"Let me get back to you" – A machine learning approach to measuring non-answers*

Andreas Barth† Sasan Mansouri‡ Fabian Woebbeking§

April 1, 2022

Abstract

Using a supervised machine learning framework on a large training set of questions and answers, we identify 1,364 trigrams that signal non-answers in earnings call Q&A. We show that this glossary has economic relevance by applying it to contemporaneous stock market reactions after earnings calls. Our findings suggest that obstructing the flow of information leads to significantly lower cumulative abnormal stock returns and higher implied volatility. As both our method and glossary are free of financial context, we believe that the measure is applicable to other fields with a Q&A setup outside the contextual domain of financial earnings conference calls.

Keywords: Econlinguistics, textual analysis, natural language processing, multinomial inverse regression, non-answers

JEL-Classification: D80, D82, G10, G14, G30.

*We thank Daniel Andrei, Doron Avramov, Turan Bali, Justin Birru, Beatriz García Osma, Alexander Hillert, Gerard Hoberg, Stephan Hollander, Michael Kirchler, Christian Leuz, Tim Loughran, Bill McDonald, Philippe Mueller, Michael Norton, Todd Rogers, Astrid Salzmann, Christian Schlag, Matt Taddy, Alexander Wagner, Jingjing Wang, Mark Wahrenburg, and Xiao-Jun Zhang for valuable comments and suggestions. We also thank Markus Bader and the department of *Psycholinguistics* at Goethe University Frankfurt for providing valuable background information. The paper was previously circulated under the working title 'Econlinguistics'.

†Saarland University Saarbruecken and Goethe University Frankfurt, andreas.barth@finance.uni-frankfurt.de

‡Goethe University Frankfurt, mansouri@finance.uni-frankfurt.de

§Goethe University Frankfurt, woebbeking@finance.uni-frankfurt.de

1 Introduction

"Senator, my — I can certainly have my team get back to you on any specifics there that I don't know, sitting here today." — Mark Zuckerberg, *US Senate Hearing, April 2018*

The asymmetric distribution of information is considered a key friction in economics. Different mechanisms aim to transfer information from the better-informed to the less informed agent, where a question and answer (Q&A) setting is the most targeted form of information exchange. Faced with a question, the addressee can respond in two ways. First, they can supply the requested information by faithfully answering the question, which requires having the specific knowledge within the context of the question.[1] Second, they can refuse to supply the requested information, which does not require a context-specific answer. While it is *relatively* easy for humans to detect whether a question has been answered or not, we teach this skill to a machine. We use a supervised machine learning framework on a large textual training set of 64,173 classified responses to questions to identify 1,364 trigrams that signal non-answers.

The Gricean norms in communication describe cooperative principles of how people achieve effective conversational communication (Grice, 1989). These principles state that effective communication (i) contains the appropriate quantity of information, (ii) is truthful, (iii) is delivered in an appropriate manner and (iv) is relevant to the topic at hand. Violating any of these principles results in 'deceptive' communication.

We use violations of Gricean norms to derive a metric that identifies the absence of requested information in an answer, i.e. non-answers in earnings call Q&A. The glossary is derived from financial markets, which are heavily characterized by and sensitive to asymmetric information. More precisely, we derive the glossary from a training set of earnings conference calls, where investors and analysts can directly question senior executives' during Q&A sessions.[2] We conduct several tests to investigate the economic relevance of the metric using a large validation set of earnings conference calls.

We document that markets react to non-answers and observe negative stock returns after calls on which management distinctly avoids answers. Moreover, we find larger implied volatilities after these calls, indicating higher investor uncertainty. Financial analysts, too, perceive non-answers as a negative signal. In particular, we find that analysts are less likely to revise their EPS forecasts upwards following a call with many non-

[1] Alternatively, the respondent could answer the question with a lie, which also requires knowing the specific context of the question. As lies typically receive heavy sanctions, this paper focuses on refusals to answer and not detecting lies.

[2] Despite being derived from a financial market context, the glossary is free of financial context. Therefore, we believe that the measure is applicable to other fields with a Q&A setup outside the contextual domain of financial earnings conference calls.

Q

Answer

Non-Answer

Cause:
Intention to withold information
Inability to provide information

Symptoms:
Rejecting, e.g., “I won‘t answer that.”
Blathering, i.e., “Beating around the bush”
Postponing, e.g., “We have to wait and see...”
. . .

Figure 1: **Anatomy of an answer.** The response to a question consists of either an attempt to provide the requested information with an appropriate answer or the lag thereof, i.e. a non-answer. Later, we will use the symptoms blathering and rejecting to classify non-answers in our training set and show that this is sufficient to capture additional symptoms, such as postponing. Non-answers have in common that they can be free of context, e.g. one may reject answering without even knowing the questions, which is an important distinction that allows for generalizing our glossary.

answers. We finally investigate to which questions managers provide non-answers and find that, within the same earnings call, managers avoid responding to arguably tougher questions, as indicated by negative-toned questions, and follow-up questions by the same analyst. Similarly, non-answers appear more frequent when questions are ‘forward looking’, hence, when managers’ may be less able to provide the requested information.

