts are usually characterized by lower uncertainty since more information is available to the markets. I also observe a significant and negative coefficient for idiosyncratic volatility[15]. This result is not surprising and it confirms that analyst precision varies with firm uncertainty.

The second set of results concern characteristics of the analyst. There is a clear and positive relation with analyst absolute experience, i.e., more experienced analysts appear to issue more precise reports. This result goes hand in hand with the idea that experience is related to forecasting performance (e.g., Clement 1999). The negative relation between broker size and precision is instead puzzling. These results further confirms the need to include analysts fixed effects.

I also estimate similar specifications for analysts' opinion characteristics: boldness (deviation from consensus), staleness (cosine similarity with reports about the same firm published in the previous 90 days), and tone. "Bold" reports are associated with higher precision, suggesting that analysts write more precise reports when in possession of some private information that is not in line with consensus. A similar conclusion can also be drawn by the strong negative relation with staleness of the report, suggesting that analysts precision does not purely stem from herding. I also examine the relation between precision and the incidence of numbers in a report and I find a strong positive relation, suggesting that more precise reports are indeed more quantitative in nature. This results is consistent with the results of Zhou (2018) for corporate disclosures and the idea that lack of numbers is used to mask uncertainty about a statement or forecast.

An equally interesting result is the strong relation between tone and precision. Particularly, optimistic tone is associated with a between 2.7% and 3.8% higher precision than pessimistic reports[16]. In other words, optimistic reports appear to be disproportionately more precise, while

[15]Idiosyncratic volatility is the standard deviation of the residuals of fitting a Fama-French plus momentum model

[16]The definition of tone used here is slightly different than from the rest of the paper due to the fact that a small subset of words is both in the "negative" and "uncertain" list (e.g. "volatile"). To avoid capturing any by-construction

pessimistic reports appear to be disproportionately less precise. Given the well-known optimism bias of analysts, these results suggest the possible existence of a second layer of distortion. Indeed, when producing more optimistic reports, analysts appear to inflate the precision of their information. This behavior is consistent with the idea that there are incentives for analysts to produce flattering and strong reports (e.g., Lin and McNichols 1998, Hong and Kubik 2003, and Barber et al. 2007), while downplaying pessimistic opinions.

Table 4 reports regressions where two dummies are used to identify reports in the top or bottom sextile in terms of precision. The results of the linear probability models are again largely consistent with what previously found about the relation between precision and the different firm, analyst, and report characteristics. Particularly striking is the result for the difference between pessimistic and optimistic reports, where there is a difference of 5-6% in probability of a report being in one of the top/bottom categories. Given a baseline probability of observing a highly or lowly precise report of around 17%, this means that the probability of optimistic reports being of high (low) precision is 30-35% higher (lower) than pessimistic reports.

Interesting is also the result for staleness. On one hand, there is a clear and strong negative relation with the probability of a report being highly precise. However, even if much weaker in terms of magnitude and statistical significance, there is a negative relation also with the probability of being of very low precision. This result suggests that there is some herding effect, s.t. analysts precision is not very low when they can rely on past information.

4.4 Precision and Ex-Post Accuracy

The previous subsections focused on questions concerning the cross-sectional and time-series determinants of precision. I now examine the relation between precision and ex-post accuracy.

Table 5 (Panel A) reports the results of a series of regressions of different measures of earnings forecasts error and standardized *PrecisionC* as well as a series of controls for analyst and firm characteristics (e.g., Stickel 1992 and Clement 1999), as well as forecast age (Brown and Mohd 2003).[17] Regardless of the measure used, the results suggest that precision and forecast error are

effect, I excluded these words when calculating the measure of tone here used. Not surprising, the results are marginally stronger if I include them.

[17]Forecast error is equal to forecast EPS minus actual EPS. "Scaled" means scaled by stock price. "Proportional" means as a proportion of the average forecast error of all forecasts. "Relative" means relative to average forecast error of all forecasts issued in the previous or next 90 days, scaled by forecast standard deviation.

negatively related. For instance, one standard deviation increase in precision is associated with around 1 cent smaller earnings-per-share errors. Table 5 (Panel B) also reports similar results where the previously defined high/low precision dummies are employed as well as their interaction with the *Tone* variable. High precision is associated with smaller forecast error and, albeit at a lower extent, low precision is associated with higher forecast error. However, the former relation is much weaker and, generally, not statistically significant for optimistic and, at a lesser extent, neutral reports. In other words, high precision optimistic reports appear to not be much more accurate than medium precision reports and, for some specifications, than low precision ones. This empirical evidence is again consistent with the idea that the precision of optimistic reports is somewhat inflated.

Lastly, Table 6 reports some other results for the relation between *Tone*, forecast error, and bias. First, neutral and optimistic reports appear to be significantly more positively biased than pessimistic reports, suggesting that the well known optimism bias of analysts' forecasts extends also to the textual content of their reports. Second, neutral and optimistic reports appear to be overall more accurate ex-post. This result is consistent with the finding of Xiao and Zang (2017) that analysts tend to not incorporate into the earnings forecasts the negative information discussed in the reports, resulting in a larger forecast error.

5 Analysts Precision, Price Reaction, and Reports Informativeness

The main question this paper tries to answer is whether there is a relation between the analysts' output precision, the magnitude of the price reaction to its publication, and, more generally, the informativeness of the output they produce.

Following existing finance and accounting literature, I use abnormal returns around the issuance of an analyst report to measure price reaction and, more generally, report informativeness. To avoid capturing the effect of overlapping events, I excluded from the empirical specifications all the reports published in the 5 trading days centered on earnings announcements and management guidance days. Other corporate news is published in other periods, but earnings announcements and guidance (as well as leaks happening the days before) are definitely the major sources of information for analysts to piggyback on.

I study the relation between cumulative abnormal returns, calculated for different horizons around the issuance of an analyst reports, and two categorical variables: *Tone* and *Precision*. I run the following empirical specification for each report j about firm i published at time t:

$$CAR[t, t+n]_{i,t,j} = \alpha + \beta Tone_j + \gamma Precision_j + \delta Tone_j \times Precision_j + \mathbf{v}Controls_{i,j,t} + \epsilon_{i,j,t}$$

CARs are estimated using the Carhart (1997) model for $n = 1, 2, 4, 6$.[18] The *Tone* measure is constructed as previously defined, s.t. it is equal to 0, 1, or 2 for, respectively, neutral, pessimistic and optimistic reports. *Precision* is constructed based on a monthly sort of the *PrecisionC* measure previously defined in three bins; reports in the bottom sextile are classified as highly precise, reports in the top sextile as low precision and all other reports as medium precision.[19] *Precision* is equal to 0, 1 or 2 for, respectively, low, medium or high precision. The list of control variables is reported in Appendix A and varies depending on the specification and includes also the interactions between these variables and both *Tone* and *Precision*. The controls includes different firm and analysts characteristics related to investors' reaction as well as a control for overall uncertainty (Baker et al. 2016).[20] All variables are winsorized at 1% and, excluding prior-CAR and analyst opinion change variables, are standardized. Summary statistics for the non-standardized variables are reported in Table 1 (Panel B). All regressions include also fixed effects; specifically I use analyst-firm pair fixed effect. Results are comparable if using analyst and firm fixed effects.

Table 7 reports the results of variations of this specification. The first two columns include fixed effects as well as the aforementioned controls. Column two, in particular, includes also controls for the change in recommendation or change in earnings forecasts. Differently, column three includes no controls nor fixed effects.

First, it is important to notice that *Tone* is, at least at the shorter horizons, a predictor of the direction the market moves, consistently with Huang et al. (2014). More interesting is the result for the interaction between *Tone* and *Precision*. Consistently with the first hypothesis, at the 2 days horizon and for pessimistic reports, the interaction coefficient is significant and between -40 and -47 bps. This result suggests that high precision is associated with a significantly stronger price reaction

[18]Results are stronger if using a simple market model

[19]Results are similar if thresholds are constructed based on reports published in the 30 days before the report is published.

[20]For example, see Stickel (1995), Gleason and Lee (2003), Jegadeesh and Kim (2010), and Loh and Stulz (2011)

and, hence, report informativeness, for pessimistic reports. The interaction coefficient rises by about 50% at the 3 days horizon. The results are stronger when fixed-effects are included, suggesting that is the within analyst-firm variation in precision that matters more.

I do not observe a significant relation for positive reports, where precision appears to have a largely non-significant effect of, at best, around 6 bps. This result is consistent with the second hypothesis that the presence of an overoptimism as well as an inflation in precision for optimistic reports result in textual precision not being a strong predictor of report informativeness. The interpretation is that investors, at least partially, "see through" these systematic distortions and discount the analysts' opinion and precision accordingly. Generally speaking and consistently with the literature about analysts' bias, the price reaction to optimistic reports appears to be limited.

Table 7b reports the results for longer CAR horizons; 5 and 7 days, respectively. The results are largely similar, but the magnitude of the precision effect for pessimistic reports tends to be greater. The price reaction for highly precise pessimistic reports appears to be unchanged within a trading week and only slightly lower at the longer 7 days horizon. These results support the hypothesis that the price reaction to highly precise reports and the relation between price reaction and precision are not short-lived and are not immediately reversed, suggesting that more precise pessimistic reports are indeed more informative. On the other hand, the price reaction to low precision pessimistic reports is significant only at the very short horizons. This result suggests that while investors appear to trade according to report precision, there is some degree of overreaction to these low precision pessimistic reports, similarly to what observed by literature about news and sentiment (e.g., Tetlock 2007).

As a robustness check, Table 7c reports the results where reports are sorted into precision classification based on quartiles instead of sextiles. The results are again largely consistent with the main results, just slightly weaker in terms of magnitude, consistently with the fact that the difference in precision is smaller. In particular, the price reaction to highly precise pessimistic reports is unchanged even at the longest horizon. Furthermore, although not reported, these main results are also robust to running separate regression for pessimistic and optimistic reports.

6 Analysts Precision, Turnover, and Volatility

The previous section established a relationship between report precision and price reaction, especially for more pessimistic reports. In this section, I investigate whether precision is related to volume and volatility. I test the hypothesis that the diffusion of more precise information is associated with higher volume and volatility.

6.1 Precision and Turnover

I calculate abnormal turnover as the difference between the log turnover and the average log turnover in the 5 days before the report is issued. The results, especially for for pessimistic reports, are the largely similar, if not stronger, if a longer window is used to calculate the average or if the predicted value from a regression of log turnover on market log turnover is used instead of the average (Umar 2017).

Specifically, I run the following empirical specification for each report j about firm i published at time t:

$$CAT[t, t+n]_{i,t,j} = \alpha + \beta Tone_j + \gamma Precision_j + \delta Tone_j \times Precision_j + \mathbf{v}Controls_{i,j,t} + \epsilon_{i,j,t}$$

Where CAT is Cumulative Abnormal Turnover and is calculated for $n = 1, 2, 4, 6$. Controls are the same as in the main specification plus the abnormal turnover in the prior 5 trading days. For the longer five and seven days horizons, I also provide the results excluding any observation whose event window overlap with earnings or guidance announcements[21].

Table 8 reports the results. At the two days horizon, the coefficient of the interaction between $Tone$ and $Precision$ is significant between 0.13 and 0.15, suggesting that higher precision is associated with larger turnover. Similarly to what observed for the price reaction, there is also a dichotomy between pessimistic and optimistic reports; high precision pessimistic reports are associated with a stronger and longer lasting market reaction than high precision optimistic reports. Indeed, abnormal turnover for optimistic reports is significant only in the first two trading days and the results for these reports are not very robust to alternative specifications.

[21] The robustness check section discusses this potential issue for the main price reaction result

6.2 Precision and Volatility

I next study the relation between precision and stock volatility. I estimate the following empirical specification for each report j about firm i published at time t:

$$CSAR[t, t+n]_{i,t,j} = \alpha + \beta Tone_j + \gamma Precision_j + \delta Tone_j \times Precision_j + \mathbf{v}Controls_{i,j,t} + \epsilon_{i,j,t}$$

Where $CSAR$ is the sum of squared abnormal returns and was calculated for $n = 1, 2, 4, 6$. Controls are the same as in the main specification and idiosyncratic volatility. As for the turnover specifications, for the longer five and seven days horizons, I also provide the results excluding any observation whose event window overlap with earnings or guidance announcements[22].

The results are reported in Table 9. At the two days horizon, the interaction coefficient between *Tone* and *Precision* is positive and significant for pessimistic report. This result suggest that high precision is associated with an increase in stock volatility and an increased market activity around the issuance of these more precise reports. As observed for price reaction, the phenomenon is non-significant for optimistic ones.

Next I calculate the change in the implied volatility of 30-days standardized at-the-money options from OptionMetrics around the publication of an analyst report. Following existing literature (e.g., Billings et al. 2015), I calculate the average of the implied volatility of call and put options. Specifically I define abnormal volatility as the difference between the natural log of implied volatility at different horizons after a report publication and the log implied volatility two days before the publication date.

Table 9b reports the results of a series of empirical specifications similar to the previous ones but for this measure of abnormal implied volatility. Besides the covariates included in previous specifications, I also control for market change in volatility by using the change in VIX in the same period. I also restrict the sample such that there is no overlap between any earning/guidance announcement and the event window even at longer horizons, since the former can have a significant effect on uncertainty[23]. The negative sign of the coefficient of the interaction between *Tone* and *Precision* for pessimistic reports suggest the existence of a negative relation between precision and volatility. Albeit not statistically significant, the negative main *Precisions* effect suggests the

[22] The results for the 2-3 days horizons are slightly stronger if the same restriction is applied

[23] The results are largely similar if this restriction is not applied.

existence of a similar relation for neutral reports. On the other hand, similarly to what observed for other specifications, the effect appears to be largely non-significant for optimistic reports. The results also suggest that the diffusion of highly precise pessimistic reports is associated with a smaller increase in firm uncertainty or even a partial resolution of it.[24]

7 Variation in the Relation between Precision and Price Reaction

An interesting question is whether the results observed in the previous section vary. I focus on two main factors that could affect the relative informativeness of highly precise reports. First, the reaction to highly precise reports will tend to depend on the general information environment. Precise reports are particularly more valuable if uncertainty about a firm is high, while investors can just rely on their priors and public information when uncertainty is low. Second, I test whether the effect of precision is greater when investors are more attentive to the whole report text. While the analyst opinion can be obtained from published numerical outputs, precision requires the reading of the whole actual reports. Therefore, precision should play a bigger role when precision or tone are obtained from specific parts of the report text.

