Awkward Silence: Is Manager Hesitation Informative?

Umit Kurucak
The University of Texas at Dallas
umit.kurucak@utdallas.edu

October 2025

Acknowledgments: I appreciate helpful comments and suggestions from Vikram Nanda, Alajandro Rivera, Alejandro Zentner, Jedson Pinto, Steven Xiao, Kelsey Wei, Han Xia, Yexiao Xu, Harold Zhang, Umit Gurun, and seminar participants at the University of Texas at Dallas. All errors are my own.

Awkward Silence: Is Manager Hesitation Informative?

Abstract

I investigate whether managers' hesitations provide insights into the future behavior of investors and analysts. Hesitation is defined as the response time (RT) between analyst questions and managerial answers, measured using AI-based speaker diarization and transcript alignment over 7,000 S&P 500 earnings calls (2019–2023). I find that longer RT is associated with lower contemporaneous and 1-quarter-ahead cumulative abnormal returns. A split sample analysis provides empirical evidence for the information uncertainty explanation of Post-Earnings Announcement Drift (PEAD). Analysts revise earnings forecasts downward and show increased uncertainty through higher dispersion. RT does not predict earning surprises, consistent with analysts promptly incorporating the hesitation signal. This paper shows that managerial response time is an additional non-verbal information channel.

JEL classification: G14, G12, G41, M41, C55
Keywords: hesitation, response time, PEAD, information asymmetry, behavioral finance, artificial intelligence, silence

1 Introduction

The use of alternative data sources is rapidly increasing with technological advancement, making it possible to extract information from sources that were previously infeasible to analyze. Specifically, developments in machine learning and artificial intelligence have enabled the effective and accurate analysis of unstructured textual data. In conjunction with this, the growing availability of computational power now allows researchers to examine even more unconventional sources, such as audio and video recordings, for financial information.

In addition to the mandatory 10–Q filings, many publicly traded firms hold voluntary conference calls to discuss their financial performance for the most recent quarter with stakeholders. These calls allow managers to qualitatively communicate messages beyond the quantitative content of the financial reports to investors, analysts, and other stakeholders. In particular, Q&A sessions within the calls serve as an interactive forum where managers respond to analyst questions. The output of an earnings conference call is twofold: an audio recording and a corresponding transcript. Neither of these unstructured sources is straightforward to analyze.

Textual analysis of earnings call transcripts began earlier, due to the availability of technology and the relatively low computational requirements. Researchers have investigated several aspects of the language used by managers and analysts, such as tone, linguistic extremity, complexity, and spontaneity. In contrast, research using audio data began more recently. Existing studies that incorporate audio cues primarily focus on vocal segments to perceive affect, voice quality, or uncertainty. Yet, the study of non-vocal segments, or silence, remains underexplored[1]. Distinct from

[1]There are few studies that mention silence or response time. Hollander et al. (2010) examine silence from a textual perspective, analyzing transcript instances where managers explicitly choose not to answer questions. Closer to this study, Burgoon et al. (2016) were the first to explore manager response time from audio recordings, but their focus was on detecting deception, for which they failed

prior work, this paper investigates whether "silence" or response time, as a proxy for managerial hesitation, conveys economically meaningful information to investors and analysts. I define response time as the delay between end of an analyst's question and start of a manager's response. I argue that response times during the Q&A section may carry meaningful information, as they can serve as a proxy for managerial hesitation. Furthermore, I argue that response time may be a more natural measure of hesitation than vocal cues, which experienced managers might consciously control. In contrast, response time reflects unconscious cognitive processing, and to the best of my knowledge, this is the first study to establish such a link.

The motivation for examining these silent intervals stems from a body of psychology literature. Research in this field suggests that the pauses before a response are not merely empty time but are indicative of cognitive processing. Longer response times have been shown to correlate with periods of cognitive planning or hesitation (Henderson et al. (1966)). Furthermore, listeners often perceive these delays as a sign of insincerity, suggesting that the speaker may be engaging in complex cognitive tasks such as thought suppression or answer fabrication (Ziano and Wang (2021)). This established link between response latency and perceived hesitation provides a strong theoretical foundation for treating silence as an informative, non-verbal cue in high–stakes settings like earnings calls.

I compute managerial response times via speaker diarization and transcript matching. The diarization algorithm processes audio recordings using artificial intelligence to generate a timeline of who speaks when. I then match the resulting speaker timelines to the corresponding transcripts to identify the moments when an analyst's question ends and a manager's response begins, and finally measure the time difference between the two. Using earnings call recordings of S&P 500 firms between 2019

to find a significant correlation.

and 2023, I construct a dataset of over 7,000 conference calls.

Using this novel dataset, I first investigate whether and how the market reacts to variation in managers' response times. If response time carries information about firm fundamentals, I expect the market to react to it. Furthermore, if longer response times are perceived as hesitation, I expect a negative market reaction. Consistent with these predictions, I find that longer response times are associated with negative cumulative abnormal returns in both the short and long run.

Next, I study whether analysts incorporate managerial hesitation into their information processing. I examine the relationship between response time and subsequent revisions in analysts' earnings forecasts. If a prolonged response time is perceived as a signal of uncertainty, analysts are expected to adjust their forecasts downward. Consistent with this interpretation, my analysis reveals that longer response times are associated with more negative revisions to analysts' earnings forecasts for the subsequent quarter. Forecast dispersion also increases with increasing response time. This result indicates that analysts act on idiosyncratic information. A cross-sectional analysis of analyst dispersion suggests that this result is likely attributable to heightened analyst uncertainty stemming from extended manager response times when the firm meets or exceeds earnings expectations.

Finally, I analyze whether response time predicts future firm performance. I regress one- and two-quarter-ahead unexpected earnings on response time, where unexpected earnings are defined as the difference between realized earnings and the median consensus forecast for the corresponding quarter. The association is not statistically significant. However, this lack of association is consistent with earlier findings: analysts appear to revise their forecasts downward following longer response times, which in turn reduces the magnitude of future earnings surprises. In this sense, the absence of a predictive relationship does not contradict the informativeness of hesitation but instead suggests that analysts may incorporate the signal promptly and efficiently.

The main contribution of this study is related to one of the most robust violations of the Efficient Markets Hypothesis: post-earnings-announcement Drift (PEAD). PEAD is the phenomenon in which positive (negative) earnings news is followed by positive (negative) stock returns in the longer period following the earnings announcement, suggesting earnings information is not fully incorporated into prices, hence contradicting market efficiency (Ball and Brown (1968); Bernard and Thomas (1989, 1990)). One strand of theoretical asset pricing literature connects this anomaly to information uncertainty (Lewellen and Shanken (2002); Brav and Heaton (2002); Liang (2003); Zhang (2006)). This study presents empirical evidence by illustrating the immediate market reaction when earnings information is reinforced by a manager hesitation signal. However, a delayed reaction occurs when conflicting signals are presented.

Distinct from prior work, this study utilizes "silence", captured through response time, as a proxy for manager hesitation and demonstrates that it can serve as an incremental source of information in the context of earnings calls. In this regard, the paper contributes to the literature on information asymmetry by introducing a novel, nonverbal channel through which managerial hesitation may be inferred.

Another contribution relates to the technologies employed. The artificial intelligence tools used in this study—such as speaker diarization and audio-based segmentation—are not yet widely adopted in financial economics. By integrating these computational methods, the study bridges the gap between information technologies and empirical finance.

Finally, the study is motivated by insights from behavioral psychology, particularly regarding the interpretation of response times within real–time Q&A settings. In this sense, this work also contributes to the growing literature on behavioral finance by highlighting the informational value of the perception of nonverbal cues.

The remainder of the paper is organized as follows: Section 2 reviews the related literature and develops hypotheses. Section 3 describes the data and the construction of the hesitation measure, and outlines the empirical methodology. Section 4 presents market reaction results, while Section 5 shows the results of analyst reaction tests. Section 6 concludes.

2 Related Research and Hypothesis

Public disclosures can improve the efficiency and liquidity of financial markets (Goldstein and Yang (2017)) by "leveling the playground" (Bowen et al. (2002)). Finance and accounting research has been trying to extract meaningful information from every form of firm disclosure. As a result, researchers have also been investigating numerous aspects of quarterly earnings conference calls. Although not mandatory, most public firms choose to hold these calls to better communicate the information contained in their financial disclosures, providing an additional information channel for stakeholders (Brown et al. (2004)).