Conceptually, a question can be understood as an illocutionary act that attempts to extract information from its addressee. The addressee can respond in two ways, as outlined in Figure 1. First, they can supply the requested information by faithfully answering the question, which requires having the specific knowledge within the context of the question. Second, they can refuse to supply the requested information, which, in contrast to effective communication, does not depend on the context of the question. This paper refers to this potentially deceptive communication in answering as a non-answer.

Non-answers are characterized by different symptoms. The most obvious symptom is openly refusing to provide the requested information, as for example, Elon Musk, the CEO of Tesla Inc, did during an earnings call in May 2018.[3] A second symptom of refusing context-specific information is the more indirect and deceptive behavior of dodging a question or “blathering”, i.e. ‘beating around the bush’.[4]

[3]On the question of Sanford Bernstein’s analyst Toni Sacconaghi: “And so where specifically will you be in terms of capital requirements?”, Musk replied: “Excuse me. Next. Boring, bonehead questions are not cool. Next?”

[4]Although there may be other symptoms, we later show that a limited number of symptoms is sufficient to train the model.

We base our measure on the two symptoms *'rejecting'* and *'blathering'*, and show that we can construct a metric that quantifies non-answers from just a few symptoms by using a multinomial inverse regression (MNIR) (Taddy, 2013b).[5] The input for the MNIR are all Q&As of earnings conference calls for financial firms in the S&P 500 for the period 2002 - 2019 (the training set). For each of management's answers during these calls, we quantify the two symptoms rejecting, as described in Gow et al. (2021), and blathering, as outlined in Barth et al. (2021). MNIR, as a supervised generative model, then maps the high-dimensional choice set of available trigrams into the two observable attributes in the classified training set.

This procedure results in a glossary of 1,364 trigrams that deduce a scoring metric for non-answers. As the glossary is determined by a machine learning algorithm, it comes with several desired features: first, it reflects not only the attribute measures classified in the training set, but captures a non-answer across several additional symptoms by association. Second, this approach reduces the subjectivity of a human interpreter (Loughran and McDonald, 2020). Third, the trigrams in the glossary are neither industry- nor context-specific. Thus, we believe that it can be applied to any economic sector or other Q&A setting, such as senate hearings, interviews with politicians, or press conferences by central banks. Furthermore, comparing the non-answer metric with other linguistic sentiment measures yields only very weak correlations and, hence, it serves as a new dimension in textual analysis.

We conduct various tests to document the plausibility of the glossary. To do so, we collect the earnings conference calls of non-financial companies in the S&P 500, i.e. firms that are not part of the training set, for the period 2002 - 2019 as a validation set. Earnings conference calls provide an ideal setting to test our glossary: the listener is more likely to detect whether a question has been answered when their attention is diverted from social goals (Rogers and Norton, 2011), and thus, we expect an immediate market reaction.

We show that not answering analysts' questions leads on average to significantly lower abnormal stock returns following the conference call. We also link our metric to option implied volatilities after earnings conference calls and show that investor uncertainty is greater if the requested information is not provided, i.e. investors are willing to pay more for insurance against adverse stock price movements. Both the stock price reaction as well as higher implied volatility suggest that the non-answer score measures an obstruction of information flow, which retards the reduction of information asymmetries between the

[5] While the usage of multinomial text regressions is novel in finance literature, it has been applied in political science studies to derive the subject-specific document sentiment in political posts (Taddy, 2013b) or to measure time trends in partisanship in congressional speeches (Gentzkow et al., 2019).

management and investors. We validate that the glossary is not a product of sheer randomness in a Monte Carlo simulation, where we draw 1,000 randomly selected dictionaries with 1,364 trigrams from all words that appear at least once in our training set. We repeat the textual analysis for each of these random draws, derive a corresponding placebo non-answer score and test for an effect of this score on cumulative abnormal returns. It turns out that it would be extremely unlikely to produce economically significant results by randomly drawing a glossary from the universe of trigrams.

Finally, we document how firms subtly control information flow (Cohen et al., 2020): during an earnings call, managers try to avoid answering tougher and more critical questions. The non-answer score is higher for managements' responses to follow-up questions by the same analyst, i.e. when the analyst asks a (typically more drilling) clarification question, as well as for managements' responses to more negative questions. Furthermore, questions with forward looking sentences, i.e., questions that refer to (potentially unknown) future outcomes, are more likely to receive a non-answer.

The remainder of this paper is organized as follows: Section 2 reviews the relevant literature. In Section 3, we describe in detail how we generate our novel glossary. We describe several tests for the validity of the word list in Section 4. Section 5 concludes.

2 Background and literature

Textual analysis is a versatile tool in finance and accounting that transforms qualitative information into quantitative measures. One common approach for quantifying language is word categorization (bag of word / dictionary approach). For example, the Harvard-IV and Lasswell dictionaries, which are part of Harvard General Inquirer Word Lists, consist of word lists about many psychological and sociological topics.[6]

To overcome issues with noise from general dictionaries, researchers have introduced finance-specific dictionaries to measure the tone of financial reports. Henry (2008), for example, published one such list for the telecommunications and computer services industries. Loughran and McDonald (2011) produced other widely recognized word lists that were extracted from 10-K reports to measure inter alia positive and negative tone, and most recently a word list to measure firm complexity (Loughran and McDonald, 2019). Harvey (2016) also created a glossary of factual finance terminology, which Loughran and McDonald (2014) use, for example, to develop a measure of financial readability of 10-K reports. We extend this literature by providing a novel glossary that quantifies the informational content of a response to a question, and thus, a new dimension of quantifiable

[6]For more details on the different available dictionaries and the list of words, see: `http://www.wjh.harvard.edu/~inquirer/homecat.htm`

language that can be used in several contexts.