7.1 Precision and Firm Uncertainty

An interesting question is how analyst precision and overall firm uncertainty interact. To identify periods of high or low firm uncertainty, I estimate idiosyncratic volatility as the standard deviation of the residuals of the aforementioned Carhart (1997) model. Then, I classify a report as being produced during a period of low firm uncertainty if idiosyncratic volatility is below the median [25]. The $FirmU$ dummy variable takes value 0 if the observation corresponds to a period of low firm uncertainty as previously defined. Specifically, I ran the following specification for each report j about firm i published at time t::

$$CAR[t,t+n]_{i,t,j} = \alpha + \beta Tone_j + \gamma Precision_j + \delta Tone_j \times Precision_j + \zeta Tone_j \times FirmU_{i,t} + \eta Precision_j \times FirmU_{i,t} + \theta Tone_j \times Precision_j \times FirmU_{i,t} + \mathbf{v}Controls_{i,j,t} + \epsilon_{i,j,t}$$

[24] I take an agnostic approach regarding the overall effect of the publication of a report on firm uncertainty. While in a purely Bayesian framework a new signal would always decrease uncertainty, it is possible that new information increase uncertainty if perceived as premonitory of future possible surprises (e.g., Rogers et al. 2009).

[25] The results are conceptually the same if I use quartile, but significance is affected due to lower power. Results are also quantitatively similar, albeit smaller in magnitude, if observations are sorted within year.

Differently from the main specification, I sort *Precision* according to quartiles instead of sextiles. This choice allows to reduce possible issues related to power or to a small set of observations driving the results. The results are similar, especially at the longer horizons, to ones obtained using sextiles, but minimum bin size is significantly larger: the smallest bin, low uncertainty-low precision-optimistic, contains around 500 observations.

Table 10a reports the results for different horizons, where the the interactions between the control variables and the *FirmU* dummy are added as controls. These results are largely consistent with the hypothesis, when they are pessimistic. Indeed, higher firm uncertainty is associated with a larger effect of precision for pessimistic reports of between 70 and 150 bps. Table 10b reports the results for similar empirical specifications where analysts' disagreement is used instead of idiosyncratic volatility. This measure has been used in literature as a measure of uncertainty (e.g., Johnson 2004) and, not surprisingly, it is correlated with volatility. The results, albeit weaker, are largely consistent with the ones obtained using the *FirmU* dummy.

7.2 Tone, Precision, and Sentence Position

My main analysis takes an agnostic view regarding where the positive or negative words are located as well as the position of sentence containing uncertainty words. Here I run empirical specifications similar to the main ones, but I change what portion of the report I use to calculate tone and precision. Indeed, consistently with the idea of limited investors' attention, it is possible to hypothesize that the effects of these two variables should be stronger if calculated based on specific parts of the report. On one hand, the beginning of a report usually consists of a summary of the analyst's opinion. A strong opinion, especially a pessimistic opinion[26], would make investors focus on this and place less importance on precision. On the other hand, the ending of longer reports are usually employed by analysts to discuss the forecasting models, risks and uncertainties concerning these, and general discussions about firm characteristics and trends. In other words, an investors with limited attention would focus on this part to gauge the level of precision.

I focus on the first 30 sentences of each report to calculate an alternative measure of *Tone* and the last 30 sentences to calculate an alternative measure of *Precision*. These correspond,

[26] A pessimistic opinion could be perceived as highly informative per se (Joos et al. 2016).

respectively, approximately to the first and last two pages of a report.[27] It is important to point out that all threshold for tone and precision classification are the same used in the main specifications.

Table 11 reports the results where tone from the first 30 sentences is used, while precision is the same used in the main specification. The results are qualitatively similar, but, consistently with the hypothesis, the precision effect for pessimistic reports is significantly weaker of around 15-20 bps. A strong opinion at the beginning of a report sways away attention of investors from precision. Alternatively, a strong negative tone in the beginning of the report is perceived as a signal that the analyst has some strong information and precision becomes somewhat less relevant. A possible issue is that, by using just the first 30 sentences, some pessimistic or optimistic reports are mislabeled as neutral. While part of the weaker results can be ascribed to this, unreported regressions on just pessimistic reports suggest that the effect of precision is indeed weaker, especially at shorter horizons.

Table 12 is the one that reports the results where precision from the last 30 sentences is used, while tone is the same used in the main specification. The results for the longer 4, 7 and, especially, two weeks horizons are stronger than in the main specification[28]. This result suggests that the last part of the reports is the one that better measure (perceived) precision. However, at shorter horizons, the relation between price reaction and precision is actually weaker than the one in the main specification. This result could stem from some inattentive investors ignoring the concluding parts of the reports, instead of focusing on them. These investors choice would, hence, lead to some overreaction to low precision reports.

8 Robustness

In this section, I present the results from robustness checks tests for the main price reaction results.

8.1 Alternative Observations Restrictions

In the main results, I follow existing literature and I exclude days around earnings announcements and around the issuance of management guidance. While these are the major groups of possibly problematic events, there could be other confounding news events. In this subsection, I define three

[27]Choosing an higher threshold (40-50), yield results closer to the main results.

[28]The interaction coefficient is only marginally statistically significant at this latter horizon for the main specification.

progressively stricter definitions of of possible confounders. First, I exclude any observation on days when more than 50% of the analysts covering a firm issued a forecast. For the second definition, I exclude days when more than five analysts issued a forecast, i.e., corresponding to 20-30% issued a forecast. For the last and most conservative forecast, I focused only on days where not more than three analysts issued a forecast, i.e., 10-20%[29]. In both these two restrictions, I also exclude days when more than one analyst changed a recommendation since these usually occur due to firm news (Loh and Stulz 2011).

The results for all the three restrictions are reported in Table 13. The first additional restriction has a relatively small effect on both the magnitude and significance of the main results. The effect of using the two more restrictive definition is stronger, but the results are still statistically significant at the 3, 5 and 7 days horizons. In particular, the effect of the restrictions is larger for shorter horizons compared to longer ones, suggesting that when analysts activity is smaller, investors take more time to process the information about report precision.

8.2 Longer Horizons and Overlapping Events

In the main results, I exclude reports issued less than 3 days before earnings or management guidance announcements. This choice does not prevent a partial overlap between these events and the longer horizons 5 and 7 days event windows. On one hand, this helps understanding whether informativeness and not just sentiment varies along the precision measure. Indeed, earnings or guidance announcements should tend to reverse any short term non-informational price reaction. On the other hand, it makes harder to properly separate the price reaction to the report publication from the price reaction to these major corporate news.

Table 14 reports the results for the longer 5 and 7 days horizons where only reports published at least 7 days before an earnings/guidance announcement are included. The results are nearly identical to the ones obtained with the full sample. It is also worth mentioning that this largely holds also for other empirical specifications for price reaction such as the ones presented in Section 7.

[29]The results are similar in magnitude when using a two analysts threshold, but more noisy

8.3 Crisis and Tone Thresholds

The construction of the *Tone* variable assumes that the threshold between pessimistic, neutral and optimistic reports is fixed across time. However, as aforementioned, it is reasonable to assume that the signal communicated by a neutral report could vary across time. Existing literature (e.g. Malmendier and Shantikumar 2007) has highlighted the fact that neutral analysts' opinions are considered as mildly negative signal by larger traders and positive ones as neutral, due to the well know optimistic analysts bias. This should be perhaps especially true during periods of crisis.

I run an alternative version of the main specification where I classify as pessimistic neutral reports and as neutral optimistic reports if issued during "bad times". I use NBER recession dates to identify "bad times" in my sample. Specifically, the period between December 2007 and June 2009 is classified as such.

The results are reported in Table 15 and they are actually stronger than the main results. High precision effect for pessimistic reports jumps to 75-99 bps depending on the horizon. Furthermore, even medium precision is associated with a generally significant albeit smaller effect of around 30-40 bps, or around half the effect observed for high precision.

8.4 Anomalous Measures

While the large majority of observations take a wide range of values for $ToneC$ and $PrecisionC$, some observations have extreme values. In particular, some reports have a value of precision equal to 1, suggesting that there was no sentence containing uncertain words. Similarly, some reports contains no positive or negative words. While these values could actually reflect the characteristics of the reports, they may corresponds to reports not actually discussing a firm. I estimate the same main price reaction specifications but excluding these observations. Results are reported in Table 16 and, despite the smaller sample, largely equivalent. This suggests that neither of these two groups of possibly anomalous measures is driving the results.

9 Conclusions

In this paper, I study the relation between the precision of the textual content of equity analysts' report and the informativeness of those reports. In particular, I study whether the magnitude of the price reaction to the publication of a report is increasing in report precision. I test whether investors put more “weight” on more precise reports and whether the price reaction is greater for highly precise optimistic or pessimistic reports. Because this relation could be negatively affected by analysts’ distortions and biases as well as by investors’ priors, I hypothesize that the relation between price reaction and precision is weaker when an analyst’s opinion and precision are more strongly distorted and when investors’ priors are more precise. Finally, I test whether due to investors’ limited attention, the placement of sentences within a report affects investors’ ability to correctly process the information and, as a consequence, the precision effect.

To measure precision at the analyst report level, I employ textual analysis to construct a novel, text-based measure of precision. I find that precision is related to different proxies for firm uncertainty such as idiosyncratic volatility, the availability of high quality information about a firm (e.g., days when earnings are announced) as well as to the ex-post accuracy of earnings forecasts. Precision is also positively correlated with the sentiment (tone) of the report. More optimistic reports also appear to be more precise. This finding suggests that analysts inflate the precision of more positive reports, possibly as a response to their well-known distorted incentives to produce flattering reports.

I find that high precision is associated with a larger and persistent price reaction for pessimistic reports, suggesting that more precise reports are also more informative. I also find that high precision is associated with higher abnormal turnover, higher realized volatility, and a smaller change in implied volatility. All these results support the hypothesis that highly precise pessimistic report are associated with increased market activity.

On the other hand, I find a weak and largely non-significant relation for the more biased optimistic reports. These results suggest that investors are at least partially able to see through the analysts’ distortions and consequently discount their opinions and their reports’ textual precision accordingly.

Concerning pessimistic reports, I also find that the relation between precision and price reaction is stronger for reports issued during periods of higher uncertainty. This result for firm uncertainty is consistent with the argument that investors put limited weight on analysts reports when their prior information is highly precise. Furthermore, this finding is also consistent with the idea that investors pay more attention to analysts during these periods and, hence, are better able to assess their precision.

Finally, I observe a relation between the effect of precision on price reaction and the part of an analyst report used to measure tone or precision; this relation is consistent with the idea that limited-attention affects the relation between precision and investors' reaction. Indeed, I find that the relation is weaker when tone is measured based on the first 30 sentences of a report, but it is stronger when precision is measured based on the last 30 sentences. I argue that these results are due to (1) investors focusing mainly on tone and less on precision when it is strongly presented in the first part of the report or (2) investors focusing on parts of the reports, such as the ending, that more often contain direct information about the risks related to the analysts' estimates.

This study provides evidence about the characteristics of analysts' precision as well as the relation between precision and financial markets, but it also opens several questions for future research. First, it would be interesting to have a better understanding of how precision is related to other variables and how it is affected by incentive, regulatory, or macroeconomic factors. This is particularly interesting because the industry is moving toward a business model where investors must purchase reports and, hence, the ability to value a report is key. In other work in progress, I provide further evidence about how precision and precision determinants changed after the implementation of the Global Research Settlement and related regulations (e.g., Kadan et al. 2009, Guan et al. 2017, Corwin et al. 2017).

It would also be interesting to employ machine learning and natural language processing to improve the understanding of analyst precision as well as to refine the textual measure of precision. For instance, in Appendix D, I use a Latent Dirilecht Allocation (LDA) algorithm to provide preliminary results concerning the relation between precision, market outcomes, and the topics of the reports. In other work in progress, I am working on a naive Bayes classifier to obtain a measure of precision that does not depend on a predefined dictionary. It would also be interesting to use similar classification algorithms to obtain measures capturing different types of (lack of) precision.

Finally, it would be interesting to employ the same methodology to study the output produced and disseminated by other information providers such as media or, especially, credit analysts. I leave this to future work.

References

[Altinkilic et al., 2013] Altinkilic, O., Balashov, V. S., & Hansen, R. S. (2013). Are Analysts' Forecasts Informative to the General Public? *Management Science*, 59(11), 2550–2565.

[Altinkilic & Hansen, 2009] Altinkilic, O. & Hansen, R. S. (2009). On the Information Role of Stock Recommendation Revisions. *Journal of Accounting and Economics*, 48, 17–36.

[Asquith et al., 2005] Asquith, P., Mikhail, M. B., & Au, A. S. (2005). Information Content of Equity Analyst Reports. *Journal of Financial Economics*, 75(2), 245–282.

[Baker et al., 2016] Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring Economic Policy Uncertainty. *The Quarterly Journal of Economics*, 131(4), 1593–1636.

[Barber et al., 2007] Barber, B. M., Lehavy, R., & Trueman, B. (2007). Comparing the Stock Recommendation Performance of Investment Banks and Independent Research Firms. *Journal of Financial Economics*, 85(2), 490–517.

[Barberis, 2018] Barberis, N. (2018). Psychology-based Models of Asset Prices and Trading Volume.

[Bellstam, 2017] Bellstam, G. (2017). Incentives, Information Production, and Bias: Evidence from Analyst Text.

[Billings et al., 2015] Billings, M. B., Jennings, R., & Lev, B. (2015). On Guidance and Volatility. *Journal of Accounting and Economics*, 60(2-3), 161–180.

[Blei et al., 2001] Blei, D. M., Ng, A. Y., & Jordan, M. I. (2001). Latent Dirichlet Allocation. *The Journal of Machine Learning Research*, 3, 993–1022.

[Bradley et al., 2014] Bradley, D., Clarke, J., Lee, S., & Ornthanalai, C. (2014). Are Analysts' Recommendations Informative? Intraday Evidence on the Impact of Time Stamp Delays. *Journal of Finance*, 69(2), 645–673.