During the early years of research, studies were geared towards quantitative analysis; later, qualitative research came into the picture (Feldman et al. (2010)). Earnings calls provide an interactive forum for managers to deliver real-time feedback. Due to this interactive nature, it is possible to convey more information through verbal and nonverbal cues. In this sense, earnings calls serve as a vital voluntary disclosure medium that helps reduce information asymmetries between managers and stakeholders (Frankel et al. (1999); Tasker (1998)).

A large body of work studies the textual content of earnings call transcripts.

Researchers have found that the tone, spontaneity, linguistic complexity, and even gendered speech patterns of managerial language influence investor and analyst perceptions. For instance, deviations from prepared scripts that demonstrate managers' spontaneity have been shown to precede earnings surprises (Lee (2016)), and the overall tone of speech —-including affective word usage—- impacts investor perceptions of firm quality Chen et al. (2018); Price et al. (2012); Feldman et al. (2010)). When managers avoid answering certain questions, this "no-answer" behavior often signals negative information, leading to adverse market reactions (Hollander et al. (2010)). Analysts have also been found to discriminate in their responses, weighting verbal cues differently based on analysts' prior experience and interactions with management Mayew (2008), and gender differences in managerial language use can result in systematic biases in forecast revisions (De Amicis et al. (2021)). Furthermore, higher linguistic complexity in call transcripts is associated with increased forecast dispersion and delayed information assimilation Bushee et al. (2018), while manager optimism—reflected in positive language use—tends to elevate investor expectations (Davis et al. (2015)). Lastly, extreme linguistic expressions can amplify market volatility in the aftermath of calls significantly increases trading volume and stock price reactions, particularly for firms with weaker information environments Bochkay et al. (2020).

As computational tools have advanced, researchers have begun to analyze vocal cues embedded in the audio recordings of earnings calls. Beyond words, vocal indicators such as tone, pitch and can signal affective states of managers and influence stakeholder perceptions. In their pioneer work, Mayew and Venkatachalam (2012) measured positive and negative affect in the voices of managers and demonstrated vocal cues can be informative about a firm's financial prospects. More recent research also incorporated machine learning and artificial intelligence to study various aspects of vocal cues like use of humor (Call et al. (2024)), delivery quality (Baik et al. (2024)), or level of uncertainty (De la Parra and Gallemore (2024)) on firm prospects. There are also some attempts so combine verbal and vocal cues from earnings calls to predict volatility (Sawhney et al. (2021, 2020)) or financial risk (Qin and Yang (2019)).

Despite extensive research on verbal and paralinguistic cues, the informativeness of non-vocal segments, such as pauses before answering questions, remains unexplored. Psychology literature suggests longer pauses indicate periods of cognitive planning or hesitation before responding (Henderson et al. (1966)). For instance, Ziano and Wang (2021) argues that longer response times are often perceived as a sign of insincerity, as they can suggest the responder is engaging in cognitive processes such as thought suppression or answer fabrication. In contrast, Burgoon et al. (2016) fails to find a significant association between response latency and fraudulent activity. On the other hand, longer pauses can also be indicative of spontaneity. Supporting this view, Lee (2016) finds that scripted answers (i.e., a lack of spontaneity) are associated with negative unexpected future firm performance, as they can lead to less informative interactions and negatively impact investor perceptions and decisions.

The questions of whether nonverbal cues such as response time are informative—and, if so, whether these signals are interpreted positively or negatively by the market or analysts—remains unanswered. This study is an attempt to fill this gap by constructing a novel measure of manager hesitation. I then examine whether this indicator of hesitation conveys information that is reflected in stock returns, analyst forecasts, and earning surprises. Building on insights from literature I develop the following hypotheses:

Hypothesis 1 (Market reaction)**.** *H_1: Longer managerial response time is associated* *with lower cumulative abnormal returns.*

Theory: If longer response time reflects hesitation or information withholding, market should price it negatively.

Hypothesis 2 (Future firm performance). H_1*: Longer managerial response time is associated with negative unexpected earnings.*

Theory: If longer response time reflects hesitation or information withholding, this should indicate firm distress and worse future performance.

Hypothesis 3 (Analyst reaction – 1). H_1*: Longer managerial response time is associated with negative forecast revisions.*

Theory: If longer response time reflects hesitation or information withholding, analyst should revise their earnings forecasts downward.

Hypothesis 4 (Analyst reaction – 2). H_1*: Longer managerial response time is associated with positive change in analyst forecast dispersion.*

Theory: If longer response time reflects hesitation or information withholding, this should increase the uncertainty among analyst forecasts.

Collectively, these hypotheses offer the first systematic investigation into whether response time—a silent, nonverbal cue extracted from earnings call audio—contains incremental information about firm fundamentals and how it is interpreted by market participants. This paper uses a novel dataset of over 7,000 earnings calls from S&P 500 firms, combined with artificial intelligence techniques for speaker diarization, to construct a precise measure of managerial hesitation. The next section outlines the data sources, the methodology used to compute response time, and the empirical strategies employed to test the proposed hypotheses.

3 Data

3.1 Sample Selection

The sample consists of firms included in the S&P 500 constituent list at the end of 2023. There are several reasons to this decision. First, these firms are large publicly traded firms for which a wide range of financial data is available. Second, these firms are more likely to hold earnings conference calls, and covered by audio datasets. Finally, because audio processing and speaker diarization are computationally intensive tasks, limiting the sample to a manageable set of firms ensures that the analysis remains feasible within reasonable time and resource constraints.

I obtained audio recordings of quarterly earnings calls of S&P 500 firms from SeekingAlpha [2]. I was able to scrape audio files of earning calls between 2018 and 2023, but with very few files in 2018 (31 files). As a result, I restrict the sample between 2019 and 2023. I obtain corresponding transcripts from CapitalIQ.

3.2 Audio Data and Hesitation Measure Construction

Social psychology literature suggests that longer pauses when answering questions can indicate uncertainty and cognitive planning (Ziano and Wang (2021)). Hence, if we can effectively measure these pauses, this data could potentially serve as a proxy to detect manager hesitation.

To this end, I utilize ’pyannote-audio’, a Python-based diarization library that leverages artificial intelligence to determine "who spoke when" in audio recordings. Diarization refers to the process of identifying and labeling speakers within an audio or video recording to determine exactly when each speaker is active (Park et al. (2022);

[2] https://seekingalpha.com

Tranter and Reynolds (2006); Anguera et al. (2012)). The origins of diarization date back to the early 1990s, when it was initially developed to assist automatic speech recognition systems by separating segments belonging to different speakers (Gish et al. (1991)). Traditionally, diarization systems comprise several sub-modules, including speech activity detection, segmentation, feature extraction, and clustering algorithms (Shum et al. (2011); Sell et al. (2018)). With recent advances in artificial intelligence technology, the efficiency and accuracy of these diarization systems have significantly improved (Snyder et al. (2018); Fujita et al. (2019); Park et al. (2022)).

Diarization procedure produces an "RTTM" file [3] for each audio file. These are actually text files that store tab separated values of speaker labels, speech timings and durations. I clean the data to get rid of any artifacts that can be caused by audio quality issues, background noise or overlapping speech. Figure 1 shows an excerpt from an RTTM file.

[Figure 1 about here.]

This process only assigns labels to the speakers, but the actual names and types of speakers (operator, manager, or analyst) still need to be determined. It Is easier to match the speaker labels for the operator and managers is easier as the conference calls are generally following a structured order. Once the operator and the manager labels are matched, the remainder of the speakers must be analysts. I remove any conference calls that fails to meet the expected structure, and hence cannot be confidently matched.

I measure the response times by calculating the time difference between the end of an analyst segment and the start of consequent manager segment. This proce-

[3] RTTM (Rich Transcription Time Marked) is a standard file format used in speech processing tasks, especially speaker diarization. Each line in an RTTM file corresponds to a single speaker segment and includes metadata such as start time, duration, speaker label, and channel.

dure generates several response times for each earnings call. I estimated the median response time within each call to use as the independent variable. This procedure yields 7,460 firm call observations from 471 firms. After merging with financial data and removing observations with missing control variables, the final sample consists of 7,035 observations. Figure 2 illustrates an example of detected question end and answer start times plotted on the waveform of an earnings call audio recording.

[Figure 2 about here.]

3.3 Returns, Analyst Behavior and Firm Performance Indicators

I follow Mayew and Venkatachalam (2012) in constructing return and firm performance measures. Normal returns are estimated using Fama–French size and book-to-market sorted portfolio returns.[4] The constructed variable CAR(i, j) denotes the cumulative abnormal return from day i to day j (both inclusive) around the earnings call.

Analyst response indicators, $FREV$ stands for analyst forecast revisions, ΔDispersion stands for change in analyst forecast dispersion. Future firm performance indicator UE_t (Unexpected earnings at period t) is calculated using I/B/E/S data. Summary statistics of the variables used in the study is provided in Table 1.