Second, we add to the interdisciplinary literature on the precise and efficient transfer of information, or lack thereof. In linguistics, for example, many studies describe how to achieve effective conversational communication, with the cooperative principles by Grice (1989) being some of the most influential content. Building on this work, several studies show how listeners perceive a violation of the cooperative principles, or in which situations a listener is more likely to detect deceptive communication.[7]

The effects of precise information sharing are also of particular interest in the field of economics, where obscuring information is associated with, e.g., lower stock returns, lower earnings, and higher risks.[8] The related studies use various proxies for the characterization of imprecise information, such as vague communication measured by the frequency of words such as "vague" and "uncertainty" in management statements (Loughran and McDonald, 2011; Dzielinski et al., 2021), the ratio of numeric to textual content in earnings conference calls (Zhou, 2018), readability of 10-K reports measured by the popular Gunning (1952) "Fog-Index" of linguistic complexity (Li, 2008; Bloomfield, 2008),[9] calling on bullish analysts in conference calls (Cohen et al., 2020), or managers reading from prepared scripts when responding to questions during earnings conference calls (Lee, 2015).

Most closely related to our work in the interdisciplinary literature is Clayman (1993), who show that evading a question is frequently characterized by the response practice to reformulate the question. In the field of economics, our work is closest to Barth et al. (2021), Hollander et al. (2010) and Gow et al. (2021). The latter two papers measure withholding information in the most direct sense by manually reviewing call transcripts to deduct regular expressions that identify answers such as '*No, we do not want to provide*

[7]See, for example, Buller and Burgoon (1996), who state in their Interpersonal Deception Theory that the process and outcome of interpersonal deception is grounded within a conversational context and an interpersonal relationship. McCornack et al. (1992) show that perceived message deceptiveness and perceived message competence are significantly influenced by the manipulations of amount, veracity, relevance, and clarity of information. The paper by Rogers and Norton (2011) shows that deception is more likely to be detected when listeners' attention was to determine the relevance of the speakers' answers, i.e. if it was diverted from social goals. It is also shown that speakers were more negatively rated once their deception was detected.

[8]There also is evidence for an amplifying effect arising from precise information sharing, as the latter positively correlates with media coverage (Mansouri, 2021).

[9]Based on this work, several studies offer different measures of linguistic complexity (Loughran and McDonald, 2014; Bonsall et al., 2017) or provide a rationale for using complex language. For example, Bushee et al. (2018) analyze the linguistic complexity of Q&As in earnings calls and argue that the source of complexity can be composed into its latent components *obfuscation* and *information*. Specifically, they argue that complex responses to complex questions should be understood as information, whereas complex responses to simpler questions should be understood as obfuscation. In line with that, Guay et al. (2016) show that managers employ voluntary disclosures as a tool to mitigate the negative impact of their complex financial statements.

that information', while Barth et al. (2021) investigate whether managers' avoidance of context specific language (jargon) is perceived by market participants as retarding the information flow in earnings conference calls.

Our glossary is a significant step towards the general identification of non-answers. Our novel approach starts by classifying a training set with different symptoms for non-answers, including symptoms laid out by the preceding literature.[10] The glossary is then determined by a machine learning algorithm, thereby reducing the subjectivity of a human interpreter (Loughran and McDonald, 2020). The resulting glossary reflects not only one very specific symptom, namely 'rejecting' or 'blathering', but captures a non-answer across several symptoms (see Figure 1) and in a much broader sense. We show that this identification is significantly more powerful than previous attempts in the literature. Moreover, the approach in Barth et al. (2021) is only valid for financial firms, while our glossary is neither specific to an industry nor to a specific context, but can be applied to any Q&A setting. Ultimately, our non-answer metric is only very weakly correlated with other linguistic sentiment measures and, hence, serves as a new dimension in textual analysis.

3 The glossary

A machine readable version of the glossary is available at econlinguistics.org

---

The glossary contains trigrams that are markers for non-answers.[11] This section lays out how these trigrams were identified from a training set of questions and answers.

As outlined in Figure 1, when faced with a question, the addressee can respond using effective communication and supply the requested information, or they can be deceptive by violating any of the four Gricean cooperation principles (Grice, 1989). When developing the glossary, we focus on the first and the fourth Gricean maxims, i.e., whether the respondent provides any information at all and whether their answer provides factual content that is relevant for the topic at hand.

[10] When applying the regular expressions for rejecting an answer of Gow et al. (2021) to the analysis of economic relevance in Section 4, we find only very weak evidence of abnormal stock returns following an earnings call and obtain hardly any variation in the distribution for any Q&A setting outside the financial domain.

[11] A trigram is a continuous sequence of three elements from a text. The literature on natural language processing shows a significant improvement in modeling language with trigrams compared to unigrams (Dave et al., 2003; Bekkerman and Allan, 2004). While higher-order n-grams better capture the sentiments of expressions, they come at the cost of lower coverage in the data (Pak and Paroubek, 2010). For many years, trigrams have been a favorite model choice, as they can simultaneously reflect the syntax and the pragmatics of the text domain (Jelinek, 1991).