[Bradshaw, 2011] Bradshaw, M. T. (2011). Analysts' Forecasts: What Do We Know after Decades of Work?

[Brown & Mohd, 2003] Brown, L. D. & Mohd, E. (2003). The Predictive Value of Analyst Characteristics. *Journal of Accounting, Auditing & Finance*, 18(4), 625–647.

[Campbell et al., 1993] Campbell, J. Y., Grossman, S. J., & Wang, J. (1993). Trading Volume and Serial Correlation in Stock Returns. *The Quarterly Journal of Economics*, 108(4), 905–939.

[Carhart, 1997] Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *The Journal of Finance*, 52(1), 57–82.

[Clement, 1999] Clement, M. B. (1999). Analyst Forecast Accuracy: Do Ability, Resources, and Portfolio Complexity Matter? *Journal of Accounting and Economics*, 27(3), 285–303.

[Corwin et al., 2017] Corwin, S. A., Larocque, S. A., & Stegemoller, M. A. (2017). Investment Banking Relationships and Analyst Affiliation Bias: The Impact of the Global Settlement on Sanctioned and Non-sanctioned Banks. *Journal of Financial Economics*, 124(3), 614–631.

[DeLong et al., 1990] DeLong, J. B. D., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise Trader Risk in Financial Markets. *Journal of Political Economy*, 98, 703–738.

[Demers & Vega, 2014] Demers, E. A. & Vega, C. (2014). The Impact of Credibility on the Pricing of Managerial Textual Content.

[Derrien & Kecskec, 2013] Derrien, F. & Kecskec, A. (2013). The Real Effects of Financial Shocks: Evidence from Exogenous Changes in Analyst Coverage. *The Journal of Finance*, 68(4), 1407–1440.

[Diether et al., 2002] Diether, K. B., Malloy, C. J., & Scherbina, A. (2002). Differences of Opinion and the Cross Section of Stock Returns. *The Journal of Finance*, 57(5), 2113–2141.

[Duffie, 2010] Duffie, D. (2010). Asset Price Dynamics with Slow-Moving Capital. *The Journal of Finance*, 65(4), 1237–1267.

[Fedyk, 2018] Fedyk, A. (2018). Front Page News: The Effect of News Positioning on Financial Markets.

[Gleason & Lee, 2003] Gleason, C. A. & Lee, C. M. C. (2003). Analyst Forecast Revisions and Market Price Discovery. *The Accounting Review*, 78(1), 193–225.

[Guan et al., 2017] Guan, Y., Lu, H., & Wong, M. H. F. (2017). Regulations and Brain Drain: Evidence from the Wall Street Star Analysts' Career Choices.

[Heston & Sinha, 2017] Heston, S. L. & Sinha, N. R. (2017). News vs. Sentiment: Predicting Stock Returns from News Stories. *Financial Analysts Journal*, 73(3), 67–83.

[Hong & Kubik, 2003] Hong, H. & Kubik, J. D. (2003). Analyzing the Analysts: Career Concerns and Biased Earnings Forecasts.

[Huang et al., 2017] Huang, A. H., Lehavy, R., Zang, A. Y., & Zheng, R. (2017). Analyst Information Discovery and Interpretation Roles: A Topic Modeling Approach. *Management Science*, (pp. 2473–2972).

[Huang et al., 2014] Huang, A. H., Zang, A. Y., & Zheng, R. (2014). Evidence on the Information Content of Text in Analyst Reports. *The Accounting Review*, 89(6), 2151–2180.

[Jackson, 2005] Jackson, A. R. (2005). Trade Generation, Reputation, and Sell-side Analysts. *Journal of Finance*, 60(2), 673–717.

[Jegadeesh & Kim, 2010] Jegadeesh, N. & Kim, W. (2010). Do Analysts Herd? An Analysis of Recommendations and Market Reactions. *Review of Financial Studies*, 23(2), 901–937.

[Johnson, 2004] Johnson, T. C. (2004). Forecast Dispersion and the Cross Section of Expected Returns. *Journal of Finance*, 59(5), 1957–1978.

[Joos et al., 2016] Joos, P., Piotroski, J. D., & Srinivasan, S. (2016). Can Analysts Assess Fundamental Risk and Valuation Uncertainty? An Empirical Analysis of Scenario-based Value Estimates. *Journal of Financial Economics*, 121(3), 645–663.

[Joos & Piotroski, 2017] Joos, P. R. & Piotroski, J. D. (2017). The Best of All Possible Worlds: Unraveling Target Price Optimism Using Analysts' Scenario-based Valuations. *Review of Accounting Studies*, 22(4), 1492–1540.

[Kadan et al., 2009] Kadan, O., Madureira, L., Wang, R., & Zach, T. (2009). Conflicts of Interest and Stock Recommendations: The Effects of the Global Settlement and Related Regulations. *Review of Financial Studies*, 22(10), 4189–4217.

[Kandel & Pearson, 1995] Kandel, E. & Pearson, N. D. (1995). Differential Interpretation of Public Signals and Trade in Speculative Markets. *Journal of Political Economy*, 103(4), 831–872.

[Kelly & Ljungqvist, 2012] Kelly, B. & Ljungqvist, A. (2012). Testing Asymmetric-Information Asset Pricing Models. *Review of Financial Studies*, 25(5), 1366–1413.

[Kim & Verrecchia, 1991] Kim, O. & Verrecchia, R. E. (1991). Trading Volume and Price Reactions to Public Announcements. *Journal of Accounting Research*, 29(2), 302–321.

[Kim & Verrecchia, 1994] Kim, O. & Verrecchia, R. E. (1994). Market Liquidity and Volume Around Earnings Announcements. *Journal of Accounting and Economics*, 17(1-2), 41–67.

[Kim & Verrecchia, 1997] Kim, O. & Verrecchia, R. E. (1997). Pre-announcement and Event-period Private Information. *Journal of Accounting and Economics*, 24(3), 395–419.

[Li et al., 2015] Li, E. X., Ramesh, K., Shen, M., & Wu, J. S. (2015). Do Analyst Stock Recommendations Piggyback on Recent Corporate News? An Analysis of Regular-Hour and After-Hours Revisions. *Journal of Accounting Research*, 53(4), 821–861.

[Lin & McNichols, 1998] Lin, H.-w. & McNichols, M. F. (1998). Underwriting Relationships, Analysts' Earnings Forecasts and Investment Recommendations. *Journal of Accounting and Economics*, 25(1), 101–127.

[Loh & Stulz, 2017] Loh, R. & Stulz, R. M. (2017). Is Sell-Side Research More Valuable in Bad Times? *Journal of Finance*, (Forthcoming).

[Loh & Stulz, 2011] Loh, R. K. & Stulz, R. M. (2011). When Are Analyst Recommendation Changes Influential? *Review of Financial Studies*, 24(2), 593–627.

[Loughran & McDonald, 2011] Loughran, T. & McDonald, B. (2011). When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *The Journal of Finance*, 66(1), 35–65.

[Loughran & McDonald, 2013] Loughran, T. & McDonald, B. (2013). IPO First-day Returns, Offer Price Revisions, Volatility, and Form S-1 Language. *Journal of Financial Economics*, 109(2), 307–326.

[Ludvigson et al., 2015] Ludvigson, S., Ma, S., & Ng, S. (2015). Uncertainty and Business Cycles: Exogenous Impulse or Endogenous Response?

[Malmendier & Shanthikumar, 2007] Malmendier, U. & Shanthikumar, D. (2007). Are Small Investors Naive About Incentives? *Journal of Financial Economics*, 85(2), 457–489.

[McNichols & O'Brien, 1997] McNichols, M. & O'Brien, P. C. (1997). Self-Selection and Analyst Coverage. *Journal of Accounting Research*, 35, 167.

[Michaely & Womack, 1999] Michaely, R. & Womack, K. L. (1999). Conflict of Interest and the Credibility of Underwriter Analyst Recommendations. *Review of Financial Studies*, 12(4), 653–686.

[Moore & Healy, 2008] Moore, D. A. & Healy, P. J. (2008). The Trouble with Overconfidence. *Psychological Review*, 115(2), 502–517.

[Ouimet & Tate, 2017] Ouimet, P. P. & Tate, G. A. (2017). Investor Experience and Attention: The Effect of Financial Shocks on Individual Trading Decisions.

[Řehůřek & Sojka, 2010] Řehůřek, R. & Sojka, P. (2010). Software Framework for Topic Modelling with Large Corpora.

[Rogers et al., 2009] Rogers, J. L., Skinner, D. J., & Van Buskirk, A. (2009). Earnings Guidance and Market Uncertainty. *Journal of Accounting and Economics*, 48(1), 90–109.

[Sievert & Shirley, 2014] Sievert, C. & Shirley, K. E. (2014). *LDAvis: A Method for Visualizing and Interpreting Topics*. Technical report.

[Stickel, 1992] Stickel, S. E. (1992). Reputation and Performance Among Security Analysts. *The Journal of Finance*, 47(5), 1811–1836.

[Subramanyam, 1996] Subramanyam, K. R. (1996). Uncertain Precision and Price Reactions to Information. *The Accounting Review*, 71, 207–219.

[Tetlock, 2007] Tetlock, P. C. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *The Journal of Finance*, 62(3), 1139–1168.

[Tetlock, 2011] Tetlock, P. C. (2011). All the News That's Fit to Reprint: Do Investors React to Stale Information? *Review of Financial Studies*, 24(5), 1481–1512.

[Umar, 2017] Umar, T. (2017). Complexity Aversion when Seeking Alpha.

[Womack, 1996] Womack, K. L. (1996). Do Brokerage Analysts' Recommendations Have Investment Value? *Journal of Finance*, 51(1), 137–167.

[Xiao & Zang, 2017] Xiao, K. & Zang, A. (2017). Hardening Soft Information: Analyst Conservative Bias.

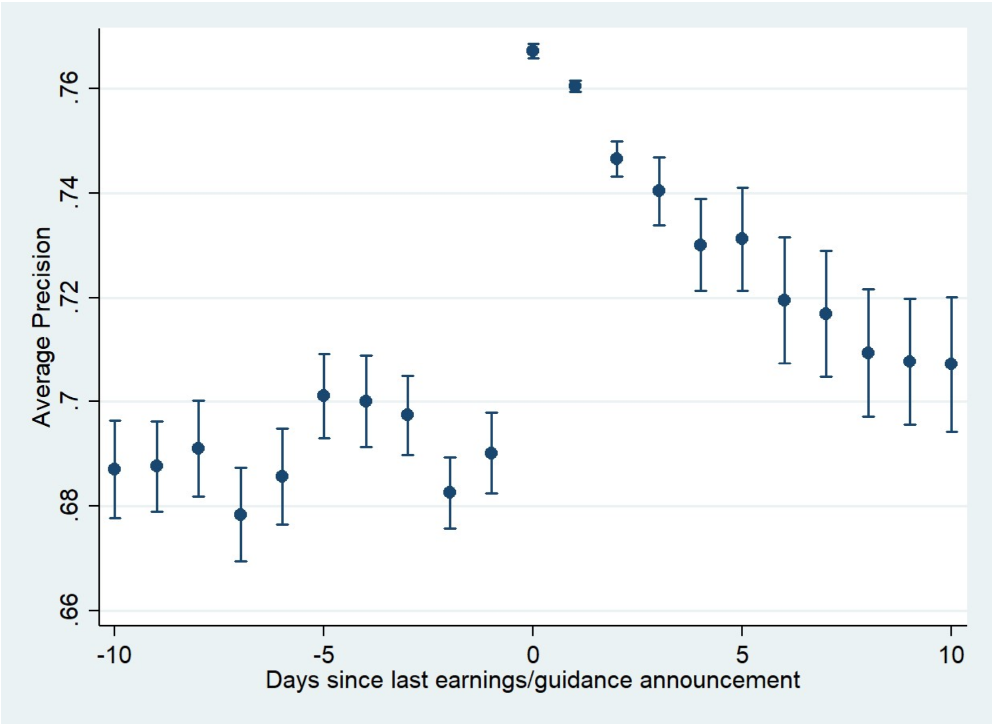
[Zerva et al., 2017] Zerva, C., Batista-Navarro, R., Day, P., & Ananiadou, S. (2017). Using Uncertainty to Link and Rank Evidence from Biomedical Literature for Model Curation. *Bioinformatics*, 33(23), 3784–3792.

[Zhang, 2006] Zhang, X. F. (2006). Information Uncertainty and Stock Returns. *Journal of Finance*, 61(1), 105–137.

[Zhou, 2018] Zhou, D. (2018). Good News in Numbers.

Figure 1: Precision around Earnings and Guidance Announcements

These figures show the evolution of precision around earnings and management guidance announcements. The first figure shows the average precisions. The second figure shows the percentage of *High* and *Low* precision reports. *High* (*Low)* is equal to one if precision is in the top (bottom) sextile. Bars represent 95% confidence intervals.