[Table 1 about here.]

The data and methodology outlined above provide the basis for examining whether response time conveys incremental information as well as insights about PEAD. The next section presents the empirical evidence.

[4] As a robustness check, I also estimate normal returns using the Fama and French (1993) and Carhart (1997) four-factor model. The results remain qualitatively unchanged and presented in Appendix A.3.

4 Market Reaction to Manager Hesitation

4.1 Contemporaneous Market Reaction

How do the market participants react to manager hesitation? If response time is a perceivable measure of manager hesitation or information withholding, then the natural prediction is that longer response times should indicate negative short term reactions. In order to test this hypothesis, I estimate the following specification [5]:

$$\begin{aligned} CAR(0,1) = \beta\,\mathrm{RT} + \gamma'\mathrm{FirmControls} \\ + \theta'\mathrm{ExecControls} + \phi'\mathrm{AudioControls} + \delta'\mathrm{TextControls} + \varepsilon \end{aligned} \tag{1}$$

Here, RT stands for the variable of interest and represents the median response time of managers to analyst questions in seconds. I follow Mayew and Venkatachalam (2012) to incorporate size, growth, and risk firm controls, which are linked to returns (Collins and Kothari (1989)). Size is expressed as the logarithm of total assets. The book-to-market ratio, calculated as the book value of shareholders' equity at the end of the current quarter divided by the market value of equity, serves as a proxy for growth. Return volatility serves as a proxy for risk. Additionally, I control for return momentum. Return volatility is measured as the standard deviation of returns, while momentum is measured as the cumulative abnormal returns between -127 and -2 days before the earnings. UE_t represents the unexpected earnings, defined as the difference between the actual I/B/E/S earnings per share and the I/B/E/S analyst summary consensus median earnings per share, scaled by the price per share two

[5] I follow the pooled OLS specification of Mayew and Venkatachalam (2012). As a robustness test, I also estimate panel regression with firm and quarter fixed effects. The results are robust and presented in Appendix A.3.

days before the conference call. I also include a dummy for dividend payments, which takes a value of 1 if the firm pays dividend in the quarter, and 0 otherwise, to proxy for uncertainty (Price et al. (2012)).

Executive controls, include CEO age and tenure to control for experience, and Holder 67 of Malmendier and Tate (2005) for manager overconfidence.

To illustrate the incremental informativeness of silence (i.e., median response time), I construct several acoustic control variables that indicate various emotional states of the managers, using the same audio recordings. These variables can be classified under 5 groups: tone, intensity, instability, voice quality, and tempo. Tone is measured by mean pitch (fundamental frequency), pitch variation, and pitch slope. A high pitch is associated with lack of confidence (Yang et al. (2023), pitch variation with consistency, and an upward slope with arousal/excitement (Burgoon et al. (2016)). Intensity id measured by loudness. Loudness is associated with arousal and negative affect (Burgoon et al. (2016) as well as confidence (Yang et al. (2023). Instability of voice perturbation is measured by jitter and shimmer. Jitter is the average instability in pitch and associated with uncertainty or anxiety (Fuller et al. (1992); De la Parra and Gallemore (2024). Shimmer is the average instability in voice amplitude associated with voice quality, health (Teixeira et al. (2013)) and uncertainty (De la Parra and Gallemore (2024)). Voice quality is measured by harmonic–to–noise ratio (HNR) (Yang et al. (2023)). Finally, tempo is measured by speech rate, which is associated with anxiety (Goberman et al. (2011)).

I use openSMILE[6] toolkit (Eyben et al. (2010)) to extract audio features from manager utterances corresponding to each manager reply during Q&A session. Each feature is normalized by the length of the utterance following De la Parra and Galle-

[6]openSMILE is an open-source toolkit for acoustic feature extraction. It is developed by Eyben et al. (2010). Accessible at https://audeering.github.io/opensmile-python/

more (2024) and median of all utterances are used for the aggregate firm–quarter–call observation.

Final set of control variables are consisting of text controls. These are manager and analyst tone to control for positive and negative tone; Manager and analyst FOG index for readability, analyst word count for question complexity, and finally logarithm of numbers of analyst participating and number of questions asked to control for intensity of questions. Detailed definitions of all variables used are given in Table A.1.

Following Mayew and Venkatachalam (2012), I estimate Eq.1 using pooled ordinary least squares regression. Table 2 reports the estimated values of the impact of response time on the contemporaneous CARs. In column (2), the coefficient on response time (RT) is -0.289 (s.e. 0.130) and statistically significant indicating that, on average, each additional second of delay in response time is associated with a -0.289% decline in contemporaneous abnormal return. This translates into a 0.16% decline for every standard deviation increase. Firms with higher book-to-market ratios and greater return volatility exhibit larger CARs, while larger firms experience smaller effects. As expected, unexpected earnings (UE_t) has a positive coefficient indicating market reaction to earning surprises. Momentum also enter positively and significantly. In line with Davis et al. (2012); Price et al. (2012), manager tone and analyst tone are strong predictors of CAR, underscoring that more positive language in the call boosts short-term returns. This result indicates that even after controlling for firm, risk, and text factors, the novel non-verbal measure of response time offers additional informational value.

[Table 2 about here.]

Before attributing economic significance to this variable, it's crucial to ensure that the hesitation measure accurately captures managerial hesitation rather than random fluctuations or unrelated delays. The next section addresses this issue.

4.2 Validity of Hesitation Measure

While the preceding analysis demonstrates the proposed hesitation measure carry information on market reaction, it remains important to show this measure can be used as a proxy for managerial hesitation or information withholding. To this end, I implement a tercile-split analysis by grouping observations based on the distribution of response time as low, mid, and high response time. Then, I estimate Eq 1 on these groups. If longer response times successfully proxy hesitation, I expect to see an association on the high response time group rather than lower groups.

Table 3 presents the results. Consistent with this prediction, the coefficient on response time is statistically significant only in the high group, indicating that the market reacts negatively when managerial delays are unusually long. Coefficient estimates in the first and second terciles are smaller and statistically insignificant. This finding supports the interpretation of response time as a valid proxy for managerial hesitation.

[Table 3 about here.]

Having established that managerial response time is associated with immediate market reactions and serves as a valid proxy for hesitation, the next section examines the predictability of longer term returns.

4.3 Future Return Predictability

The evidence presented so far presents informative value of managerial hesitation and investors incorporate this information in their decision processes. Post-earnings announcement drift (PEAD) is a well-documented phenomenon that suggests that positive (negative) earnings news is followed by positive (negative) CARs in the longer period following the earnings announcement (Bernard and Thomas (1989)). This indicates that the information contained in the earnings news is not fully incorporated in the contemporaneous returns. Therefore, I anticipate managerial hesitation to have some predictive power on longer–term returns.

The specification in Eq. 1 is suitable for testing this intuition. If there is a systematic association that is not fully incorporated by contemporaneous returns, my hesitation measure should be able to predict longer term stock returns. I predict a negative association between response time and longer horizon CARs. To this end, I estimate the following for 3 and 6 – month CARs (i.e., $j = 90, 180$ days):

$$\begin{aligned} CAR(2,j) &= \beta\,\mathrm{RT} + \gamma'\mathrm{FirmControls} \\ &\quad + \theta'\mathrm{ExecControls} + \phi'\mathrm{AudioControls} + \delta'\mathrm{TextControls} + \varepsilon \end{aligned} \tag{2}$$

Table 4 present the results of Eq. 2 for one and two-quarter horizons. Managerial response time significantly predicts abnormal returns over the subsequent quarter but loses predictive power for longer horizons. One second increase is associated with a 51.9 basis points (S.E. 0.256) decrease in one-quarter ahead CARs. This finding is consistent with the notion that investors receive new earnings-related information each quarter, potentially diluting the effect of earlier signals. While the coefficient on unexpected earnings remains statistically significant across all specifications, its sign reverses over longer horizons. This pattern may reflect an initial overreaction by the market to earnings surprises, followed by a mean-reverting adjustment in returns

De Bondt and Thaler (1985). This result contradicts the conventional interpretation of PEAD, which suggests that investors underreact to earnings news and fully incorporate the information in longer horizons (Ball and Brown (1968)). However, parameter uncertainty model of Lewellen and Shanken (2002) and subsequent empirical work (e.g. Zhang (2006)) claims that the slower response can stem from information uncertainty.

The next section delves into heterogeneity tests to shed light on the channels contributing to this result.

[Table 4 about here.]