To understand general factuality in a linguistic sense, one would need to understand the context of the question and the expected information gain. The supervised machine learning approach, however, can derive a glossary that is context-independent. Thus, we do not approach factuality from a context and audience specific perspective but rather focus on vocabulary that indicates the intention not to respond to a question in a broader, more general sense.

3.1 Multinomial Inverse Regression

Multinomial inverse regression is a supervised generative model developed by Taddy (2013b, 2015) that allows for mapping a high-dimensional choice set of words within a text into an observable attribute. A text is defined as a combination of several tokens, where a token is a single word or a combination of n words (n-gram). For a given tokenization, a document i in the universe of all available documents $\mathcal{I}$ is represented by a sparse vector of token counts $\mathbf{x_i} = [x_{i1}, ..., x_{ip}]'$ and frequencies $\mathbf{f_i} = \mathbf{x_i}/m_i$, where $m_i = \sum_{j=1}^{p} x_{ij}$, for all available tokens p in $\mathcal{I}$.

A naive approach would be to fit a linear regression model of the attribute measure on the token counts,

$$\mathbf{y} = \beta \mathbf{x}^{\intercal} + \epsilon,$$

where the factor loading β represents each token's contribution to the attribute measure. However, as the choice set of tokens within a text and thus, the dimension of $\mathbf{x}$, is usually quite large, a normal regression cannot provide an appropriate estimate of the conditional distribution of $\mathbf{y}$.

To shrink dimensionality in pursuit of a parsimonious model, we turn to a least absolute shrinkage and selection operator (Lasso) regression type of model (also known as L1-regularization), see e.g. Hastie et al. (2009). A special case of Lasso, which builds on the pioneering work of Cook et al. (2007), is the multinomial inverse regression (MNIR) that has been developed in Taddy (2013b, 2015). MNIR, as a 'Gamma-Lasso' scheme, applies inverse regressions, "wherein the inverse conditional distribution for text given sentiment is used to obtain low-dimensional document scores that preserve information relevant to y." This methodology, which was specifically designed for textual analyses, produces significant computational improvements over the classical Lasso approach.

As described in Taddy (2013b), a basic MNIR model is given by

$$\begin{aligned}
\mathbf{x_y} &\sim MN(\mathbf{q_y}, m_y), \quad \text{with} \\
\mathbf{x_y} &= \sum_{i:y_i=y} \mathbf{x_i}, \\
m_y &= \sum_{i:y_i=y} m_i, \\
q_{yj} &= \frac{exp[\alpha_j + y\phi_j]}{\sum_{k=1}^{p} exp[\alpha_k + y\phi_k]} \\
\text{for} \;\; j &= 1, ..., p, y \in \mathcal{Y} \;\; \text{and} \;\; m_i = \sum_{j=1}^{p} x_{ij}.
\end{aligned}$$

Each MN is a p-dimension multinomial distribution of size m and probabilities $\mathbf{q}$ that are a linear function of y through a logistic link with token loadings ϕ.

The parameters ϕ of the model, i.e. , each token's contribution to the attribute measure, can be fitted via maximum a posteriori (MAP) estimation. While a classical Lasso estimation can be interpreted as a MAP estimation with independent and identical Laplace priors for the regression parameters, Taddy (2013b) use independent Gamma-Laplace priors to fit the MNIR model.[12]

In this paper, we focus on the refined estimation procedure proposed in Taddy (2015), which is available through the *textir* package in R. In contrast to the procedure in Taddy (2013b), which requires explicit shape and rate hyperpriors for the Gamma distribution, the modified approach in Taddy (2015) only has one relevant parameter, the *"Gamma-Lasso weight"* γ. We follow the author and set $\gamma = 1$; however, we find that alternative specifications (including $\gamma = 0$ as in classical Lasso) do not change our overall results. Ultimately, our results do not hinge on the regularization model, we simply follow Taddy (2015) because it is computationally very efficient, allows more flexibility for the sentiment input (higher number of categories), and requires the fewest assumptions.

The response factor does not have to be a single attribute measure y_i, but MNIR can be generalized to support K-dimensional response factors $\mathbf{v_i}$, in which case the multinomial model collapses to the levels of $\mathbf{x_v}$. In our case, using several response factors has the benefit that we can employ measures for different violations of the Gricean norms of effective conversational communication (Grice, 1989).

We use measures for the violation of the first and fourth Gricean maxim as response

[12]See Taddy (2013a) for a comparison of their Gamma-Lasso framework with standard Lasso and alternatives with concave penalization.

factors.[13] A violation of the first Gricean norm, the lack of appropriate quantity of information, is derived in the spirit of Gow et al. (2021), using regular expressions that identify answers with a direct rejection to provide the requested information. The violation of the fourth Gricean maxim, the relevance of the response to the topic at hand, is proxied with the metric for blathering introduced in Barth et al. (2021).

By using these response factors, the resulting list of trigrams and their corresponding weights indicate the degree to which a given trigram predicts a violation of the effective conversational communication principles of Grice (1989) in any response to a question.