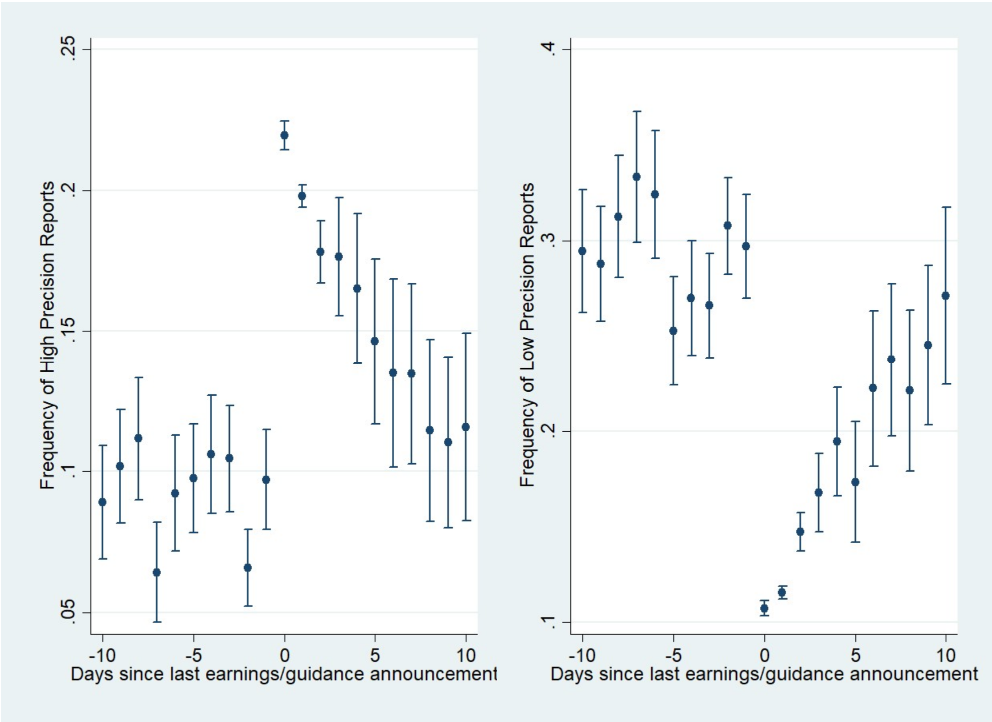


Figure 2: Time Series of Precision

This figure shows the evolution of average precision across time. Precision measure is de-trended. The period corresponding to the height of the 2008-2009 financial crisis is highlighted.

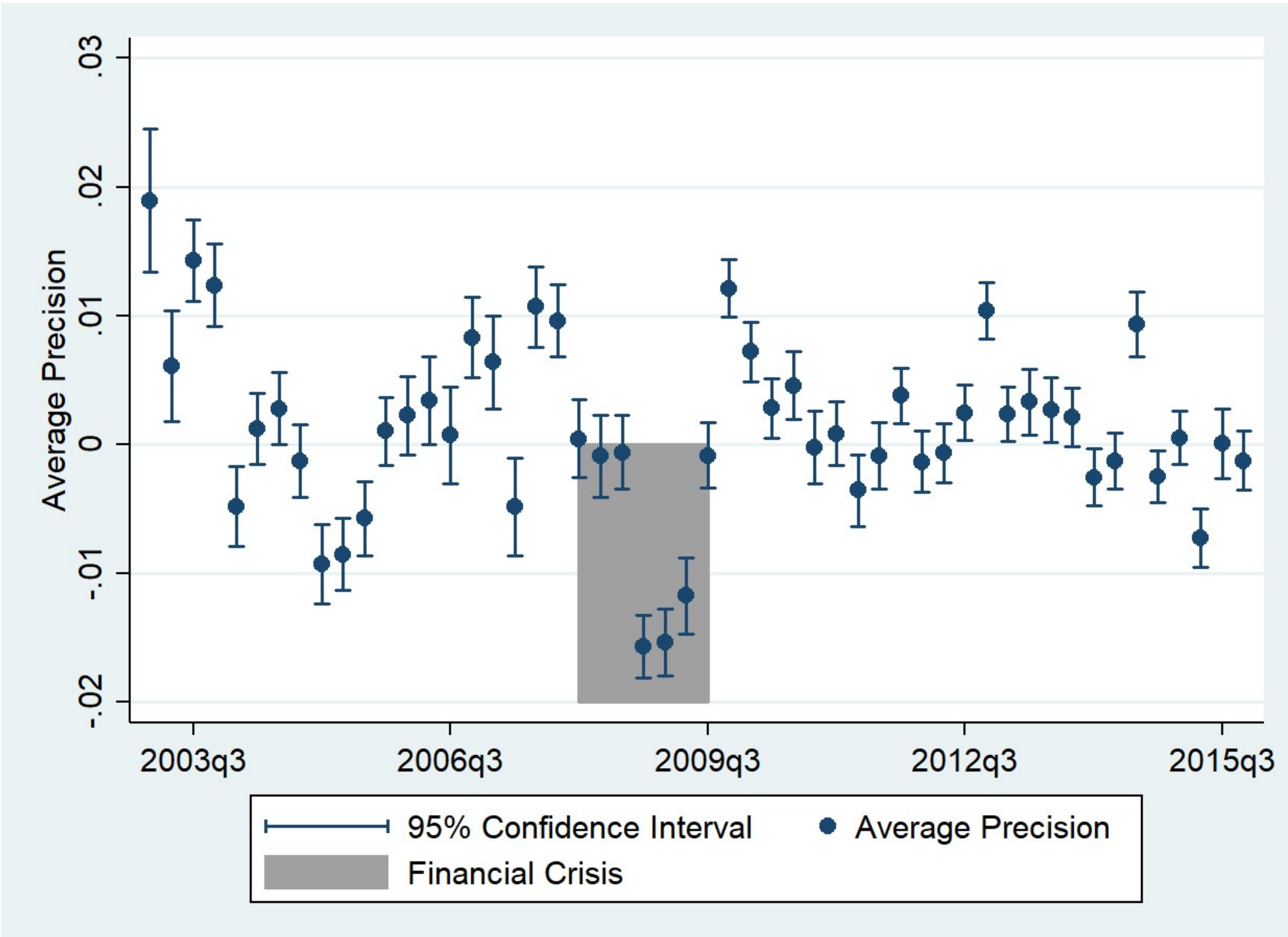


Table 1: Summary Statistics

This table reports some summary statistics about the two textual measures as well as several of the covariates used in the analyses. Definitions of the variables are presented in Appendix A.

PANEL A

	Mean	StDev	25th	Median	75th
ToneC	0.003	0.021	-0.009	0.003	0.016
PrecisionC	0.738	0.120	0.667	0.750	0.824
Observations	98,914				

PANEL B

	Mean	StDev	25th	Median	75th
Size	9.26	1.25	8.36	9.18	10.15
Book-to-Market	0.41	0.26	0.23	0.35	0.54
Policy Uncertainty	119.84	45.75	87.42	108.51	146.12
Length	3.85	0.88	3.18	3.76	4.42
Broker Size	3.95	0.80	3.33	4.14	4.58
Number of Analysts	2.94	0.44	2.71	3.00	3.26
Absolute Experience	3.75	0.91	3.26	3.97	4.48
Relative Experience	2.52	1.26	1.79	2.77	3.47
Deviation from Consensus	-0.01	0.20	-0.02	0.00	0.02
ABS(Deviation from Consensus)	0.09	0.21	0.01	0.02	0.06
Observations	98,914				

Table 2: Precision Fixed Effects

This table reports the percentage of variance in *Precision* explained by different fixed effects. Industry is Fama-French 12 industries.

	(1)	(2)	(3)	(4)	(5)
VARIABLES	PrecisionC	PrecisionC	PrecisionC	PrecisionC	PrecisionC
Observations	98,914	98,914	98,832	98,914	98,038
Adjusted R-squared	0.019	0.026	0.320	0.062	0.383
Month FE	YES	YES	YES	YES	YES
Industry FE	NO	YES	NO	NO	NO
Analyst FE	NO	NO	YES	NO	NO
Firm FE	NO	NO	NO	YES	NO
Analyst-Firm FE	NO	NO	NO	NO	YES

Table 3: Precision Determinants

This table reports the results of different regressions of *Precision* (in percentage points) on a set of different firm and analysts characteristics. *Earnings* is a dummy equal to one if the report is published within two days from an earnings or management guidance announcement. *Staleness* is the maximum cosine similarity between a report and any report published about the same firm in the previous 90 days. Standard errors are double clustered at analyst and firm level. t-stats in parenthesis. ***, ** and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	PrecisionC	PrecisionC	PrecisionC	PrecisionC	PrecisionC	PrecisionC
Size	-0.067	0.141	-0.117			
	(-0.38)	(1.26)	(-0.24)			
Book-to-Market	-0.154	-0.205**	-0.094			
	(-1.00)	(-1.98)	(-0.73)			
Id Volatility	-0.968***	-0.787***	-0.336***			
	(-7.31)	(-8.55)	(-3.83)			
Number of Analysts	0.404**	-0.177	-0.061			
	(2.48)	(-1.65)	(-0.43)			
Absolute Experience				0.674***	0.527***	-1.104
				(3.14)	(2.65)	(-0.56)
Relative Experience				0.238	-0.001	1.750
				(1.20)	(-0.01)	(1.42)
Broker Size				-0.892***	-0.777***	-0.350
				(-4.53)	(-3.79)	(-1.18)
Observations	98,914	98,832	98,038	98,459	98,459	97,620
Adjusted R-squared	0.008	0.359	0.419	0.010	0.106	0.413
Earnings	NO	YES	YES	NO	YES	YES
Month FE	NO	YES	YES	NO	YES	NO
Firm FE	NO	YES	NO	NO	NO	NO
Analyst FE	NO	NO	NO	NO	YES	NO
Analyst-Firm FE	NO	NO	YES	NO	NO	YES

Table 3: Precision Determinants – Cont.

VARIABLES	(7) PrecisionC	(8) PrecisionC	(9) PrecisionC	(10) PrecisionC	(11) PrecisionC	(12) PrecisionC
Abs(Deviation from Consensus)	0.163	0.234***	0.178***			
	(1.37)	(3.58)	(3.30)			
Pessimistic	-1.944***	-1.366***	-1.228***			
	(-9.45)	(-7.57)	(-10.67)			
Optimistic	1.909***	1.618***	1.453***			
	(8.72)	(9.37)	(13.86)			
Staleness				-1.840***	-1.707***	-1.198***
				(-13.01)	(-12.80)	(-14.43)
Log(# of numbers)				2.344***	1.791***	1.689***
				(14.05)	(11.78)	(16.68)
Observations	98,914	98,914	98,038	94,699	94,699	93,837
Adjusted R-squared	0.013	0.108	0.424	0.038	0.123	0.431
Earnings	NO	YES	YES	NO	YES	YES
Month FE	NO	YES	YES	NO	YES	YES
Analyst FE	NO	YES	NO	NO	YES	NO
Analyst-Firm FE	NO	NO	YES	NO	NO	YES

Table 4: High and Low Precision Determinants

This table reports the results of different regressions of *High* and *Low Precision* dummy variables on a set of different firm and analysts characteristics. *High* (*Low)* is equal to one if *PrecisionC* is in the top (bottom) sextile. *Earnings* is a dummy equal to one if the report is published within two days from an earnings or management guidance announcement. *Staleness* is the maximum cosine similarity between a report and any report published about the same firm in the previous 90 days. Standard errors are double clustered at analyst and firm level. t-stats in parenthesis. ***, ** and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

VARIABLES	(3) High Precision	(4) High Precision	(5) Low Precision	(6) Low Precision
Size	0.009**	0.002	0.003	0.002
	(2.07)	(0.12)	(0.62)	(0.17)
Book-to-Market	-0.005	-0.005	-0.003	-0.005
	(-1.43)	(-1.45)	(-0.78)	(-1.38)
Id Volatility	-0.013***	-0.004*	0.015***	0.008***
	(-4.06)	(-1.82)	(4.10)	(2.84)
Number of Analysts	-0.003	0.002	-0.007*	-0.001
	(-0.75)	(0.55)	(-1.80)	(-0.32)
Absolute Experience	0.012**	-0.050	-0.009*	-0.113
	(2.36)	(-0.34)	(-1.87)	(-0.88)
Relative Experience	0.003	0.055	-0.002	-0.025
	(0.61)	(1.45)	(-0.38)	(-0.59)
Broker Size	-0.024***	-0.009	0.018***	0.010
	(-4.81)	(-1.22)	(3.49)	(1.23)
Abs(Deviation from Consensus)	0.007***	0.001	-0.006***	-0.005***
	(3.34)	(0.55)	(-2.91)	(-3.05)
Staleness	-0.049***	-0.039***	0.025***	0.018***
	(-13.21)	(-15.45)	(8.13)	(8.00)
Log(# of numbers)	0.017***	0.008***	-0.058***	-0.057***
	(3.82)	(2.66)	(-15.51)	(-19.14)
Pessimistic	-0.018***	-0.013***	0.035***	0.026***
	(-3.92)	(-4.03)	(6.24)	(6.74)
Optimistic	0.046***	0.038***	-0.032***	-0.028***
	(8.40)	(10.39)	(-6.51)	(-8.71)
Observations	94,265	93,440	94,265	93,440
Adjusted R-squared	0.045	0.248	0.061	0.258
Earnings	YES	YES	YES	YES
Month FE	YES	YES	YES	YES
Analyst-Firm FE	NO	YES	NO	YES

Table 5: Precision and Ex-Post Accuracy

This table reports the results of regressions of different measures of earnings forecast accuracy on *Precision.* Panel A includes the standardized value of *Precision*, while Panel B contains two dummies for *High (Low) Precision.* Controls include size, B/M, number of analysts, idiosyncratic volatility, absolute and relative analyst experience, deviation from consensus, broker size, and distance from earnings announcement. Standard errors are double clustered at analyst and firm level. t-stats in parenthesis. ***, ** and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A	(1)	(2)	(3)	(4)
VARIABLES	Abs Error	Abs Error (Scaled)	Proportional Abs Error (Scaled)	Relative Abs Error
PrecisionC (standardized)	-0.012***	-0.048***	-0.029***	-0.014***
	(-2.72)	(-2.60)	(-6.44)	(-3.84)
Observations	97,539	97,539	97,539	97,539
Adjusted R-squared	0.352	0.273	0.237	0.081
Year FE	YES	YES	YES	YES
Analyst-Firm FE	YES	YES	YES	YES
Controls	YES	YES	YES	YES

Panel B	(1)	(2)	(3)
VARIABLES	Abs Error (Scaled)	Proportional Abs Error (Scaled)	Relative Abs Error
High Precision	-0.090***	-0.051***	-0.045***
	(-3.43)	(-4.20)	(-2.99)
Low Precision	0.029	0.021*	0.025
	(1.06)	(1.76)	(1.56)
Neutral#High Precision	0.063**	0.038***	0.029
	(2.09)	(2.65)	(1.65)
Optimistic#High Precision	0.077**	0.036**	0.035*
	(2.32)	(2.25)	(1.90)
Neutral#Low Precision	0.008	0.004	0.006
	(0.32)	(0.30)	(0.35)
Optimistic#Low Precision	0.010	0.014	0.012
	(0.33)	(0.86)	(0.59)
Observations	97,539	97,539	97,539
Adjusted R-squared	0.524	0.294	0.075
Year FE	YES	YES	YES
Analyst-Firm FE	YES	YES	YES
Controls	YES	YES	YES