4.4 Market Response under Varying Levels of Scrutiny

Earnings conference calls are held shortly after the quarterly financial reports are released. Investors and analysts are aware of any unexpected earning news before joining the call, which may influence their perception of the managers' verbal and nonverbal cues. I expect a stronger negative association between returns and manager hesitation when firms are under greater scrutiny.

To test this prediction, as per Mayew and Venkatachalam (2012), I divided the RT data into two groups: RT^{HS} and RT^{LS}, based on the level of scrutiny faced by the firms. RT^{HS} represents the median response time of high scrutiny firms, which are the firms that fail to meet earnings expectations, i.e., $UE_t < 0$. RT^{HS} is equal to zero otherwise. On the contrary, median response time for low scrutiny firms, RT^{LS}, takes the value of RT if $UE_t \geq 0$ and 0 otherwise. This helps to further understand the source of stock market reaction to delays in response time. I estimate the following modified specification:

$$
\begin{aligned}
CAR(i,j) = \beta_0 RT^{HS} + \beta_1 RT^{LS} + \gamma'\text{FirmControls} \\
+ \theta'\text{ExecControls} + \phi'\text{AudioControls} + \delta'\text{TextControls} + \varepsilon
\end{aligned}
\tag{3}
$$

Here, i, j takes values of 0 and 1 for contemporaneous cumulative abnormal returns. For one and two–quarter–ahead CARs, $i = 2$ and $j \in 90, 180$.

Table (5) displays results. Column (2), shows the estimates of contemporaneous market reaction. As predicted, high-scrutiny firms exhibit a stronger negative association, while low-scrutiny firms do not have any significant association. The coefficient on RT^{HS} is -1.253% (s.e. 0.167), significant at the 1% level, implying that for the firms facing greater scrutiny, each extra second of response time erodes nearly 1% of CAR. This pattern suggests that, in the short run, markets particularly penalize response delays at firms subject to greater scrutiny.

Long-term CARs, however, reveal some surprising results. Contrary to expectations, the association appears to originate from low-scrutiny firms for both 3 and 6-month CARs (columns (4) and (6)). These results collectively indicate that the market reacts to response time regardless of the level of scrutiny. However, it quickly takes action on recent negative news about high-scrutiny firms while being cautious and slow to react to perceived hesitation from low-scrutiny firms.

These results support the informational uncertainty explanation of PEAD. When a negative earning surprise signal is accompanied by other negative signals, such as a manager's hesitation signal, the information is certain, and the market reacts swiftly (possibly overreacting) to these signals. However, when the signals point in opposite directions (for instance, when the firm meets its earning expectations but the manager is hesitant), this creates information uncertainty, leading to a slower response and hence, drift.

[Table 5 about here.]

4.5 Future Firm Performance

After establishing the predictive power of managerial response times on market reactions, in this section, I investigate whether the response time can also predict future firm performance. I follow Mayew and Venkatachalam (2012) to use unexpected earnings, UE_t, as a proxy for future firm performance. I estimate the following for two quarters after the earnings call:

$$\begin{aligned} UE_{t+1} = \beta_0\,\mathrm{RT} + \beta_1 FREV + \beta_2 \Delta Dispersion \\ + \gamma'\mathrm{FirmControls} + \theta'\mathrm{ExecControls} \\ + \phi'\mathrm{AudioControls} + \delta'\mathrm{TextControls} + \varepsilon \end{aligned} \tag{4}$$

In this specification analyst forecast revision, $FREV$, and analyst forecast dispersion,$\Delta Dispersion$, are added to the specification to control form analyst reaction to current earning news. These variables are as defined in Section 5. All other variables are same as in Eq. (1). Table 6 presents the estimates from Eq. (4).

[Table 6 about here.]

I predict a positive correlation between response time and future earning surprises. However, as shown in Table 6, I cannot confirm this prediction and fail to reject the null hypothesis, $\beta_1 = 0$. This result suggests that analysts incorporate hesitation information into their forecasts, effectively eliminating predictability of future unexpected earnings, Section 5 investigates this prediction in detail.

Next section extends the discussion into information asymmetry implications of response time measure.

4.6 Information Asymmetry

The evidence presented thus far indicates that investors incorporate the information contained in managerial response time. If this source of information offers any incremental value, it should reduce the information gap between informed and uninformed investors. To test this prediction, I split the sample into three groups according to the level of institutional ownership and estimate Eq. 1 for these subsamples. Level of institutional ownership is often used as a proxy for investor sophistication (Hollander et al. (2010)). Firms with high levels of institutional ownership should have access to more information than firms with low levels of institutional ownership. Therefore, I expect the firms with low institutional ownership should benefit from additional information sources.

The results are presented in Table 7 and confirms the prediction. The coefficient of managerial response time is negative, statistically significant and more pronounced for low–institutional ownership firms then mid– and high–institutional ownership firms. This result suggests that investors with informational disadvantages incorporate the information conveyed in managerial response time into their decisions, thereby effectively reducing information asymmetry. This result also confirms the value of quarterly earnings calls to mitigate with information asymmetry.

[Table 7 about here.]

This result concludes the analysis on the market reaction. The next section investigates analyst perception and reaction to extended response time information.

5 Analyst Reaction to Manager Hesitation

Previous sections suggest response time carry information about manager hesitation and this information is perceived by the investors. How about analysts? Do analysts incorporate this information channel in their forecast decisions as well. I attempt to answer this question in this section.

If analysts perceive managerial hesitation, they are expected to revise their earnings forecasts for the next quarter downward. Moreover, this signal is not expected to be perceived equally by all analysts. Therefore, the dispersion of analysts' forecasts for the next quarter is also expected to increase; accordingly, there should be a positive difference between the dispersion of forecasts made immediately after the earnings call and those made immediately before it. To test these hypotheses, I estimate the following specification [7]. In this setup, the dependent variable Y is either the forecast revision (FREV) or the change in dispersion (ΔDispersion).

$$\begin{aligned} Y_{i,t} = \beta_0 RT_{i,t} + \beta_1 CAR(0,1) + \gamma' \text{FirmControls}_{i,t} \\ + \theta' \text{ExecControls}_{i,t} + \phi' \text{AudioControls}_{i,t} \\ + \delta' \text{TextControls}_{i,t} + \alpha_i + \varepsilon_{i,t} \end{aligned} \tag{5}$$

FREV is the one-quarter-ahead analyst forecast revision, defined as the difference between the median forecast for quarter t+1 issued after and before the quarter t earnings announcement, scaled by the stock price two days prior to the announcement. ΔDispersion is the change in the standard deviation of forecasts for the same quarter, calculated as the difference between post-call and pre-call dispersion, scaled by the stock price at the beginning of the quarter. For both measures, I use the first (post-event) and last (pre-event) forecasts issued by I/B/E/S analysts within a 90-day

[7] All estimates are qualitatively robust to inclusion of time fixed effects.

window surrounding the earnings announcement or conference call date. I included contemporaneous market reaction, CAR(0,1), as a control to proxy for other earnings information. Other controls are the same as Eq, (1)

Table 8 presents the estimates of Eq.(5). The results are in agreement with the predictions, suggesting analysts react to non verbal hesitation cues in the earnings calls. On average, a one standard deviation increase in manager response time is associated with 5% downward revision in analyst forecasts, along with a 4% increase in analyst forecast dispersion. The analysts may be finding the information conveyed by the managers during the call less reliable, possibly due to perceived hesitation. Consequently, they might be relying more on other personal information sources in their forecasts, which could lead to an increase in dispersion. This result also confirms the findings of De la Parra and Gallemore (2024).

[Table 8 about here.]

Now that I have shown how managers' hesitation is processed differently by investors at different rates for high- and low-scrutiny firms, in order to understand how the analysts incorporate this information under varying levels of scrutiny. This setting also allows for testing information uncertainty theory. I anticipate a negative correlation between forecast revisions for high-scrutiny firms, as any hesitation on top of negative news would be perceived as even more concerning. As for forecast dispersion, I anticipate a positive correlation between low scrutiny response time and analyst dispersion. This is because a perceived hesitation in otherwise well-performing firms could lead to confusion among analysts.