3.2 Training set

The glossary is extracted from textual data that originates from earnings conference calls. These calls offer a relatively standardized Q&A format in a controlled contextual environment, have an economically relevant impact and are available in regular intervals and large numbers. Earnings calls do not have an identical structure, yet they often follow a similar pattern: first, the management, typically the CEO or CFO, presents the latest financial results and earnings outlooks in a speech that is usually prepared by the investor relations department. A question and answer session between the management and financial analysts then follows this presentation.

We collected every transcript of earnings calls held by companies listed in the S&P 500 index available from Thomson Reuters' StreetEvents for the period 2002 - 2019. These calls are released quarterly and usually take place on the same day as the corresponding earnings release.[14] As we focus on a question and answer setup as a specific form of communication rather than the prepared presentation, we exclude all earnings calls without a Q&A session. We also restrict our sample to managements' responses that contain at least five words to mitigate any bias in our attribute measures.[15]

The full sample of earnings calls is divided into a training set $\mathcal{I}$, where we have a clean measure for both response factors, and a validation set that *validates* our glossary by showing its economic relevance. We split the sample across industries and use all

[13]As described below, we derive the glossary based on Q&As in financial markets. As lies typically receive heavy sanctions, and answers by management were usually delivered in an appropriate manner, we abstract from measures of violations of the second and third Gricean norms.

[14]92.8% of all calls in our sample take place on the same day as the earnings release and 6.9% take place one day after the earnings announcements. In only six cases is the call scheduled for more than one day after the earnings announcement.

[15]Shorter answers need to be removed because of the high signal-to-noise ratio for the attribute measure blathering. For example, sentences like "This is correct" result in a blathering score of 1 and would thus appear with a high probability in the non-answer glossary. These sentences, however, are obviously very precise answers.

financial firms in the training set.[16] More precisely, we use each single answer given by a financial firm's management in response to an analyst's question as an observation in the training set $\mathcal{I}$. We derive our two response factors for these answers.

As a first attribute measure, we use regular expressions to identify the rejections according to Gow et al. (2021). Rejections can take several forms, such as the refusal to provide the requested information ("we do not provide this disclosure") or the inability to provide the requested information ("I do not know"). We flag these answers with a dummy $y_{ijt}^{\text{Rejecting}}$ that equals 1 if the response j in an earnings call of company i at time t contains any rejection phrase, i.e.

$$y_{ijt}^{\text{Rejecting}} = \begin{cases} 1 & \text{, if rejection phrase} \in \text{response } j \\ 0 & \text{, else.} \end{cases} \tag{1}$$

As a second attribute measure for non-answers, we calculate blathering as introduced by Barth et al. (2021). Blathering – from the Oxford English Dictionary: *"[To] talk in a long-winded way without making very much sense."* – is capturing 'information' that is volunteered but either does not meet or purposefully avoids a precise answer.

The degree of blathering in the response j in an earnings call of company i at time t is defined as

$$y_{it}^{\text{Blathering}} = 1 - \frac{\text{Finance glossary words}_{it}}{\text{Total words}_{it}}, \tag{2}$$

where words are classified as financial words based on the Hypertextual Finance Glossary by Campbell R. Harvey, consisting of more than 8,500 entries.[17] This metric assumes that for *financial firms*, the factual content in managements' responses to analysts' questions is mirrored by the usage of *finance-related words.*

Fitting the model requires us to turn $y_{it}^{\text{Blathering}}$ into a categorical variable. We therefore min-max normalize $y_{it}^{\text{Blathering}}$ and truncate to one decimal place. This leaves us with 10 categories that reflect different blathering intensities.[18]

To emphasize the usage of financial firms as a training set, consider the signal-to-noise ratio of the blathering measure for financial versus non-financial firms. The measure would be equally applicable to any industry if earnings calls were solely concerned with

[16] We classify all firms with industry code 44 or 47 in Fama-French's 48 industry portfolios as financial firms.

[17] See Campbell Harvey's webpage at Duke University, `http://people.duke.edu/~charvey/`.

[18] The results are also robust to changes in the number of categories, e.g., when truncating to two digits. We decide in favor of a low number of categories, which, on one hand, ensures sufficient observations per category and, on the other hand, is computationally more efficient.

a company's financial situation. However, those calls typically include additional topics, such as a discussion of the firm's products and other revenue drivers. The linguistic context of financial firm earnings calls is therefore unique, as their products and other topics are largely finance related.[19] Thus, non-financial words are a strict subset of non-answers, that is, they consist of noise that is not required for answering the initial question, as well as of words that are required to formulate a sentence in plain English. Think in contrast of a technology firm like Apple, which tends to discuss non-finance product and brand specifics during earnings calls. This discussion provides factual and relevant information, but the blathering metric would falsely classify these as non-factual, resulting in a low signal-to-noise ratio of the measure.[20]

To further reduce noise in this attribute measure for non-answers, we restrict our training set to answers that respond to questions with at least one finance-related word. This ensures that we do not involuntarily assign a high score for blathering to a response to a question that was unrelated to a finance context.

Our training set comprises 64,173 management answers from 2,124 earnings calls for 42 financial firms listed in the S&P 500, which accounts for roughly 10% of the textual data from earnings calls of all S&P 500 firms. As the remaining textual data of earnings calls of non-financial firms in the S&P 500 is used for validation, we end up with a very large validation set, which prevents typical problems of machine learning algorithms, such as over-fitting and data mining.