Table 6: Tone, Bias, and Ex-Post Accuracy

This table reports the results of regressions of different measures of bias and earning forecasts accuracy on *Tone.* Controls include size, B/M, number of analysts, idiosyncratic volatility, absolute and relative analyst experience, deviation from consensus, broker size, and distance from earnings announcement. Standard errors are double clustered at analyst and firm level. t-stats in parenthesis. ***, ** and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Error	Error (Scaled)	Abs Error	Proportional Abs Error (Scaled)	Proportional Abs Error (Scaled)	Relative Abs Error
Neutral	0.039***	0.137***	-0.016***	-0.086***	-0.026***	0.007
	(4.89)	(4.50)	(-3.68)	(-6.23)	(-3.48)	(0.81)
Optimistic	0.080***	0.230***	-0.025***	-0.127***	-0.061***	-0.016
	(7.10)	(5.49)	(-3.91)	(-6.89)	(-6.15)	(-1.58)
Observations	97,539	97,539	97,539	97,539	97,539	97,539
Adjusted R-squared	0.202	0.192	0.478	0.544	0.294	0.075
Year FE	YES	YES	YES	YES	YES	YES
Analyst-Firm FE	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES

Table 7: Price reaction to analyst report precision

This table reports the relation between price reaction (Cumulative Abnormal Returns multiplied by 100) to the publication of an analyst report and its precision. Reports are sorted in three *Tone* categories and monthly sorted in three *Precision* categories based on sextiles. Controls include size, B/M, number of analysts, uncertainty index, prior-CAR, absolute and relative analyst experience, deviation from consensus, broker size, and report length. Rev Controls include change in recommendation and change in EPS forecast. Standard errors are double clustered at analyst and industry-week level. t-stats in parenthesis. ***, ** and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	CAR[0,1]	CAR[0,1]	CAR[0,1]	CAR[0,2]	CAR[0,2]	CAR[0,2]
Tone = 1, Pessimistic	-0.442***	-0.250	-0.570***	-0.268	-0.081	-0.484***
	(-2.77)	(-1.59)	(-4.75)	(-1.47)	(-0.45)	(-3.53)
Tone = 2, Optimistic	0.417**	0.290	0.345**	0.363*	0.229	0.290*
	(2.13)	(1.57)	(2.37)	(1.74)	(1.13)	(1.93)
Precision = 1, Medium	-0.165	-0.147	-0.094	-0.099	-0.081	-0.027
	(-1.42)	(-1.29)	(-1.15)	(-0.77)	(-0.64)	(-0.29)
Precision = 2, High	-0.219	-0.218	-0.141	-0.127	-0.125	-0.114
	(-1.52)	(-1.53)	(-1.38)	(-0.80)	(-0.80)	(-1.05)
Pessimistic#Medium	-0.176	-0.224	-0.125	-0.266	-0.309	-0.151
	(-0.99)	(-1.28)	(-0.88)	(-1.31)	(-1.53)	(-0.92)
Pessimistic#High	-0.395	-0.466*	-0.284	-0.609**	-0.677***	-0.387*
	(-1.63)	(-1.96)	(-1.34)	(-2.30)	(-2.59)	(-1.71)
Optimistic#Medium	0.028	0.043	-0.079	-0.029	-0.014	-0.080
	(0.13)	(0.21)	(-0.49)	(-0.13)	(-0.07)	(-0.48)
Optimistic#High	-0.070	-0.062	-0.104	-0.050	-0.049	0.018
	(-0.32)	(-0.29)	(-0.55)	(-0.21)	(-0.21)	(0.09)
Observations	21,361	21,361	23,530	21,361	21,361	23,530
R-squared	0.216	0.236	0.013	0.205	0.222	0.009
Controls	YES	YES	NO	YES	YES	NO
Rev Controls	NO	YES	NO	NO	YES	NO
Analyst-Firm FE	YES	YES	NO	YES	YES	NO

Table 7b: Price reaction to analyst report precision – Longer horizons

This table reports the relation between price reaction (Cumulative Abnormal Returns, multiplied by 100) to the publication of an analyst report and its precision. Reports are sorted in three *Tone* categories and monthly sorted in three *Precision* categories based on sextiles. Controls include size, B/M, number of analysts, uncertainty index, prior-CAR, absolute and relative analyst experience, deviation from consensus, broker size, and report length. Rev Controls include change in recommendation and change in EPS forecast. Standard errors are double clustered at analyst and industry-week level. t-stats in parenthesis. ***, ** and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	CAR[0,4]	CAR[0,4]	CAR[0,4]	CAR[0,6]	CAR[0,6]	CAR[0,6]
Tone = 1, Pessimistic	-0.207	-0.042	-0.511***	0.144	0.303	-0.219
	(-0.96)	(-0.19)	(-2.90)	(0.58)	(1.22)	(-1.12)
Tone = 2, Optimistic	0.253	0.115	0.213	0.224	0.118	0.308
	(0.96)	(0.45)	(1.11)	(0.75)	(0.41)	(1.45)
Precision = 1, Medium	-0.140	-0.125	-0.018	-0.066	-0.055	-0.018
	(-0.99)	(-0.89)	(-0.13)	(-0.38)	(-0.32)	(-0.13)
Precision = 2, High	-0.146	-0.140	-0.157	-0.036	-0.030	0.042
	(-0.78)	(-0.76)	(-1.19)	(-0.16)	(-0.14)	(0.27)
Pessimistic#Medium	-0.152	-0.175	0.014	-0.376	-0.393	-0.193
	(-0.64)	(-0.74)	(0.07)	(-1.35)	(-1.41)	(-0.84)
Pessimistic#High	-0.627**	-0.697**	-0.305	-0.882**	-0.959***	-0.577*
	(-1.98)	(-2.23)	(-1.11)	(-2.48)	(-2.70)	(-1.86)
Optimistic#Medium	0.122	0.130	0.014	-0.023	-0.019	-0.220
	(0.44)	(0.48)	(0.07)	(-0.08)	(-0.06)	(-0.97)
Optimistic#High	0.055	0.049	-0.071	-0.110	-0.097	-0.372
	(0.18)	(0.16)	(-0.28)	(-0.32)	(-0.28)	(-1.38)
Observations	21,361	21,361	23,530	21,361	21,361	23,530
R-squared	0.201	0.212	0.004	0.199	0.206	0.002
Controls	YES	YES	NO	YES	YES	NO
Rev Controls	NO	YES	NO	NO	YES	NO
Analyst-Firm FE	YES	YES	NO	YES	YES	NO

Table 7c: Price reaction to analyst report precision – Precision quartiles

This table reports the relation between price reaction (Cumulative Abnormal Returns, multiplied by 100) to the publication of an analyst report and its precision. Reports are sorted in three *Tone* categories and monthly sorted in three *Precision* categories based on quartiles. Controls include size, B/M, number of analysts, uncertainty index, prior-CAR, absolute and relative analyst experience, deviation from consensus, broker size, and report length. Rev Controls include change in recommendation and change in EPS forecast. Standard errors are double clustered at analyst and industry-week level. t-stats in parenthesis. ***, ** and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
VARIABLES	CAR[0,1]	CAR[0,2]	CAR[0,4]	CAR[0,6]
Tone = 1, Pessimistic	-0.338**	-0.191	-0.184	0.097
	(-2.45)	(-1.24)	(-1.03)	(0.46)
Tone = 2, Optimistic	0.274*	0.173	0.121	0.102
	(1.84)	(1.06)	(0.61)	(0.42)
Precision = 1, Medium	-0.087	-0.049	-0.140	-0.062
	(-0.83)	(-0.43)	(-1.09)	(-0.42)
Precision = 2, High	-0.084	-0.009	-0.106	0.006
	(-0.73)	(-0.07)	(-0.70)	(0.04)
Pessimistic#Medium	-0.074	-0.121	0.111	0.008
	(-0.44)	(-0.64)	(0.51)	(0.03)
Pessimistic#High	-0.393**	-0.576***	-0.613**	-0.935***
	(-2.02)	(-2.66)	(-2.35)	(-3.05)
Optimistic#Medium	0.132	0.117	0.237	0.118
	(0.83)	(0.68)	(1.08)	(0.46)
Optimistic#High	-0.118	-0.084	-0.096	-0.251
	(-0.65)	(-0.44)	(-0.39)	(-0.85)
Observations	21,361	21,361	21,361	21,361
R-squared	0.235	0.221	0.213	0.207
Controls	YES	YES	YES	YES
Rev Controls	YES	YES	YES	YES
Analyst-Firm FE	YES	YES	YES	YES

Table 8: Turnover and Precision

This table reports the relation between turnover (cumulative abnormal log turnover) response to the publication of an analyst report and its precision. Abnormal log turnover is measured as the difference between log turnover and the average log turnover in the previous five trading days. No overlap for longer horizons exclude observations whose 7 days event window overlap with earnings/guidance announcement. Reports are sorted in three *Tone* categories and monthly sorted in three *Precision* categories based on sextiles. Controls include size, B/M, number of analysts, uncertainty index, prior-CAR, prior-CAT, absolute and relative analyst experience, deviation from consensus, broker size, and report length. Rev Controls include change in recommendation and change in EPS forecast. Standard errors are double clustered at analyst and industry-week level. t-stats in parenthesis. ***, ** and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	CAT[0,1]	CAT[0,2]	CAT[0,4]	CAT[0,4]	CAT[0,6]	CAT[0,6]
Tone = 1, Pessimistic	-0.059*	-0.092**	-0.132*	-0.177**	-0.130	-0.198*
	(-1.77)	(-2.07)	(-1.84)	(-2.23)	(-1.33)	(-1.85)
Tone = 2, Optimistic	-0.065	-0.048	-0.036	-0.002	-0.011	0.020
	(-1.60)	(-0.86)	(-0.41)	(-0.02)	(-0.09)	(0.16)
Precision = 1, Medium	0.021	0.037	0.074	0.031	0.077	0.008
	(0.72)	(0.94)	(1.20)	(0.45)	(0.93)	(0.09)
Precision = 2, High	-0.022	-0.026	-0.030	-0.054	-0.067	-0.124
	(-0.60)	(-0.52)	(-0.40)	(-0.65)	(-0.64)	(-1.10)
Pessimistic#Medium	0.071*	0.109**	0.146*	0.198**	0.147	0.221*
	(1.93)	(2.20)	(1.87)	(2.30)	(1.38)	(1.89)
Pessimistic#High	0.151***	0.232***	0.344***	0.376***	0.389***	0.464***
	(3.14)	(3.59)	(3.44)	(3.46)	(2.87)	(3.15)
Optimistic#Medium	0.031	0.036	0.029	0.004	0.031	0.007
	(0.75)	(0.61)	(0.32)	(0.04)	(0.25)	(0.05)
Optimistic#High	0.134***	0.157**	0.198*	0.178	0.213	0.231
	(2.75)	(2.29)	(1.87)	(1.54)	(1.43)	(1.44)
Observations	21,255	21,255	21,255	18,519	21,255	18,519
R-squared	0.263	0.259	0.268	0.281	0.284	0.293
No Overlap for Longer Horizons			NO	YES	NO	YES
Controls	YES	YES	YES	YES	YES	YES
Rev Controls	NO	YES	NO	YES	NO	YES
Analyst-Firm FE	YES	YES	YES	YES	YES	YES

Table 9: Volatility and Precision

This table reports the relation between realized volatility (cumulative squared abnormal returns, multiplied by 100) response to the publication of an analyst report and its precision. Reports are sorted in three *Tone* categories and monthly sorted in three *Precision* categories based on sextiles. No overlap for longer horizons exclude observations whose 7 days event window overlap with earnings/guidance announcement. Controls include size, B/M, number of analysts, uncertainty index, prior-CAR, absolute and relative analyst experience, deviation from consensus, broker size, report length, and idiosyncratic volatility. Rev Controls include change in recommendation and change in EPS forecast. Standard errors are double clustered at analyst and industry-week level. t-stats in parenthesis. ***, ** and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	CSAR[0,1]	CSAR[0,2]	CSAR[0,4]	CSAR[0,4]	CSAR[0,6]	CSAR[0,6]
Tone = 1, Pessimistic	-0.024	-0.018	0.012	-0.028	0.026	-0.017
	(-1.63)	(-1.16)	(0.54)	(-1.41)	(1.01)	(-0.76)
Tone = 2, Optimistic	-0.014	-0.006	0.017	-0.008	0.012	-0.014
	(-0.66)	(-0.27)	(0.62)	(-0.30)	(0.39)	(-0.49)
Precision = 1, Medium	-0.010	-0.005	0.020	0.001	0.018	-0.002
	(-0.90)	(-0.45)	(1.42)	(0.10)	(1.02)	(-0.12)
Precision = 2, High	-0.000	-0.000	0.021	0.001	0.020	0.001
	(-0.02)	(-0.01)	(1.20)	(0.03)	(0.92)	(0.05)
Pessimistic#Medium	0.054***	0.049**	0.015	0.055**	0.010	0.054*
	(2.88)	(2.44)	(0.58)	(2.18)	(0.34)	(1.89)
Pessimistic#High	0.066**	0.068**	0.045	0.086**	0.024	0.066*
	(2.52)	(2.41)	(1.32)	(2.47)	(0.63)	(1.73)
Optimistic#Medium	-0.003	-0.005	-0.028	-0.005	-0.017	0.006
	(-0.13)	(-0.23)	(-0.96)	(-0.20)	(-0.51)	(0.18)
Optimistic#High	-0.012	-0.014	-0.024	-0.005	-0.024	-0.005
	(-0.54)	(-0.59)	(-0.84)	(-0.18)	(-0.72)	(-0.16)
Observations	21,361	21,361	21,361	18,605	21,361	18,605
R-squared	0.316	0.338	0.372	0.391	0.391	0.409
No Overlap for Longer Horizons			NO	YES	NO	YES
Controls	YES	YES	YES	YES	YES	YES
Rev Controls	YES	YES	YES	YES	YES	YES
Analyst-Firm FE	YES	YES	YES	YES	YES	YES