I test this predictions by estimation the modified specification below:

$$
\begin{aligned}
Y_{i,t} = {} & \beta_0 RT^{HS}_{i,t} + \beta_1 RT^{HS}_{i,t} + \beta_2 CAR(0,1) + \gamma' \text{FirmControls}_{i,t} \\
& + \theta' \text{ExecControls}_{i,t} + \phi' \text{AudioControls}_{i,t} \\
& + \delta' \text{TextControls}_{i,t} + \alpha_i + \varepsilon_{i,t}
\end{aligned}
\tag{6}
$$

All variables in Eq. (6) are the same as Eq. (5) except for partitioned RT. The results are reported in Table 9. The coefficients of managerial response time are negative and statistically significant for both high- and low-scrutiny firms. This result confirms the prediction of further downward revision for high-scrutiny firms. Moreover, analysts also recognize the hesitation in providing information and revise their forecasts even for low-scrutiny firms. The results of dispersion analysis also confirms the prediction. This result might be attributed to analysts' perceived hesitation in valuing the messages conveyed by managers, leading them to rely more on their idiosyncratic information (De la Parra and Gallemore (2024)).

[Table 9 about here.]

6 Conclusion

In this study, I provide empirical evidence that silence, measured by managerial response time, serves as a meaningful information channel. The results indicate that both investors and analysts respond to perceived non-vocal cues in the quarterly earning calls.

Longer response times are associated with negative contemporaneous and one-quarter ahead cumulative abnormal returns, indicating a negative investor perception of delayed responses. This association is particularly pronounced when response times are abnormally long. Additionally, I present evidence suggesting that analysts react negatively to prolonged response times by revising one-quarter ahead earning forecasts downward. When faced with conflicting signals, market reaction slows down, and analyst forecast dispersion increases. Notably, I fail to find any association between future earning surprises and managerial response time. This result implies that analysts efficiently incorporate hesitation information into their forecast revisions.

This study contributes to several strands of finance and accounting literature. First, the results presented provide evidence supporting the information uncertainty explanation of the PEAD anomaly. The findings suggest that the market reacts swiftly when negative earnings news is reinforced by hesitation, but it responds with a delay when conflicting signals emerge (e.g., positive earnings but a hesitant manager), leading to a drift. Second, the analysis contributes to the information asymmetry literature by demonstrating that earning conference calls can help reduce information asymmetry. Third, the study highlights how subtle, non-verbal cues are perceived and incorporated into the decisions of investors and analysts, contributing to behavioral finance. Finally, the study presents a novel application of artificial intelligence tools, such as speaker diarization, in finance research.

This study is among the first to provide evidence that silence, measured by response time, can provide incremental information. This finding paves the way for innovative research approaches that leverage unconventional data sources.

References

Anguera, X., Bozonnet, S., Evans, N., Fredouille, C., Friedland, G., and Vinyals, O. (2012). Speaker diarization: A review of recent research. *IEEE Transactions on Audio, Speech, and Language Processing*, 20(2):356–370.

Baik, B., Kim, A. G., Kim, D. S., and Yoon, S. (2024). Vocal delivery quality in earnings conference calls. *Journal of Accounting and Economics*, page 101763.

Ball, R. and Brown, P. (1968). An empirical evaluation of accounting income numbers. *Journal of Accounting Research*, 6(2):159–178.

Bernard, V. L. and Thomas, J. K. (1989). Post-earnings-announcement drift: delayed price response or risk premium? *Journal of Accounting Research*, 27:1–36.

Bernard, V. L. and Thomas, J. K. (1990). Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. *Journal of Accounting and Economics*, 13(4):305–340.

Bochkay, K., Hales, J., and Chava, S. (2020). Hyperbole or reality? investor response to extreme language in earnings conference calls. *The Accounting Review*, 95(2):31–60.

Bowen, R. M., Davis, A. K., and Matsumoto, D. A. (2002). Do conference calls affect analysts' forecasts? *The Accounting Review*, 77(2):285–316.

Brav, A. and Heaton, J. B. (2002). Competing theories of financial anomalies. *The Review of Financial Studies*, 15(2):575–606.

Brown, S., Hillegeist, S. A., and Lo, K. (2004). Conference calls and information asymmetry. *Journal of Accounting and Economics*, 37(3):343–366.

Burgoon, J., Mayew, W. J., Giboney, J. S., Elkins, A. C., Moffitt, K., Dorn, B., Byrd, M., and Spitzley, L. (2016). Which spoken language markers identify deception in high-stakes settings? evidence from earnings conference calls. *Journal of Language and Social Psychology*, 35(2):123–157.

Bushee, B. J., Gow, I. D., and Taylor, D. J. (2018). Linguistic complexity in firm disclosures: Obfuscation or information? *Journal of Accounting Research*, 56(1):85–121.

Call, A. C., Flam, R. W., Lee, J. A., and Sharp, N. Y. (2024). Managers' use of humor on public earnings conference calls. *Review of Accounting Studies*, 29(3):2650–2687.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, 52(1):57–82.

Chen, J. V., Nagar, V., and Schoenfeld, J. (2018). Manager-analyst conversations in earnings conference calls. *Review of Accounting Studies*, 23:1315–1354.

Collins, D. W. and Kothari, S. P. (1989). An analysis of intertemporal and cross-sectional determinants of earnings response coefficients. *Journal of Accounting and Economics*, 11(2-3):143–181.

Davis, A. K., Ge, W., Matsumoto, D., and Zhang, J. L. (2015). The effect of manager-specific optimism on the tone of earnings conference calls. *Review of Accounting Studies*, 20:639–673.

Davis, A. K., Piger, J. M., and Sedor, L. M. (2012). Beyond the numbers: Measuring the information content of earnings press release language. *Contemporary Accounting Research*, 29(3):845–868.

De Amicis, C., Falconieri, S., and Tastan, M. (2021). Sentiment analysis and gender differences in earnings conference calls. *Journal of Corporate Finance*, 71:101809.

De Bondt, W. F. and Thaler, R. (1985). Does the stock market overreact? *The Journal of Finance*, 40(3):793–805.

De la Parra, D. and Gallemore, J. (2024). The sound of uncertainty: Examining managerial acoustic uncertainty in conference calls. *Available at SSRN 4883575*.

Eyben, F., Wöllmer, M., and Schuller, B. (2010). Opensmile: the munich versatile and fast open-source audio feature extractor. In *Proceedings of the 18th ACM International Conference on Multimedia*, pages 1459–1462.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56.

Feldman, R., Govindaraj, S., Livnat, J., and Segal, B. (2010). Management's tone change, post earnings announcement drift and accruals. *Review of Accounting Studies*, 15:915–953.

Frankel, R., Johnson, M., and Skinner, D. J. (1999). An empirical examination of conference calls as a voluntary disclosure medium. *Journal of Accounting Research*, 37(1):133–150.

Fujita, Y., Kanda, N., Horiguchi, S., Nagamatsu, K., and Watanabe, S. (2019). End-to-end neural speaker diarization with permutation-free objectives. *arXiv preprint arXiv:1909.05952*.

Fuller, B. F., Horii, Y., and Conner, D. A. (1992). Validity and reliability of nonverbal voice measures as indicators of stressor-provoked anxiety. *Research in Nursing & Health*, 15(5):379–389.

Gish, H., Siu, M.-H., and Rohlicek, J. R. (1991). Segregation of speakers for speech recognition and speaker identification. In *ICASSP*, volume 91, pages 873–876.

Goberman, A. M., Hughes, S., and Haydock, T. (2011). Acoustic characteristics of public speaking: Anxiety and practice effects. *Speech Communication*, 53(6):867–876.

Goldstein, I. and Yang, L. (2017). Information disclosure in financial markets. *Annual Review of Financial Economics*, 9(1):101–125.

Henderson, A., Goldman-Eisler, F., and Skarbek, A. (1966). Sequential temporal patterns in spontaneous speech. *Language and Speech*, 9(4):207–216.

Hollander, S., Pronk, M., and Roelofsen, E. (2010). Does silence speak? An empirical analysis of disclosure choices during conference calls. *Journal of Accounting Research*, 48(3):531–563.

Lee, J. (2016). Can investors detect managers' lack of spontaneity? Adherence to predetermined scripts during earnings conference calls. *The Accounting Review*, 91(1):229–250.

Lewellen, J. and Shanken, J. (2002). Learning, asset-pricing tests, and market efficiency. *The Journal of Finance*, 57(3):1113–1145.

Liang, L. (2003). Post-earnings announcement drift and market participants' information processing biases. *Review of Accounting Studies*, 8(2):321–345.

Loughran, T. and McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, 66(1):35–65.

Malmendier, U. and Tate, G. (2005). CEO overconfidence and corporate investment. *The Journal of Finance*, 60(6):2661–2700.

Mayew, W. J. (2008). Evidence of management discrimination among analysts during earnings conference calls. *Journal of Accounting Research*, 46(3):627–659.

Mayew, W. J. and Venkatachalam, M. (2012). The power of voice: Managerial affective states and future firm performance. *The Journal of Finance*, 67(1):1–43.