3.3 Fitted glossary

For each answer in the training set, we form an answer-term-matrix of trigrams. The response factors are metrics for rejection and blathering as described above. We employ two cleaning procedures to make sure that the answers contain meaningful words and that the resulting glossary is of general use. First, we aim to avoid company specific trigrams to achieve the most general language in the glossary. Thus, we focus on the most common trigrams of all responses that appear in at least 100 of the answers. Second, we want to remove common trigrams consisting mostly of (meaningless) stop words. In order to

[19]Northern Trust provides us with an example for a financial firm using financial terminology when discussing an individual product during an earnings call (21. October 2009). When asked: *"The PFS [Personal Financial Services] breakout between equity, fixed, and short-term? You gave the C&IS and assets under management, I'm not sure, did you give the PFS?"*. Using financial terminology, the management team responded with: *"PFS AUM was 34% equity. 32% fixed income, 34% short duration"*.

[20]Take Apple's earnings calls from 24. January 2012 as an example. Therein, the question: *"... as we're doing the math here, it seems like the ASP is up sequentially for the iPhone, so it seems like there was a very good mix of the 4s and the higher capacity 4s. Could you just comment a little bit on that versus the iPhone 4 which you cut the price of?"* was answered without any financial terminology with: *"The iPhone 4s was the most popular iPhone during the quarter. And consistent with most launches, we typically see a higher mix at the front end of a launch."*.

provide a directly applicable glossary to spoken English sentences, we do not filter for stop words before forming trigrams, but would like to remove those trigrams from the glossary that appear in more than 50% of the answers.[21] This cleaning procedure leaves us with around 3,400 trigrams.[22]

The model returns 568 (1290) trigrams with a positive (negative) loading for the rejection response factor and 1099 (970) trigrams with a positive (negative) loading for the blathering response factor. Unlike a non-answer, an answer is always specific to the context of the question. Hence, trigrams with a finance meaning show a strong negative loading by construction. However, these trigrams are only meaningful in a context-specific Q&A setting of earnings conference calls of financial firms.[23]

We only keep the 1,364 trigrams with a positive factor loading for either the attribute measure $y^{\text{Rejecting}}$ or $y^{\text{Blathering}}$ in order to increase the scope of the glossary and to allow the glossary to measure a non-answer independent of the context.[24]

Figure 2 shows trigrams from the glossary with font sizes weighted by their respective factor loading ϕ.[25] We find that phrases like "back to you", "top of my" or "not sure I" are particularly strong markers for non-answers.[26]

A non-answer does not require a specific context. As both our method and glossary are free of financial context, we believe that the metric is applicable to other fields with Q&A sessions. To corroborate this claim, consider Mark Zuckerberg's responses to the US Senate during the Cambridge Analytica hearing as anecdotal evidence. His response "Senator, my — I can certainly have my team get **back to you** on any specifics there that **I don't know**, sitting here today." is clearly a non-answer, and would have been identified by the highlighted glossary entries. In Appendix A.6, we also briefly explore Q&A data from interviews of US presidents.[27] The presidential data shows substantial variation in *NonAnswer*, a necessary condition for the measure to be informative. Particularly high examples for non-answer scores are found in interviews with President Clinton around the time when sexual assault allegations surfaced that later became the

[21] Note that this restriction is not binding, i.e., we do not delete any trigram by applying this filter.

[22] Loosening the first filtering criteria will result in a less generic and lengthier word-list that contains more trigrams specific to our training set, i.e., the financial industry. It does not, however, affect the results shown later in the validation analysis.

[23] Context-specific trigrams with a high negative factor load include, e.g., money market funds, basis points in, the investment portfolio, net interest margin, the securities portfolio, or the loan portfolio.

[24] The intersection of the two word-lists contains 263 trigrams. Interestingly, the intersecting trigrams often show high loadings $\phi^{\text{Rejecting}}$ and $\phi^{\text{Blathering}}$ for both of the attribute measures $y^{\text{Rejecting}}$ and $y^{\text{Blathering}}$.

[25] The factor loading of each trigram ϕ is defined as $\max(\phi^{Rejecting}, \phi^{Blathering})$.

[26] A full glossary of trigrams with a positive loading is provided in the Appendix and at econlinguistics.org.

[27] These interviews were collected by UCSB's American Presidency Project, see www.presidency.ucsb.edu.