Table 9b: Volatility and Precision

This table reports the relation between volatility (percentage change in implied volatility of 30-days ATM option) response to the publication of an analyst report and its precision. Reports are sorted in three *Tone* categories and monthly sorted in three *Precision* categories based on sextiles. Controls include size, B/M, number of analysts, uncertainty index, prior-CAR, absolute and relative analyst experience, deviation from consensus, broker size, report length, and change in the VIX. Rev Controls include change in recommendation and change in EPS forecast. Standard errors are double clustered at analyst and industry-week level. t-stats in parenthesis. ***, ** and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
VARIABLES	aIVOL1	aIVOL2	aIVOL4	aIVOL6
Tone = 1, Pessimistic	0.159	0.523	0.271	0.405
	(0.47)	(1.42)	(0.56)	(0.79)
Tone = 2, Optimistic	-0.358	0.410	0.198	0.064
	(-0.81)	(0.83)	(0.33)	(0.11)
Neutral#Medium	-0.190	-0.092	-0.290	-0.238
	(-0.73)	(-0.32)	(-0.77)	(-0.65)
Neutral#High	-0.098	-0.255	-0.583	-0.646
	(-0.27)	(-0.69)	(-1.22)	(-1.34)
Pessimistic#Medium	-0.244	-0.487	-0.412	-0.560
	(-0.80)	(-1.40)	(-0.97)	(-1.24)
Pessimistic#High	-0.925**	-0.955*	-1.196*	-1.374**
	(-2.06)	(-1.94)	(-1.86)	(-2.12)
Optimistic#Medium	-0.292	-0.459	-0.468	-0.364
	(-0.69)	(-0.98)	(-0.82)	(-0.63)
Optimistic#High	0.259	-0.225	-0.069	-0.037
	(0.52)	(-0.40)	(-0.10)	(-0.05)
Observations	17,995	17,995	17,995	17,995
R-squared	0.465	0.476	0.448	0.494
Controls	YES	YES	YES	YES
Rev Controls	YES	YES	YES	YES
Abnormal VIX	YES	YES	YES	YES
Analyst-Firm FE	YES	YES	YES	YES

Table 10a: Variation in price reaction to analyst report precision – Firm Uncertainty

This table reports the relation between price reaction (Cumulative Abnormal Returns, multiplied by 100) to the publication of an analyst report and its precision. Reports are sorted in three *Tone* categories and monthly sorted in three *Precision* categories based on quartiles. *Firm Uncertainty* is equal to one if idiosyncratic volatility is above median. Controls include size, B/M, number of analysts, uncertainty index, prior-CAR, absolute and relative analyst experience, deviation from consensus, broker size, and report length. Rev Controls include change in recommendation and change in EPS forecast. Coefficients of the interaction between *Firm Uncertainty* and *Tone* and between *Firm Uncertainty* and *Precision* are not reported for compactness. Standard errors are double clustered at analyst and industry-week level. t-stats in parenthesis. ***, ** and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

VARIABLES	(1) CAR[0,1]	(2) CAR[0,2]	(3) CAR[0,4]	(4) CAR[0,6]
Tone = 1, Pessimistic	-0.481***	-0.344**	-0.340*	-0.164
	(-3.76)	(-2.44)	(-1.89)	(-0.77)
Tone = 2, Optimistic	0.064	-0.010	0.141	0.157
	(0.38)	(-0.05)	(0.65)	(0.61)
Confidence = 1, Medium	-0.124	-0.073	-0.117	-0.130
	(-1.04)	(-0.55)	(-0.79)	(-0.76)
Confidence = 2, High	-0.039	0.035	-0.079	-0.170
	(-0.27)	(0.22)	(-0.43)	(-0.83)
Firm Uncertainty	0.076	0.077	0.198	-0.014
	(0.43)	(0.38)	(0.82)	(-0.05)
Pessimistic#Medium	0.208	0.065	0.239	0.196
	(1.51)	(0.42)	(1.21)	(0.84)
Pessimistic#High	0.064	-0.098	-0.023	-0.158
	(0.43)	(-0.57)	(-0.10)	(-0.57)
Optimistic#Medium	0.336*	0.286	0.225	0.109
	(1.93)	(1.48)	(1.01)	(0.40)
Optimistic#High	0.169	0.270	0.159	0.186
	(0.91)	(1.26)	(0.62)	(0.62)
Pessimistic#Medium#Firm Uncertainty	-0.415	-0.277	-0.204	-0.292
	(-1.41)	(-0.85)	(-0.54)	(-0.64)
Pessimistic#High#Firm Uncertainty	-0.694**	-0.788**	-1.087**	-1.485***
	(-1.97)	(-2.07)	(-2.40)	(-2.80)
Optimistic#Medium#Firm Uncertainty	-0.205	-0.153	0.089	0.174
	(-0.67)	(-0.46)	(0.21)	(0.35)
Optimistic#High#Firm Uncertainty	-0.320	-0.495	-0.496	-0.882
	(-0.86)	(-1.18)	(-0.99)	(-1.51)
Observations	21,361	21,361	21,361	21,361
R-squared	0.243	0.229	0.219	0.212
Controls	YES	YES	YES	YES
Rev Controls	YES	YES	YES	YES
Analyst-Firm FE	YES	YES	YES	YES

Table 10b: Variation in price reaction to analyst report precision – Analysts' Disagreement

This table reports the relation between price reaction (Cumulative Abnormal Returns, multiplied by 100) to the publication of an analyst report and its precision. Reports are sorted in three *Tone* categories and monthly sorted in three *Precision* categories based on quartiles. *High Disagreement* is equal to one if analysts' disagreement (sd of analysts' EPS forecasts) is above median. Controls include size, B/M, number of analysts, uncertainty index, prior-CAR, absolute and relative analyst experience, deviation from consensus, broker size, and report length. Rev Controls include change in recommendation and change in EPS forecast. Coefficients of the interaction between *High Disagreement* and *Tone* and between *High Disagreement* and *Precision* are not reported for compactness. Standard errors are double clustered at analyst and industry-week level. t-stats in parenthesis. ***, ** and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
VARIABLES	CAR[0,1]	CAR[0,2]	CAR[0,4]	CAR[0,6]
Tone = 1, Pessimistic	-0.557***	-0.358*	-0.648**	-0.480
	(-2.96)	(-1.67)	(-2.55)	(-1.58)
Tone = 2, Optimistic	0.158	0.205	0.282	0.318
	(0.80)	(0.92)	(0.99)	(0.97)
Confidence = 1, Medium	0.008	0.068	0.018	0.117
	(0.05)	(0.42)	(0.10)	(0.51)
Confidence = 2, High	-0.188	-0.100	-0.284	-0.068
	(-1.09)	(-0.53)	(-1.29)	(-0.26)
High Disagreement	0.043	0.075	0.005	0.061
	(0.24)	(0.39)	(0.02)	(0.22)
Pessimistic#Medium	-0.048	-0.247	0.135	0.176
	(-0.19)	(-0.90)	(0.41)	(0.47)
Pessimistic#High	0.025	-0.156	0.039	-0.034
	(0.10)	(-0.55)	(0.11)	(-0.08)
Optimistic#Medium	0.288	0.080	-0.102	-0.173
	(1.37)	(0.35)	(-0.34)	(-0.50)
Optimistic#High	0.128	0.136	0.032	-0.247
	(0.58)	(0.54)	(0.10)	(-0.63)
Pessimistic#Medium#High Disagreement	0.052	0.236	-0.011	-0.177
	(0.17)	(0.67)	(-0.03)	(-0.36)
Pessimistic#High#High Disagreement	-0.477	-0.524	-0.929*	-1.264**
	(-1.39)	*(-1.41)*	*(-1.94)*	*(-2.32)*
Optimistic#Medium#High Disagreement	-0.079	0.216	0.601	0.591
	(-0.25)	(0.64)	(1.44)	(1.20)
Optimistic#High#High Disagreement	-0.156	-0.139	-0.124	0.066
	(-0.44)	(-0.37)	(-0.27)	(0.12)
Observations	21,361	21,361	21,361	21,361
R-squared	0.243	0.229	0.219	0.213
Controls	YES	YES	YES	YES
Rev Controls	YES	YES	YES	YES
Analyst-Firm FE	YES	YES	YES	YES

Table 11: Variation in price reaction to analyst report precision – Tone Position

This table reports the relation between price reaction (Cumulative Abnormal Returns, multiplied by 100) to the publication of an analyst report and its precision. Reports are sorted in three *Tone* categories and monthly sorted in three *Precision* categories based on sextiles. *Tone* is measured based on the first 30 sentences reports. Controls include size, B/M, number of analysts, uncertainty index, prior-CAR, absolute and relative analyst experience, deviation from consensus, broker size, and report length. Rev Controls include change in recommendation and change in EPS forecast. Standard errors are double clustered at analyst and industry-week level. t-stats in parenthesis. ***, ** and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
VARIABLES	CAR[0,1]	CAR[0,2]	CAR[0,4]	CAR[0,6]
Tone = 1, Pessimistic	-0.341**	-0.171	-0.111	0.158
	(-2.38)	(-1.02)	(-0.52)	(0.65)
Tone = 2, Optimistic	0.248	0.213	0.086	-0.002
	(1.53)	(1.12)	(0.36)	(-0.01)
Precision = 1, Medium	-0.197*	-0.144	-0.194	-0.179
	(-1.67)	(-1.07)	(-1.29)	(-0.97)
Precision = 2, High	-0.263*	-0.168	-0.265	-0.195
	(-1.71)	(-0.98)	(-1.26)	(-0.82)
Pessimistic#Medium	-0.186	-0.234	-0.117	-0.275
	(-1.10)	(-1.22)	(-0.49)	(-0.97)
Pessimistic#High	-0.308	-0.482*	-0.546*	-0.734**
	(-1.27)	(-1.83)	(-1.71)	(-2.11)
Optimistic#Medium	0.164	0.137	0.268	0.288
	(0.97)	(0.71)	(1.09)	(1.04)
Optimistic#High	0.008	-0.005	0.350	0.314
	(0.04)	(-0.02)	(1.17)	(0.92)
Observations	21,361	21,361	21,361	21,361
R-squared	0.236	0.222	0.213	0.206
Controls	YES	YES	YES	YES
Rev Controls	YES	YES	YES	YES
Analyst-Firm FE	YES	YES	YES	YES

Table 12: Variation in price reaction to analyst report precision – Precision Position

This table reports the relation between price reaction (Cumulative Abnormal Returns, multiplied by 100) to the publication of an analyst report and its precision. Reports are sorted in three *Tone* categories and monthly sorted in three Precision categories based on sextiles. *Precision* is measured based on the last 30 sentences of the reports. Controls include size, B/M, number of analysts, uncertainty index, prior-CAR, absolute and relative analyst experience, deviation from consensus, broker size, and report length. Rev Controls include change in recommendation and change in EPS forecast. Standard errors are double clustered at analyst and industry-week level. t-stats in parenthesis. ***, ** and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)
VARIABLES	CAR[0,1]	CAR[0,2]	CAR[0,4]	CAR[0,6]	CAR[0,10]
Tone = 1, Pessimistic	-0.363**	-0.319*	-0.136	0.189	0.311
	(-2.25)	(-1.78)	(-0.64)	(0.79)	(1.03)
Tone = 2, Optimistic	0.329*	0.276	0.182	0.116	0.305
	(1.78)	(1.42)	(0.74)	(0.40)	(0.87)
Precision = 1, Medium	-0.044	0.050	0.082	0.152	0.100
	(-0.42)	(0.42)	(0.58)	(0.90)	(0.48)
Precision = 2, High	-0.076	0.025	0.106	0.226	0.191
	(-0.54)	(0.16)	(0.59)	(1.08)	(0.71)
Pessimistic#Medium	-0.065	0.023	-0.014	-0.200	-0.254
	(-0.36)	(0.11)	(-0.06)	(-0.75)	(-0.76)
Pessimistic#High	-0.386*	-0.524**	-0.767**	-1.037***	-0.998**
	(-1.69)	(-2.05)	(-2.50)	(-2.88)	(-2.19)
Optimistic#Medium	0.001	-0.069	0.057	-0.002	-0.432
	(0.01)	(-0.34)	(0.22)	(-0.01)	(-1.18)
Optimistic#High	-0.161	-0.128	-0.063	-0.175	-0.506
	(-0.70)	(-0.54)	(-0.20)	(-0.48)	(-1.09)
Observations	21,361	21,361	21,361	21,361	21,361
R-squared	0.235	0.221	0.212	0.206	0.201
Controls	YES	YES	YES	YES	YES
Rev Controls	YES	YES	YES	YES	YES
Analyst-Firm FE	YES	YES	YES	YES	YES

Table 13: Price reaction to analyst report precision – Events Restrictions

This table reports the relation between price reaction (Cumulative Abnormal Returns, multiplied by 100) to the publication of an analyst report and its precision. Reports are sorted in three *Tone* categories and monthly sorted in three Precision categories based on sextiles. The first restriction excludes days when more than 50% of analyst covering a firm issued an earnings forecast. The second restrictions limits the sample to days when five or less forecasts where issued and the third restriction when three or less. Controls include size, B/M, number of analysts, uncertainty index, prior-CAR, absolute and relative analyst experience, deviation from consensus, broker size, and report length. Rev Controls include change in recommendation and change in EPS forecast. Standard errors are double clustered at analyst and industry-week level. t-stats in parenthesis. ***, ** and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Restriction 1				*Restriction 2*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	CAR[0,1]	CAR[0,2]	CAR[0,4]	CAR[0,6]	CAR[0,1]	CAR[0,2]	CAR[0,4]	CAR[0,6]
Tone = 1, Pessimistic	-0.183	-0.008	0.010	0.367	-0.124	0.036	0.061	0.386
	(-1.19)	(-0.05)	(0.05)	(1.49)	(-0.82)	(0.20)	(0.29)	(1.63)
Tone = 2, Optimistic	0.175	0.145	0.085	0.055	0.126	0.091	0.067	0.052
	(0.92)	(0.70)	(0.32)	(0.18)	(0.66)	(0.43)	(0.25)	(0.17)
Precision = 1, Medium	-0.111	-0.045	-0.075	0.006	-0.134	-0.072	-0.099	-0.037
	(-1.04)	(-0.37)	(-0.54)	(0.03)	(-1.24)	(-0.57)	(-0.69)	(-0.21)
Precision = 2, High	-0.167	-0.061	-0.036	0.071	-0.123	0.004	0.019	0.127
	(-1.26)	(-0.41)	(-0.20)	(0.33)	(-0.92)	(0.02)	(0.10)	(0.57)
Pessimistic#Medium	-0.204	-0.307	-0.147	-0.383	-0.115	-0.211	-0.063	-0.278
	(-1.20)	(-1.57)	(-0.63)	(-1.39)	(-0.70)	(-1.10)	(-0.28)	(-1.06)
Pessimistic#High	-0.380*	-0.575**	-0.572*	-0.925***	-0.348	-0.565**	-0.625**	-0.835**
	(-1.73)	(-2.33)	(-1.90)	(-2.67)	(-1.62)	(-2.29)	(-2.11)	(-2.41)
Optimistic#Medium	0.070	-0.020	0.082	-0.043	0.119	0.004	0.062	-0.115
	(0.34)	(-0.09)	(0.29)	(-0.14)	(0.58)	(0.02)	(0.22)	(-0.37)
Optimistic#High	0.026	0.019	0.075	-0.053	0.085	0.026	0.026	-0.061
	(0.12)	(0.08)	(0.25)	(-0.15)	(0.39)	(0.11)	(0.08)	(-0.17)
Observations	20,826	20,826	20,826	20,826	19,896	19,896	19,896	19,896
R-squared	0.230	0.217	0.211	0.205	0.237	0.222	0.214	0.210
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Rev Controls	YES	YES	YES	YES	YES	YES	YES	YES
Analyst-Firm FE	YES	YES	YES	YES	YES	YES	YES	YES

Table 13: Price reaction to analyst report precision – Events Restrictions – Cont.