Park, T. J., Kanda, N., Dimitriadis, D., Han, K. J., Watanabe, S., and Narayanan, S. (2022). A review of speaker diarization: Recent advances with deep learning. *Computer Speech & Language*, 72:101317.

Price, S. M., Doran, J. S., Peterson, D. R., and Bliss, B. A. (2012). Earnings conference calls and stock returns: The incremental informativeness of textual tone. *Journal of Banking & Finance*, 36(4):992–1011.

Qin, Y. and Yang, Y. (2019). What you say and how you say it matters: Predicting stock volatility using verbal and vocal cues. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 390–401.

Sawhney, R., Aggarwal, A., and Shah, R. (2021). An empirical investigation of bias in the multimodal analysis of financial earnings calls. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 3751–3757.

Sawhney, R., Khanna, P., Aggarwal, A., Jain, T., Mathur, P., and Shah, R. (2020). Voltage: Volatility forecasting via text audio fusion with graph convolution networks for earnings calls. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 8001–8013.

Sell, G., Snyder, D., McCree, A., Garcia-Romero, D., Villalba, J., Maciejewski, M., Manohar, V., Dehak, N., Povey, D., Watanabe, S., et al. (2018). Diarization is hard: Some experiences and lessons learned for the jhu team in the inaugural dihard challenge. In *Interspeech*, pages 2808–2812.

Shum, S., Dehak, N., Chuangsuwanich, E., Reynolds, D. A., and Glass, J. R. (2011). Exploiting intra-conversation variability for speaker diarization. In *Interspeech*, volume 11, pages 945–948.

Snyder, D., Garcia-Romero, D., Sell, G., Povey, D., and Khudanpur, S. (2018). X-vectors: Robust dnn embeddings for speaker recognition. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 5329–5333. IEEE.

Tasker, S. C. (1998). Bridging the information gap: Quarterly conference calls as a medium for voluntary disclosure. *Review of Accounting Studies*, 3:137–167.

Teixeira, J. P., Oliveira, C., and Lopes, C. (2013). Vocal acoustic analysis–jitter, shimmer and hnr parameters. *Procedia Technology*, 9:1112–1122.

Tranter, S. E. and Reynolds, D. A. (2006). An overview of automatic speaker diarization systems. *IEEE Transactions on Audio, Speech, and Language Processing*, 14(5):1557–1565.

Yang, Y., Qin, Y., Fan, Y., and Zhang, Z. (2023). Unlocking the power of voice for financial risk prediction: A theory-driven deep learning design approach. *MIS Quarterly*, 47(1).

Zhang, X. F. (2006). Information uncertainty and stock returns. *The Journal of Finance*, 61(1):105–137.

Ziano, I. and Wang, D. (2021). Slow lies: Response delays promote perceptions of insincerity. *Journal of Personality and Social Psychology*, 120(6):1457.

```
SPEAKER waveform 1 37.797 2.329 <NA> <NA> SPEAKER_03 <NA> <NA>
SPEAKER waveform 1 40.582 1.316 <NA> <NA> SPEAKER_03 <NA> <NA>
SPEAKER waveform 1 42.252 2.396 <NA> <NA> SPEAKER_03 <NA> <NA>
SPEAKER waveform 1 46.353 5.771 <NA> <NA> SPEAKER_01 <NA> <NA>
SPEAKER waveform 1 52.748 4.084 <NA> <NA> SPEAKER_01 <NA> <NA>
SPEAKER waveform 1 57.355 2.244 <NA> <NA> SPEAKER_01 <NA> <NA>
SPEAKER waveform 1 60.055 2.481 <NA> <NA> SPEAKER_01 <NA> <NA>
SPEAKER waveform 1 62.688 2.312 <NA> <NA> SPEAKER_01 <NA> <NA>
SPEAKER waveform 1 65.489 3.983 <NA> <NA> SPEAKER_01 <NA> <NA>
```

Figure 1: Excerpt from an RTTM file

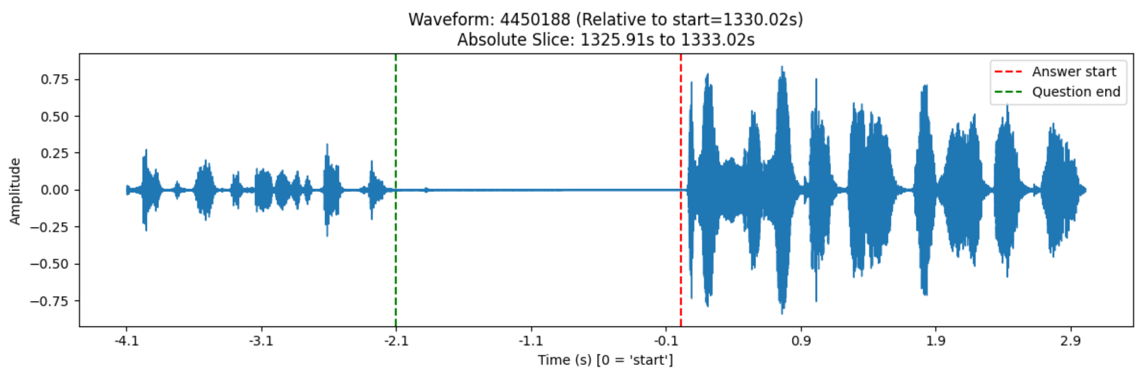

Figure 2: Example of detected question-end and answer-start timestamps overlaid on the earnings call audio waveform.

Table 1: Summary Statistics

This table reports summary statistics for the main variables used in the analysis. The sample covers 7,035 earnings calls from 471 S&P 500 firms between 2019 and 2023. Variables are winsorized at the 1st and 99th percentiles. Variable definitions are provided in Appendix A.1.

	N	Mean	SD	Median	25%	75%
Main Variables						
RT	7,035	1.599	0.504	1.536	1.248	1.889
RT^{HS}	7,035	0.325	0.693	0.000	0.000	0.000
RT^{LS}	7,035	1.274	0.775	1.384	0.937	1.772
CAR(0,1)	7,035	0.147	5.406	0.198	-3.113	3.378
CAR(2,90)	7,035	0.389	15.512	0.290	-9.265	9.812
CAR(2,180)	7,035	0.697	21.228	0.140	-12.578	13.201
FREV	6,816	0.660	1.816	0.004	-0.070	0.203
$\Delta Dispersion$	6,811	-0.244	0.720	-0.019	-0.152	0.008
Firm Controls						
Size	7,035	10.123	1.280	10.019	9.246	10.874
Book-to-Market	7,033	0.348	0.325	0.263	0.115	0.506
DIV	7,035	0.779	0.415	1.000	1.000	1.000
UE_t	6,898	0.102	0.345	0.059	0.005	0.175
Momentum	7,024	1.967	26.200	1.219	-15.442	17.494
Return Vol.	7,022	2.101	0.934	1.843	1.456	2.451
Executive Controls						
Age	7,029	59.538	5.663	60.000	56.000	63.000
Tenure	7,029	3.201	9.234	0.000	0.000	0.000
Overconfidence	7,029	0.379	0.485	0.000	0.000	1.000
Text Controls						
ToneMng	7,035	3.782	2.567	3.140	2.109	4.655
ToneAnalyst	6,881	2.332	2.346	1.667	1.000	2.800
FogMng	7,035	12.103	1.393	12.086	11.174	12.983
FogAnalyst	7,035	11.338	1.276	11.271	10.470	12.119
#WordsAnalyst	7,034	15.859	4.891	15.573	12.380	18.915
Log#Q	7,035	2.951	0.403	2.996	2.708	3.219
Log#Analyst	7,035	2.333	0.319	2.398	2.197	2.565
Pitch	7,028	0.653	0.738	0.521	0.387	0.712
Audio Controls						
StDev Pitch	7,028	0.005	0.006	0.004	0.003	0.006
Pitch Slope +	7,028	3.039	3.817	2.385	1.599	3.443
Pitch Slope -	7,028	7.017	8.816	5.498	3.859	7.857
Loudness	7,028	0.016	0.020	0.012	0.008	0.018
Shimmer	7,028	0.038	0.044	0.030	0.022	0.041
Jitter	7,028	0.001	0.001	0.001	0.001	0.001
HNR	7,028	0.058	0.082	0.046	0.026	0.072
Tempo	7,028	0.063	0.077	0.050	0.036	0.069
Length	7,034	1.007	0.199	1.011	0.915	1.088

Table 2: Managerial Hesitation and Contemporaneous Cumulative Abnormal Returns

This table reports estimates from pooled OLS regressions examining the relationship between managerial response time (RT) and contemporaneous cumulative abnormal returns (CARs). The dependent variable is $CAR(0,1)$, defined as the cumulative abnormal return from day 0 to day 1 relative to the earnings call date, computed using Fama–French size and book-to-market sorted portfolios as benchmarks. Variable definitions are provided in Appendix A.1. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively, based on two-tailed tests.