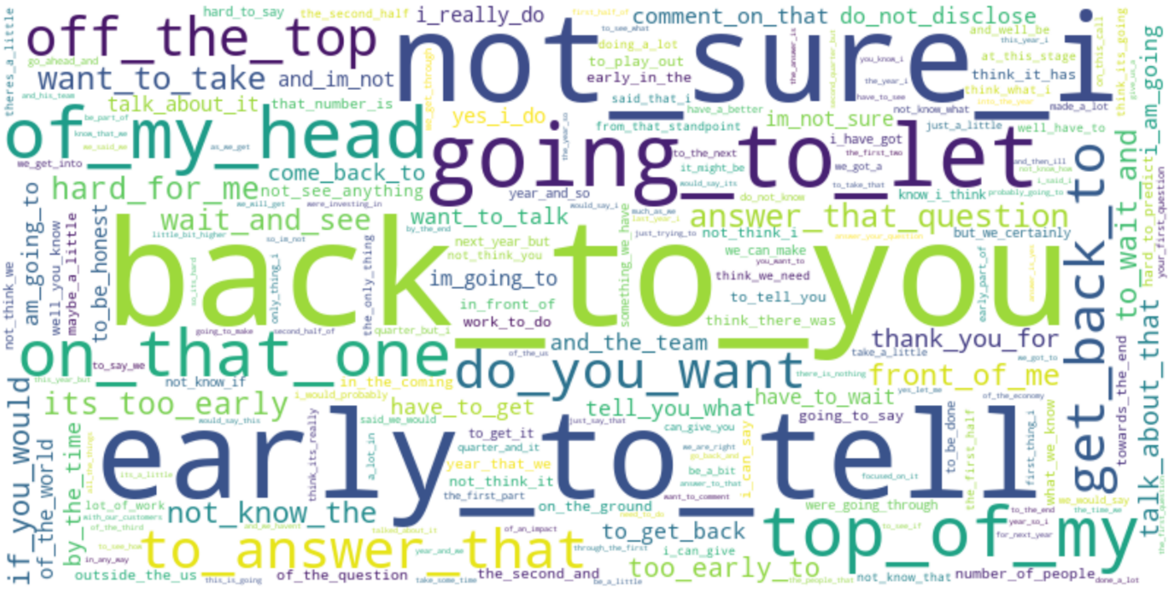


Figure 2: **Summary of the glossary.** The glossary consists of 1,364 trigrams. A larger font size indicates a higher factor load.

basis for an impeachment charge of perjury.[28]

[28]For example, during a telephone interview with Morton Kondracke, Bill Clinton responded to the question: *"Okay. Let me just ask you one more question about this. You said in a statement today that you had no improper relationship with this intern. What exactly was the nature of your relationship with her?"* with the words *"Well, let me say, the relationship's not improper, and I think that's important enough to say. But because the investigation is going on and because I don't know what is out—what's going to be asked of me, I think I need to cooperate, answer the questions, but I think it's important for me to make it clear what is not. And then, at the appropriate time, I'll try to answer what is. But let me answer, it is not an improper relationship, and I know what the word means."*.

4 Economic relevance

As financial economists, we naturally focus on applying *NonAnswer* on textual data related to our discipline. Thus, we measure non-answers for management responses in earnings calls and conduct a variety of tests to evaluate the plausibility of our glossary. We examine how markets react to non-answers by studying cumulative abnormal stock returns, implied volatilities and analysts' EPS forecasts in the wake of earnings conference calls, and investigating which questions managers try to avoid.

Financial markets are perfectly suited to assess the economic relevance of our measure for two reasons. First, economic theory gives a prior expectation of the effect that we would expect for avoiding answering analysts' questions. Second, in a financial markets' context, 'artful dodgers', as described in Rogers and Norton (2011), should be detected, as social evaluation does not play a role and the listeners' attention is directed towards the goal of identifying whether a person is answering a question. Thus, in the context of finance, we have a clear prior of a negative perception of avoiding an answer.

Investors participate in the Q&A sessions of earnings conference calls in order to reduce uncertainty about a firm's expected future performance. The theoretical asset pricing literature suggests that higher uncertainty translates to larger risk premia (Andrei and Hasler, 2014).[29] An investor's uncertainty is reduced less by a non-answer compared to a precise, context specific response. For a given prior expectation, a context-free response might even increase uncertainty. Thus, we expect market participants to react more negatively in response to non-answers. In fact, empirical literature in line with this expectation shows that not conveying information leads to a negative stock market reaction. For example, Zhou (2018) argues that obscuring information by increasing textual rather than numeric content is associated with lower cumulative abnormal returns around the earnings call date. Similarly, Hollander et al. (2010) shows that stock returns in a 90 or 120-minute window after an earnings conference call react significantly more negatively if the management refused to answer a question in the call. Taking these results at face value, a necessary condition for the validity of our glossary is to observe a negative correlation between $NonAnswer$ and stock returns after an earnings conference call.

[29] For a discussion on the impact of policy uncertainty on risk premia, see also Pástor and Veronesi (2013) or Liu et al. (2017). For the effect of disagreement as some special cases of uncertainty on asset pricesCarlin et al. (2014).

4.1 Data

We collect earnings calls of all S&P 500 non-financial companies, i.e., all calls of those S&P 500 firms that were not used to train the model. These 23,815 earnings calls of the 'validation set' is significantly larger than the training set (2,124 earnings calls) and, hence, minimizes the risk of over-fitting.

Non-answer score We apply our glossary to derive a metric that captures non-answers. For the earnings call of company i in quarter t, we count the occurrence of trigrams from the glossary in all responses of the Q&A session and divide by the total number of words, hence,

$$NonAnswer_{it} = \frac{\text{Non-answer glossary tokens}_{it}}{\text{Total words}_{it}}. \tag{3}$$

In addition, to incorporate the information on the loadings, we measure a $NonAnswer^{\phi}$ by weighting each trigram in the glossary with its respective factor loading,

$$NonAnswer_{it}^{\phi} = \frac{\sum_{k=1}^{K} \phi_k \times \text{Non-answer glossary token}_{itk}}{\text{Total words}_{it}}, \tag{4}$$

where ϕ_k is the loading associated with trigram $k \in \{1, 2, ..., K\}$.