	Restriction 3			
	(5)	(6)	(7)	(8)
VARIABLES	CAR[0,1]	CAR[0,2]	CAR[0,4]	CAR[0,6]
Tone = 1, Pessimistic	-0.088	0.094	0.096	0.471*
	(-0.58)	(0.51)	(0.44)	(1.89)
Tone = 2, Optimistic	-0.002	-0.083	-0.199	-0.288
	(-0.01)	(-0.40)	(-0.72)	(-0.96)
Precision = 1, Medium	-0.157	-0.135	-0.147	-0.054
	(-1.41)	(-1.03)	(-0.98)	(-0.29)
Precision = 2, High	-0.164	-0.056	-0.032	0.048
	(-1.18)	(-0.35)	(-0.16)	(0.20)
Pessimistic#Medium	-0.089	-0.225	-0.059	-0.336
	(-0.54)	(-1.15)	(-0.26)	(-1.22)
Pessimistic#High	-0.337	-0.565**	-0.606**	-0.934**
	(-1.51)	(-2.23)	(-1.97)	(-2.58)
Optimistic#Medium	0.274	0.236	0.355	0.226
	(1.43)	(1.11)	(1.26)	(0.75)
Optimistic#High	0.198	0.238	0.354	0.345
	(0.91)	(0.97)	(1.11)	(0.96)
Observations	18,026	18,026	18,026	18,026
R-squared	0.245	0.227	0.222	0.217
Controls	YES	YES	YES	YES
Rev Controls	YES	YES	YES	YES
Analyst-Firm FE	YES	YES	YES	YES

Table 14: Price reaction to analyst report precision – Long Horizons no Overlap

This table reports the relation between price reaction (Cumulative Abnormal Returns, multiplied by 100) to the publication of an analyst report and its precision. Reports are sorted in three *Tone* categories and monthly sorted in three *Precision* categories based on sextiles. Controls include size, B/M, number of analysts, uncertainty index, prior-CAR, absolute and relative analyst experience, deviation from consensus, broker size, and report length. Rev Controls include change in recommendation and change in EPS forecast. Standard errors are double clustered at analyst and industry-week level. t-stats in parenthesis. ***, ** and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

VARIABLES	(1) CAR[0,4]	(2) CAR[0,6]
Tone = 1, Pessimistic	-0.042	0.269
	(-0.19)	(1.07)
Tone = 2, Optimistic	0.152	0.129
	(0.59)	(0.44)
Precision = 1, Medium	-0.130	-0.119
	(-0.85)	(-0.67)
Precision = 2, High	-0.124	0.018
	(-0.63)	(0.08)
Pessimistic#Medium	-0.220	-0.372
	(-0.87)	(-1.29)
Pessimistic#High	-0.702**	-0.922**
	(-2.04)	(-2.43)
Optimistic#Medium	0.067	0.040
	(0.24)	(0.13)
Optimistic#High	-0.105	-0.218
	(-0.35)	(-0.63)
Observations	18,605	18,605
R-squared	0.227	0.222
Controls	YES	YES
Rev Controls	YES	YES
Analyst-Firm FE	YES	YES

Table 15: Price reaction to analyst report precision – Neutral Reports and Crisis

This table reports the relation between price reaction (Cumulative Abnormal Returns, multiplied by 100) to the publication of an analyst report and its precision. Reports are sorted in three *Tone* categories and monthly sorted in three *Precision* categories based on sextiles. *Neutral* reports are classified as *Pessimistic* if published during the 2008-09 financial crisis. Controls include size, B/M, number of analysts, uncertainty index, prior-CAR, absolute and relative analyst experience, deviation from consensus, broker size, and report length. Rev Controls include change in recommendation and change in EPS forecast. Standard errors are double clustered at analyst and industry-week level. t-stats in parenthesis. ***, ** and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
VARIABLES	CAR[0,1]	CAR[0,2]	CAR[0,4]	CAR[0,6]
Tone = 1, Pessimistic	-0.109	0.032	0.102	0.271
	(-0.83)	(0.20)	(0.55)	(1.26)
Tone = 2, Optimistic	0.290	0.311	0.270	0.112
	(1.56)	(1.54)	(1.09)	(0.40)
Precision = 1, Medium	-0.037	0.023	0.001	0.041
	(-0.34)	(0.18)	(0.01)	(0.24)
Precision = 2, High	0.008	0.103	0.057	0.021
	(0.05)	(0.64)	(0.31)	(0.10)
Pessimistic#Medium	-0.327**	-0.385**	-0.335	-0.520**
	(-2.17)	(-2.14)	(-1.57)	(-2.12)
Pessimistic#High	-0.760***	-0.926***	-0.905***	-0.888***
	(-3.87)	(-3.80)	(-3.23)	(-2.75)
Optimistic#Medium	-0.070	-0.151	-0.042	-0.104
	(-0.35)	(-0.72)	(-0.16)	(-0.36)
Optimistic#High	-0.259	-0.239	-0.154	-0.025
	(-1.22)	(-1.05)	(-0.53)	(-0.08)
Observations	22,040	22,040	22,040	22,040
R-squared	0.233	0.218	0.209	0.204
Controls	YES	YES	YES	YES
Rev Controls	YES	YES	YES	YES
Analyst-Firm FE	YES	YES	YES	YES

Table 16: Price reaction to analyst report precision – Abnormal Measures

This table reports the relation between price reaction (Cumulative Abnormal Returns, multiplied by 100) to the publication of an analyst report and its precision. Reports are sorted in three *Tone* categories and monthly sorted in three Precision categories based on sextiles. Reports with *Precision* equal to 1 or *Tone* equal to 0 are excluded. Controls include size, B/M, number of analysts, uncertainty index, prior-CAR, absolute and relative analyst experience, deviation from consensus, broker size, and report length. Rev Controls include change in recommendation and change in EPS forecast. Standard errors are double clustered at analyst and industry-week level. t-stats in parenthesis. ***, ** and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
VARIABLES	CAR[0,1]	CAR[0,2]	CAR[0,4]	CAR[0,6]
Tone = 1, Pessimistic	-0.324*	-0.105	-0.066	0.265
	(-1.88)	(-0.55)	(-0.29)	(0.98)
Tone = 2, Optimistic	0.151	0.161	0.142	0.163
	(0.82)	(0.80)	(0.54)	(0.55)
Precision = 1, Medium	-0.213	-0.146	-0.191	-0.119
	(-1.61)	(-1.02)	(-1.17)	(-0.60)
Precision = 2, High	-0.239	-0.108	-0.089	-0.023
	(-1.47)	(-0.61)	(-0.42)	(-0.09)
Pessimistic#Medium	-0.111	-0.194	-0.010	-0.224
	(-0.61)	(-0.93)	(-0.04)	(-0.76)
Pessimistic#High	-0.480*	-0.683**	-0.746**	-0.935**
	(-1.95)	(-2.50)	(-2.32)	(-2.53)
Optimistic#Medium	0.215	0.134	0.202	0.037
	(1.08)	(0.65)	(0.73)	(0.12)
Optimistic#High	0.078	0.038	-0.006	-0.099
	(0.36)	(0.16)	(-0.02)	(-0.28)
Observations	20,137	20,137	20,137	20,137
R-squared	0.248	0.234	0.223	0.215
Controls	YES	YES	YES	YES
Rev Controls	YES	YES	YES	YES
Analyst-Firm FE	YES	YES	YES	YES

Appendix A - Variables Description

Variable	Description
Size	ln(Total Assets)
Book-to-Market	Book Value of Common Equity / Market Cap
Number of Analysts	ln(Number of analysts issuing earning forecasts about a firm)
Report Length	ln(Number of Sentences)
Prior-CAR	CAR[-5,-1]
Uncertainty Index	US Economic Policy Uncertainty Index
Deviation from Consensus	EPS - Consensus EPS
Abs(Deviation from Consensus)	|EPS - Consensus EPS|
Absolute Analyst Experience	ln(Number of Quarters since first forecast)
Relative Analyst Experience	ln(Number of Quarters since first forecast about a firm)
Broker Size	ln(Number of analysts issuing forecasts for a broker)
Change in Earning Forecasts	%Change in EPS forecast, zero if first forecast
Change in Recommendation	Change in Recommendation, zero if first or no recommendation

Appendix B - Sample Construction

In this Appendix, I provide more details about how the sample was constructed.

Reports are provided by Thomson One in PDF format. I converted the PDF files to txt files, since they are easier to manipulate. Not all files can be successfully converted (some are, for instance, just pictures or have an uncommon codification) and, hence, must be excluded. Analysts often publish reports that contain just few sentences (for instance, a reminder about date and time of a company event). Given the difficulty of using textual analysis techniques on texts that are very short, I excluded all these reports. I also excluded industry reports that cover several firms since it would be hard to capture to which firm the precision measure refers to.

I removed from the text files disclaimers, disclosures and analyst certifications sections as well as numerical tables. I also removed other parts of the report that are not content like contact information and footnotes. I then removed numbers, common stop words (like "the" or "are", from NLTK Python package list) and, for analysis not involving sentences, punctuation. Following some of the existing finance literature using textual analysis (e.g. Huang et al. 2014), I did not perform "stemming" (i.e. using just the root form of a word) due to being problematic in finance applications (using stemming, for instance, "market" and "marketing" are equivalent as well as "operations" and "operating"). Last but not least, I converted each report in an array of words. Following the previous argument, I excluded reports with limited textual data. Specifically, I excluded reports with less than 10 sentences or 100 words[1]. I also excluded the capitalized word "May" since it commonly refers to the month and not the verb[2].

For each report, Thomson One database provides information, among others, about the primary analyst name, the broker ("Contributor") and the date the report was published. First, I matched the different Contributors to IBES estimator IDs. Then, I used the IBES recommendations database to match I/B/E/S analysts' names and analysts' IDs[3] to the Investext names. The same analyst has a unique name in IBES, but can have different names in Thomson One. I was able to match around 4200 Thomson One analysts' names based on unique date-firm-broker pairs. In few words, I

[1] Results barely change if using lower thresholds like 50 words

[2] The results are the same if I do not make this modification

[3] I/B/E/S Academic officially does not provide anymore a table linking the earning forecasts and the recommonedations databases, but it is still possible to match a significant amount of observations using the analyst codes.

exploited the fact some brokers have a unique analyst issuing recommendations about a specific firm in a specific period, so I can match the I/B/E/S analyst code to the reports. I matched also other around 1500 names based on broker and small variations of the Investext analyst name (an analyst can appear as “Smith, John”, others as “Smith, John et al” or “Smith, John and team”) . Further, I manually matched around 420 analysts’ names that were not matched in the previous step, but that issued more than 50 reports. These around 6100 names cover approximately 95% of the reports. Remaining reports were issued by analysts without a name (e.g. “Research Department”) or analysts I was not able to reliably match to I/B/E/S.

Appendix C - Example of Uncertain Sentences

I report here some examples of extract from analysts' reports containing several "uncertainty" words (in bold) according to the list of Loughran and McDonald (2011). Firm names have been redacted.

As we flagged repeatedly recently, the outlook for the thermal coal market is poor in our view. Now, it **appears** the met coal market is also feeling the pain, as the lack of a rebound in global steel production is taking a toll on US met coal demand/exports. While we maintain our view that met coal prices will rebound during the 2nd half of 2012, we **believe** the starting point is getting lower, and we are revising our [. . .] estimates accordingly.

Another factor that **could** cause the stock to rally would be completion of the [...] spin-off faster than we **anticipate**. The stock **may** also perform better than expected if greenhouse gas (GHG) emissions limits are adopted more quickly, or are more accretive to [...] than we project. As with any regulated utility, our rating **could** also be positively or negatively impacted by a significant change, positive or negative, in the regulatory outlook vs. our projection

An increase in aluminum prices is **possible** and can provide further upside. Higher aluminum prices over the forecast period **could** enhance earnings lift from execution. Supply demand characteristics **appear** favorable, and aluminum is currently cheap relative to energy, and competing materials. In addition, we **believe** progress securing long-term power agreements in key operating regions for [...] will allow the company to make attractive incremental investments to expand capacity.

The once-untouchable [. . .] now knows what it is like to compete, in our view. While we **believe** its cost advantage and high margins afford it the most flexibility to price as necessary to defend its turf, we expect the margin impact will be more significant than previously forecast. Although we remain positive longer term, pricing **uncertainty** has been introduced to the [. . .] story, dampening near-term excitement.