	(1) CAR(0,1)	(2) CAR(0,1)
RT	-0.243*	-0.289**
	(0.137)	(0.130)
UE_t		2.066***
		(0.261)
Momentum		0.0433***
		(0.00312)
ToneMng		0.128***
		(0.0247)
ToneAnalyst		0.107***
		(0.0320)
N	7035	6721
Adj. R^2	0.000373	0.0845
Firm Controls	–	✓
Executive Controls	–	✓
Text Controls	–	✓
Audio Controls	–	✓

Table 3: Estimates of CAR(0,1) Regressions by Response Time Terciles

This table reports results from estimating Eq. 1 separately for each tercile of the distribution of median response times. Observations are sorted into terciles based on response time. The dependent variable is CAR(0,1), defined as the cumulative abnormal return from day 0 to day 1 relative to the earnings call date, computed using Fama–French size and book-to-market sorted portfolios as benchmarks. Variable definitions are provided in Appendix A.1. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively, based on two-tailed tests.

	Low	Mid	High
RT	-0.225	-0.156	-0.590**
	(0.335)	(0.317)	(0.250)
UE_t	2.191***	1.863***	2.167***
	(0.398)	(0.355)	(0.424)
Momentum	0.0366***	0.0515***	0.0425***
	(0.00483)	(0.00479)	(0.00506)
ToneMng	0.129***	0.0884**	0.175***
	(0.0443)	(0.0423)	(0.0431)
ToneAnalyst	0.0881	0.133*	0.0983**
	(0.0546)	(0.0678)	(0.0486)
N	2348	2239	2134
Adj. R^2	0.0738	0.100	0.0760
Firm Controls	✓	✓	✓
Executive Controls	✓	✓	✓
Text Controls	✓	✓	✓
Audio Controls	✓	✓	✓

Table 4: Managerial Hesitation and Future Returns

This table reports estimates from pooled OLS regressions examining the relationship between managerial response time (RT) and 3 and 6–month CARs. The dependent variable is CAR$(2, j)$, defined as the cumulative abnormal return from day 2 to day 90 or 180 relative to the earnings call date, computed using Fama–French size and book-to-market sorted portfolios as benchmarks. Variable definitions are provided in Appendix A.1. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively, based on two-tailed tests.

	CAR(2,90)		CAR(2,180)	
	(1)	(2)	(3)	(4)
RT	-0.659*	-0.519**	-0.483	-0.554
	(0.385)	(0.256)	(0.669)	(0.402)
UE_t		-3.383***		-3.840***
		(0.515)		(0.744)
Momentum		0.333***		0.446***
		(0.00490)		(0.00849)
ToneMng		-0.113**		0.0363
		(0.0564)		(0.0926)
ToneAnalyst		-0.0834		-0.202*
		(0.0711)		(0.105)
N	7035	6721	7035	6721
Adj. R^2	0.000316	0.313	-0.0000107	0.309
Firm Controls	–	✓	–	✓
Executive Controls	–	✓	–	✓
Text Controls	–	✓	–	✓
Audio Controls	–	✓	–	✓

Table 5: Relationship Between Managerial Hesitation and Stock Returns, Split by Firm Scrutiny

This table presents pooled OLS regressions of cumulative abnormal returns (CAR) on managerial response time (RT), estimated over various post-call horizons. Column (2) covers the two-day window surrounding the earnings call; columns (4) and (6) cover one- and two-quarter horizons. To examine heterogeneity, RT is split into two variables: RT^{HS} (high scrutiny) and RT^{LS} (low scrutiny). RT^{HS} equals RT when firms miss earnings expectations ($UE_t < 0$); RT^{LS} equals RT when expectations are met or exceeded ($UE_t \geq 0$); both are zero otherwise. Variable definitions are provided in Appendix A.1. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively, based on two-tailed tests.

	CAR(0,1)		CAR(2,90)		CAR(2,180)	
	(1)	(2)	(3)	(4)	(5)	(6)
RT^{HS}	-1.253***	-0.949***	-0.0187	0.385	-0.101	0.286
	(0.167)	(0.163)	(0.446)	(0.341)	(0.780)	(0.514)
RT^{LS}	0.117	-0.0279	-0.888**	-0.877***	-0.619	-0.886**
	(0.137)	(0.129)	(0.397)	(0.265)	(0.671)	(0.420)
UE_t		1.137***		-2.110***		-2.657***
		(0.277)		(0.584)		(0.835)
Momentum		0.0429***		0.333***		0.446***
		(0.00313)		(0.00493)		(0.00852)
ToneMng		0.114***		-0.0949*		0.0535
		(0.0250)		(0.0571)		(0.0922)
ToneAnalyst		0.0967***		-0.0693		-0.189*
		(0.0315)		(0.0710)		(0.104)
N	7035	6721	7035	6721	7035	6721
Adj. R^2	0.0300	0.0942	0.00163	0.315	0.000123	0.310
Firm Controls	–	✓	–	✓	–	✓
Executive Controls	–	✓	–	✓	–	✓
Text Controls	–	✓	–	✓	–	✓
Audio Controls	–	✓	–	✓	–	✓

Table 6: Estimation of relationship between 1 and 2 quarter ahead unexpected returns

This table reports estimates from panel regressions examining the relationship between managerial response time (RT) and future firm performance, measured by unexpected earnings (UE_t) over two quarters following the earnings call. The dependent variable is UE_{t+k} for $k = 1, 2$, defined as the difference between actual and forecasted earnings scaled by stock price. Variable definitions are provided in Appendix A.1. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively, based on two-tailed tests.

	(1) UE_{t+1}	(2) UE_{t+2}
RT	0.00661	-0.00466
	(0.00943)	(0.0120)
FREV	0.0104**	0.00763
	(0.00492)	(0.00507)
UE_t	0.254***	0.159***
	(0.0295)	(0.0302)
Momentum	0.00126***	0.00127***
	(0.000212)	(0.000236)
ToneMng	0.000409	-0.000477
	(0.00167)	(0.00213)
ToneAnalyst	-0.000912	-0.00116
	(0.00184)	(0.00192)
N	5526	5151
Adj R^2	0.115	0.0798
Firm FE	✓	✓
Firm Controls	✓	✓
Executive Controls	✓	✓
Text Controls	✓	✓
Audio Controls	✓	✓

Table 7: Relationship Between Managerial Hesitation and Contemporaneous CARs, Split by Institutional Ownership

This table reports results from estimating Eq. 1 separately for each tercile of the distribution of level of institutional ownership. Observations are sorted into terciles based on level of institutional ownership. The dependent variable is CAR$(0,1)$, defined as the cumulative abnormal return from day 0 to day 1 relative to the earnings call date, computed using Fama–French size and book-to-market sorted portfolios as benchmarks. Variable definitions are provided in Appendix A.1. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively, based on two-tailed tests.

	Low	Mid	High
RT	-0.358*	-0.173	-0.217
	(0.211)	(0.225)	(0.236)
UE_t	2.419***	2.251***	1.504***
	(0.384)	(0.384)	(0.427)
Momentum	0.0424***	0.0415***	0.0461***
	(0.00492)	(0.00515)	(0.00471)
ToneMng	0.108***	0.110**	0.160***
	(0.0393)	(0.0456)	(0.0412)
ToneAnalyst	0.129**	0.121**	0.0469
	(0.0523)	(0.0572)	(0.0616)
N	2395	2209	2001
Adj R^2	0.0866	0.0867	0.0708
Firm Controls	✓	✓	✓
Executive Controls	✓	✓	✓
Text Controls	✓	✓	✓
Audio Controls	✓	✓	✓

Table 8: Analyst Reaction to Manager Hesitation

This table reports estimates from firm fixed effects regressions examining the relationship between managerial response time (RT) and two analyst outcomes: one-quarter-ahead forecast revisions (FREV) and changes in forecast dispersion (ΔDispersion). Columns (1) and (2) reports results of analyst forecast revisions and columns (3) and (4) reports results for analyst forecast dispersion. Variable definitions are provided in Appendix A.1. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively, based on two-tailed tests.