Sustained viral responses (SVR) of 6 months after cessation of therapy are required for HCV drug approval, indicating that although early data is extremely impressive, clinical **risk** remains. Moreover, although [. . .] remains the latest-stage PI out of three others known in development, every-8-hour dosing required of this agent keeps the barrier to entry low for a competing drug that **may** be dosed less frequently without compromising efficacy, in our view.

Our recommendations are based on historical trading patterns and our profit expectations, which are subject to a high degree of **risk**. Our profit expectations for [. . .] hinge on our revenue **assumptions**, which **depend** entirely on **assumptions** we make regarding economic growth, the demand for leisure and business travel, the impact of competition from low-fare carriers, and industry-wide aircraft capacity decisions. Our profit expectations also **depend** on **assumptions** we make about the cost of jet fuel (historically a **volatile** commodity), the impact on revenue and expenses of potential labor disruptions, and the impact of any number of geopolitical events and terrorism **risks** on the demand for air travel, among other things.

Appendix D - Precision and Reports' Topics

The main analysis distinguishes reports based on their tone and, when identifying which reports to exclude, whether they are issued around earnings announcements or guidance publications. However, the analysis ignores any variation in terms of topic. An issue could arise if the observed relation between tone or precision and price reaction could be driven by the topic(s) covered by the reports. In this appendix, I describe reports according to their topics and, then, include this new variable to the list of controls in my main specification for price reaction. I also present preliminary results concerning the relation between precision and reports' topics.

For the topic analysis, I employ Latent Dirichlet Allocation (LDA) introduced by Blei et al (2003). LDA is a bag-of-words technique that can be employed to discover latent topics in a collection of documents. The approach is similar to the one used by to Huang et al (2017) to compare topics in conference calls and analyst reports as well as by Fedyk (2018) for Bloomberg news.

LDA and Data Preparation

The assumption behind the LDA algorithm is that a document can be represented as a random mixture over different topics, where each topic is characterized by its specific distribution over the tokens (words). Specifically, the idea is that each word in a document is generated in two main steps:

1. A topic is randomly drawn from the document topics distribution.

2. A word is randomly drawn from the words distribution of the topic obtained in step 1.

A document is, then, generated by repeating the two steps till the random length of the document is reached. Both topics and words distributions are assumed to follow multinomial distributions whose parameters are randomly drawn from Dirichlet distributions with known parameters. The LDA algorithm relies on this model of documents generation to find the topics and words distributions that best fit the documents used to train it.

To summarize, given a corpus of documents, it is possible to use LDA to identify latent topics and the corresponding words distribution. A LDA model can succesively be used to analyze a new document and obtain the contribution of each latent topic to this document.

Before training the LDA model, in addition to the pre-processing explained in the previous part of the paper, I also remove any word that is not a noun (e.g. verbs). I also exclude words that appear in less than 2 reports or more than 70% of the reports. Out of this words list, I keep the 2000 more frequent ones.

Using this dictionary, I train the LDA model on all the reports sample with different numbers of topics (5 to 30, with steps of 5)[4]. The used LDA algorithm is the one available in the Python library "Gensim" (Rehurek and Sojka 2010). I rely on the commonly used perplexity score to identify the optimal model. The idea is to choose a sufficiently diverse set of latent topics; too few topics and each one will tend to capture only very broad concepts, too many topics and several will tend to largely overlap. Figure A1 reports the perplexity scores and it suggests that 20 topics is the optimal number, since the goodness-of-fit does not improve with adding more topics.

Results

Table A1 reports the topics as well as well as their frequency and a list of 15 most relevant words associated with them. To select the most relevant words I use the measure of Sievert and Shirley (2014), with $\lambda = 0.6$. This metrics takes into account the overall frequency of each words and, hence, decreases the relevance of more frequent terms. Based on these most relevant words, I assign a label to each topic. Not surprisingly, a large portion of the reports discuss topics related to valuation, firm performance, management guidance, and management. As aforementioned, the trained LDA model can be used to express each report as a combination of the different topics. I construct 20 different variables containing the proportion of the corresponding topic in each report.

An interesting question is whether there is a relation between topics and precision. I constructed a series of dummy variables for each topic. The dummies take value equal to 1 if the proportion of a topic in a report is greater than 25%. Table A2 shows the results of a regression of $PrecisionC$ on these dummies.[5] Unsurprisingly, topics related to forecasts and expectations such as valuation and management are associated with lower precision, while reports discussing results or guidance appear to be more precise.

Finally, I use the 20 proportion variables as control variables in the main price reaction specifi-

[4]The optimal approach would be to have a training sample separate from the actual sample you want to analyze. A possible approach would be to use "excluded" reports as the training sample

[5]There is no report where topic #20 contributes more than 25%

cation.[6] The results are reported in Table A2 and are largely consistent with the main results if not somewhat stronger, suggesting that report topics are not driving the results.

As future work, it would be interesting to understand whether the relation between precision and different firm or analyst characteristics as well as the relation between precision and market outcomes vary depending on the topic distribution of a report.

[6]The results are equivalent, if not stronger, if the dummy variables are used

Figure A1: Perplexity Scores

This plot depicts the perplexity score of LDA models with different numbers of topics. Perplexity is equal to $e^{-LogLikelihood\ per\ word}$

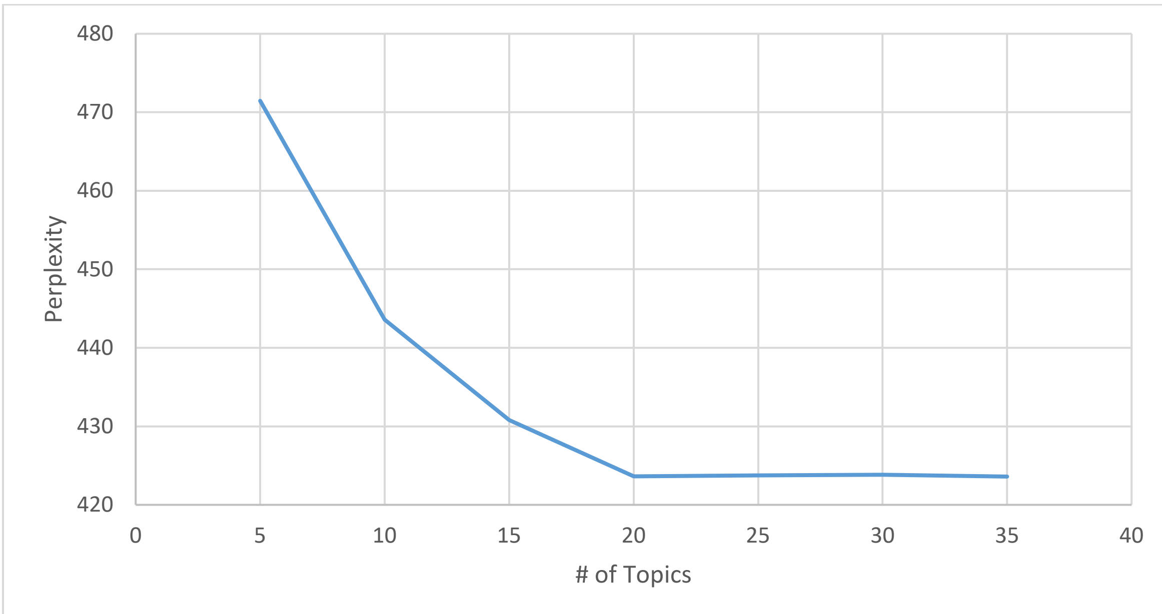


Table A1: LDA Topics

This table reports the 20 topics identified by the LDA algorithm as well as their frequency and a list of 15 most relevant words associated with them. To select the most relevant words I used the measure of Sievert and Shirley (2014), with λ=0.6. I exclude variations of the same word.

#	Topic Label	Most Relevant Words	Frequency
1	Valuation	Target, valuation, stock, earnings, risk, rating, shares, estimates, discount, group, market, p/e, analysis, multiples, value	9.8%
2	Management	Management, years, time, CEO, president, business, strategy, today, opportunity, focus, people, number, meeting, way, investors	9.0%
3	Performance (results)	Quarter, year, sales, income, basis, earnings, increase, expense, tax, results, share, management, operating, points, profit	8.3%
4	Performance (changes and trends)	Margins, cost, volume, pricing, improvement, EBIT, pressure, expansion, cost, savings, leverage, trends, recovery, yoy, headwinds	6.6%
5	Guidance	Guidance, consensus, expectations, street, results, line, estimates, call, range, management, midpoint, conference, outlook, end, forecast	6.2%
6	Retail	Stores, comps, sales, merchandise, apparel, traffic, retailers, inventory, footage, margin, fashion, samestore, fashion, women, department	5.2%
7	Software	Software, storage, product, enterprise, applications, technology, customers, security, market, solutions, platform, products, data, vendors, cloud	5.2%
8	Oil/Energy	Production, oil, gas, rig, drilling, wells, exploration, reserves, crude, activity, prices, Gulf, play, resource, shale	4.9%
9	Strategy & International Business	Sales, China, food, brand, customers, markets, currency, products, Japan, Europe, category, countries, consumers, innovation, categories	4.9%
10	Accounting	Revenues, yoy, bookings, qoq, services, days, business, margin, GAAP, strength, estimates, seasonality, revs, nonGAAP, segment	4.8%
11	Balance Sheet	Cash, flow, dividend, capital, share, debt, sheet, balance, yield, shareholders, buyback, CAPEX, value, repurchase, return	4.7%
12	Technology/Electronics	Semiconductor, equipment, orders, communication, products, systems, markets, segment, wireless, revenues, order, backlog, technology, contracts, electronics	4.1%
13	Power Utility	Fuel, prices, capacity, coal, costs, energy, yield, earnings, power, utility, fleet, plant, gas, contracts, weather	4.0%
14	Real Estate	Property, credit, assets, land, debt, estate, interest, value, housing, sale, portfolio, equity, development, rate, construction	3.7%
15	Supply and Demand	Demand, industry, inventory, markets, units, supply, shipments, capacity, pricing, share, levels, utilization, product, days, production	3.6%

16	Healthcare	Patients, drug, phase, study, data, treatment, trial, approval, patent, pipeline, safety, processing, FDA, reimbursement, disease	3.5%
17	Periodical Data	Week, year, month, trends, checks, March, June, day, period, season, occupancy, July, data, September, December	3.3%
18	Media	Advertising, media, network, wireless, cable, internet, video, content, access, entertainment, subscribers, service, online, users, revenues	2.9%
19	M&A	Acquisition, deal, transaction, synergies, care, merger, integration, agreement, business, accretion, stake, dilution, purchase, EBITDA, assets	2.6%
20	Analyst Notes	Report, research, week, processing, analyst, day, investment, recommendation, firm, information, news, data, spread, metrics, rating	2.6%

Table A2: Precision and Report Topics

This table reports the results of different regressions of *Precision* (in percentage points) on a set of dummies for each LDA topic. Each dummy is equal to 1 if the topic contributes more than 25% to the report. *Earnings* is a dummy equal to one if the report is published within two days from an earnings or management guidance announcement. Standard errors are double clustered at analyst and firm level. t-stats in parenthesis. ***, ** and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Topic Label	(1) PrecisionC	Topic Label	(1) PrecisionC
Valuation	-4.684***	Power Utility	-0.010
	(-20.22)		(-0.03)
Management	-1.209***	Real Estate	-0.557
	(-4.18)		(-1.06)
Performance (results)	3.270***	Suppy & Demand	-1.622***
	(16.85)		(-4.59)
Performance (trends)	0.815***	Healthcare	-1.901***
	(3.71)		(-3.42)
Guidance	3.171***	Periodical Data	0.659
	(14.92)		(0.84)
Retail	2.206***	Media	1.188*
	(6.97)		(1.74)
Software	-0.246	M&A	-2.099***
	(-0.77)		(-3.40)
Oil/Energy	-0.036		
	(-0.07)	Observations	99,499
Strategy & International Business	1.542***	Adjusted R-squared	0.442
	(4.85)	Earnings	YES
Accounting	4.041***	Month FE	YES
	(21.59)	Analyst-Firm FE	YES
Balance Sheet	-0.101		
	(-0.20)		
Technology	2.129***		
	(7.73)		

Table A3: Price reaction to analyst report precision – Controlling for topic

This table reports the relation price reaction (Cumulative Abnormal Return) to the publication of an analyst report and its precision. Reports are sorted in three *Tone* categories and monthly sorted in three *Precision* categories based on sextiles. Controls include size, B/M, number of analysts, uncertainty index, prior-CAR, absolute and relative analyst experience, deviation from consensus, broker size, report length, and report topics distribution. Rev Controls include change in recommendation and change in EPS forecast. Standard errors are double clustered at analyst and industry-week level. t-stats in parenthesis. ***, ** and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
VARIABLES	CAR[0,1]	CAR[0,2]	CAR[0,4]	CAR[0,6]
Tone = 1, Pessimistic	-0.248	-0.078	-0.037	0.298
	(-1.52)	(-0.41)	(-0.17)	(1.12)
Tone = 2, Optimistic	0.251	0.243	0.075	0.022
	(1.31)	(1.12)	(0.28)	(0.07)
Precision = 1, Medium	-0.125	-0.052	-0.110	-0.045
	(-1.10)	(-0.40)	(-0.74)	(-0.25)
Precision = 2, High	-0.262*	-0.145	-0.263	-0.127
	(-1.73)	(-0.85)	(-1.27)	(-0.54)
Pessimistic#Medium	-0.193	-0.285	-0.139	-0.346
	(-1.07)	(-1.37)	(-0.56)	(-1.19)
Pessimistic#High	-0.489**	-0.723***	-0.691**	-0.912**
	(-2.01)	(-2.62)	(-2.04)	(-2.35)
Optimistic#Medium	0.028	-0.035	0.162	0.076
	(0.14)	(-0.16)	(0.58)	(0.25)
Optimistic#High	-0.183	-0.221	-0.045	-0.094
	(-0.83)	(-0.89)	(-0.14)	(-0.26)
Observations	21,361	21,361	21,361	21,361
R-squared	0.245	0.229	0.217	0.210
Controls	YES	YES	YES	YES
Rev Controls	YES	YES	YES	YES
Analyst-Firm FE	YES	YES	YES	YES