	Revision	$\Delta Dispersion$
RT	-0.180***	0.0504**
	(0.0591)	(0.0226)
CAR(0,1)	0.00374	-0.00220
	(0.00461)	(0.00179)
UE_t	-0.302***	0.0871**
	(0.112)	(0.0422)
Return Vol.	-0.405***	0.0417***
	(0.0304)	(0.0135)
ToneMng	0.00682	0.00629
	(0.0104)	(0.00412)
N	6649	6515
Adj R^2	0.0559	0.0665
Firm FE	✓	✓
Firm Controls	✓	✓
Executive Controls	✓	✓
Text Controls	✓	✓
Audio Controls	✓	✓

Table 9: Relationship Between Managerial Hesitation and Analyst Behavior, Split by Firm Scrutiny

This table presents pooled OLS regressions of cumulative abnormal returns (CAR) on managerial response time (RT), estimated over various post-call horizons. Column (2) covers the two-day window surrounding the earnings call; columns (4) and (6) cover one- and two-quarter horizons. To examine heterogeneity, RT is split into two variables: RT^{HS} (high scrutiny) and RT^{LS} (low scrutiny). RT^{HS} equals RT when firms miss earnings expectations ($UE_t < 0$); RT^{LS} equals RT when expectations are met or exceeded ($UE_t \geq 0$); both are zero otherwise. Variable definitions are provided in Appendix A.1. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively, based on two-tailed tests.

	Revision	$\Delta Dispersion$
RT^{HS}	-0.163**	0.0344
	(0.0729)	(0.0259)
RT^{LS}	-0.185***	0.0555**
	(0.0587)	(0.0230)
CAR(0,1)	0.00400	-0.00244
	(0.00464)	(0.00181)
UE_t	-0.279**	0.0661
	(0.126)	(0.0467)
Return Vol.	-0.406***	0.0421***
	(0.0304)	(0.0135)
ToneMng	0.00725	0.00588
	(0.0104)	(0.00415)
N	6649	6515
Adj R^2	0.0558	0.0666
Firm FE	✓	✓
Firm Controls	✓	✓
Executive Controls	✓	✓
Text Controls	✓	✓
Audio Controls	✓	✓

A Appendix

A.1 Variable Definitions

Table A.1: Variable Definitions

Variable	Definition
Panel A: Main Hesitation and Return Variables	
RT	Median managerial response time (in seconds) between the end of analyst questions and the start of managerial answers during the Q&A session.
RT^{HS}	RT value for high-scrutiny firms ($UE_t < 0$), and zero otherwise.
RT^{LS}	RT value for low-scrutiny firms ($UE_t \geq 0$), and zero otherwise.
CAR(i, j)	Cumulative abnormal return from day i to day j relative to the earnings call date, benchmarked using Fama-French size and book-to-market portfolios.
Panel B: Analyst Behavior Variables	
FREV	One-quarter-ahead analyst forecast revision: the change in median earnings forecast for quarter $t + 1$ issued after vs. before the quarter t earnings announcement, scaled by the stock price two days before the announcement.

Continued on next page

Table A.1 – continued from previous page

Variable	Definition
ΔDispersion	Change in dispersion of analyst earnings forecasts for the same fiscal quarter, computed as the post-call forecast standard deviation minus the pre-call standard deviation, scaled by stock price at the beginning of the quarter.
Panel C: Firm Performance and Control Variables	
Size	Logarithm of total assets.
Book-to-Market	The ratio of the book value of shareholders' equity to the market value of equity at the end of the current quarter.
DIV	A dummy variable equal to 1 if the firm pays a dividend in the quarter, and 0 otherwise.
UE_t	Unexpected earnings: the difference between the actual I/B/E/S earnings per share and the I/B/E/S analyst summary consensus median earnings per share, scaled by the price per share two days before the conference call.
Momentum	Cumulative abnormal returns between -127 and -2 days before the earnings call.
Return Vol.	Standard deviation of daily stock returns between -127 and -2 days before the earnings call.
Panel D: Executive Control Variables	
Age	Age of the CEO.
Tenure	Tenure of the CEO in years.

Continued on next page

Table A.1 – continued from previous page

Variable	Definition
Overconfidence	Holder 67 dummy for manager overconfidence, based on Malmendier and Tate (2005).
Panel E: Textual Control Variables	
ToneMng	The ratio of positive to negative words uttered by managers, based on the Loughran and McDonald (2011) dictionary[8]
ToneAnalyst	The ratio of positive to negative words uttered by analysts, based on the Loughran and McDonald (2011) dictionary.
FogMng	Gunning Fog Index for the readability of manager speech.
FogAnalyst	Gunning Fog Index for the readability of analyst speech.
WordsAnalyst	The average number of words per analyst question.
LogQ	The natural logarithm of the total number of questions asked during the Q&A session.
LogAnalyst	The natural logarithm of the total number of analysts participating in the call.
Panel F: Acoustic Control Variables	
All acoustic variables are extracted using the openSMILE toolkit and are normalized by the length of the manager's response utterance.	
Pitch	Mean pitch (fundamental frequency, F0) of the manager's voice. (openSMILE: *F0semitoneFrom27.5Hz_sma3nz_amean*)

Continued on next page

[8]I exclude the words question(s) from negative words list following De la Parra and Gallemore (2024) as they are not used in a negative meaning in the context of earnings calls.

Table A.1 – continued from previous page

Variable	Definition
StDev Pitch	Standard deviation of the manager's pitch, measuring pitch variation. (openSMILE: *F0semitoneFrom27.5Hz_sma3nz_stddevNorm*)
Pitch Slope +	Measure of upward pitch slope, associated with arousal/excitement. (openSMILE: *F0semitoneFrom27.5Hz_sma3nz_meanRisingSlope*)
Pitch Slope -	Measure of downward pitch slope. (openSMILE: *0semitoneFrom27.5Hz_sma3nz_meanFallingSlope*)
Loudness	The acoustic intensity of the manager's voice. (openSMILE: *loudness_sma3_amean*)
Shimmer	Average instability in voice amplitude, a measure of voice perturbation. (openSMILE: *shimmerLocaldB_sma3nz_amean'*)
Jitter	Average instability in pitch, a measure of voice perturbation. (openSMILE: *jitterLocal_sma3nz_amean*)
HNR	Harmonic-to-Noise Ratio, a measure of voice quality. (openSMILE: *HNRdBACF_sma3nz_amean'*)
Tempo	Speech rate of the manager. (openSMILE: *VoicedSegmentsPerSec*)
Length	Length of the conference call in hours.

A.2 Accuracy of Automated Response Time Measurements

I employ an AI-based diarization tool in conjunction with earnings call transcripts to identify the precise moments when a question ends and an answer starts. To assess the accuracy of the automated process, I randomly selected a sample of one hundred question-answer pairs. Subsequently, I created audio clips that commence two seconds prior to the end of a question and three seconds after the start of an answer. Each audio clip has a total duration of 2 seconds, the calculated response time, and 3 seconds.

I then manually listened to these audio clips in Audacity.[9] I checked if the clip accurately captured the end of a question and the beginning of an answer, and I measured the response time by marking the end of the question and the start of the answer.

Out of 100 observations, 99 of the clips capture a question-answer pair. The correlation coefficient between manually and algorithmically measured response times is 95.58%. These results suggest that the algorithm used in this study is capable of accurately measuring managerial response times.

A.3 Robustness

[Table 10 about here.]

[9]Audacity® software is copyright © 1999-2025 Audacity Team. Website: https://audacityteam.org/. It is free software distributed under the terms of the GNU General Public License. The name Audacity® is a registered trademark.

Table A.2: Robustness Tests

This table present results of several robustness tests. The first column shows pooled OLS regressions of two–day cumulative abnormal returns (CAR) calculated using Fama and French (1993) and Carhart (1997) four-factor model on managerial response time (RT). Column (2) - (5) covers the results of short and long term CARs, analyst forecast revisions and change in analyst forecast dispersion with added firm and quarter fixed effects. Results are qualitatively robust. Variable definitions are provided in Appendix A.1. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively, based on two-tailed tests.

	Alternate CAR	Alternative Model			
	(1) CAR(0,1)	(2) CAR(0,1)	(3) CAR(2,90)	(4) Revision	(5) $\Delta Dispersion$
RT	-0.266**	-0.328*	-0.855**	-0.0686*	0.0284
	(0.123)	(0.188)	(0.413)	(0.0367)	(0.0203)
N	6719	6717	6717	6649	6515
Adj. R^2	0.0703	0.0927	0.333	0.629	0.352
Firm Controls	✓	✓	✓	✓	✓
Executive Controls	✓	✓	✓	✓	✓
Text Controls	✓	✓	✓	✓	✓
Audio Controls	✓	✓	✓	✓	✓
Firm FE	–	✓	✓	✓	✓
Quarter FE	–	✓	✓	✓	✓