The distribution of $NonAnswer$ for the earnings calls in our validation set is shown in Figure A1 and the sample average of $NonAnswer$ over time is shown in Figure A2. It is interesting to note that we observe a peak of non-answers during and in the immediate aftermath of the financial crisis.

Cumulative abnormal returns We obtain daily adjusted stock returns from CRSP and calculate daily abnormal return for the stock of company i at time t with the Fama-French three-factor (1993) and five-factor (2015) model returns,[30]

$$r_{i,t}^{abnormal} = r_{i,t} - r_{i,t}^{FF}.$$

We investigate the short-term effect using cumulative abnormal returns from the day of the earnings call to the day after, $CAR_{i,t}^{0;1}$.

Option implied volatility The implied volatility derived from prices of exchange-traded equity options reflects the premium that investors are willing to pay for insuring

[30]The model is calibrated to 40 trading days preceding an earnings call, with data from the Fama-French data library at `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`.

against price movements in the underlying and thus proxies for investor uncertainty.[31] We collect the daily implied volatility $\sigma_{i,t}$ derived from liquid at-the-money options with 91-day maturity from OptionMetrics, LLC. We calculate two metrics to capture the instantaneous update of investors beliefs on future volatility after a conference call. The first approach follows Rogers et al. (2009), and compares the implied volatility of company i on the day just after the call to that of the day just before the call,

$$IV_{i,t}^{-1;1} = \ln\left(\frac{\sigma_{i,t+1}}{\sigma_{i,t-1}}\right).$$

Second, we compare the change in $\sigma_{i,t}$ with a counterfactual change in the implied volatility, which we calculate as the average change in implied volatility for the 60 trading days preceding the earnings call,

$$\Delta IV_{i,t} = \frac{\sigma_{i,t+1} - \sigma_{i,t-1}}{2} - \frac{\sigma_{i,t-1} - \sigma_{i,t-60}}{59}.$$

Alternative speech characteristics The literature provides evidence that investors recognize tone sentiment and the uncertainty of the language used in earnings calls.[32] As we want to test whether our measure of non-answers is not purely capturing management tone and uncertainty, we compute standard metrics from the literature to control for these language characteristics.

For tone, we count the number of negative words in earnings calls that appear on the negative word list by Loughran and McDonald (2011). Then, we define *Negativity* of company i's earnings call in quarter t as the ratio of negative words relative to total words:

$$Negativity_{it} = \frac{\text{Negative words}_{it}}{\text{Total Words}_{it}},$$

To measure uncertainty we use the word list from Loughran and McDonald (2011).[33] Similar to the tone measure, we quantify the uncertainty of statements by counting the number of words in the earnings call that appear on this word list. *Uncertainty* for the earnings call of company i at time t is then defined as the ratio of uncertain words to

[31] See Rogers et al. (2009) for a discussion of the advantages of using implied volatility to measure investor uncertainty compared to other possible measures such as realized volatility or the dispersion in analyst forecasts.

[32] See, e.g., Price et al. (2012), Blau et al. (2015), Brockman et al. (2015) or Davis et al. (2015) for evidence on tone sentiment and Dzielinski et al. (2021) for evidence on uncertainty.

[33] Note that this word list contains also the word list of weak modals from Loughran and McDonald (2011).

total words:

$$Uncertainty_{it} = \frac{\text{Uncertainty words}_{it}}{\text{Total words}_{it}},$$

We follow the approach in Zhou (2018) to generate a variable $Numbers$ that accounts for managements' usage of numbers relative to textual words in their answers. Specifically, we use a regular expression to capture all numbers preceded by a space or a dollar sign and calculate $Numbers_{it}$ for the earnings call of company i at time t :

$$Numbers_{it} = \frac{\text{Number count}_{it}}{\text{Total words}_{it} + \text{Number count}_{it}}.$$

We further calculate for the responses of the management the complexity score proposed by Loughran and McDonald (2019).[34] Using the list of 255 words that proxy for complexity, we build the measure for the earnings call of company i at time t as follows:

$$Complexity_{it} = \frac{\text{Complex words}_{it}}{\text{Total words}_{it}},$$

Finally, we flag follow-up questions by the same analyst during an earnings call and derive a measure of forward-looking words within analysts' questions. For a single question q during the earnings call of company i at time t, we define the share of forward-looking words according to the word-lists provided by Bozanic et al. (2018) and Matsumoto et al. (2011):

$$ForwardSentiment_{qit} = \frac{\text{Forward-looking words}_{qit}}{\text{Total words}_{qit}}.$$

Earnings surprise, analyst forecast revisions, and firm characteristics We collect analyst data from IBES and calculate earnings surprises as the difference between the actual and consensus forecast earnings, divided by the share price at five trading days before the announcement in every quarter. Thus, any positive (negative) number indicates better (worse) performance than expected. As in Dzielinski et al. (2021), we rank all firms' earnings surprises into deciles and categorize earnings surprises from 1 (most negative) to 5 (least negative) and from 6 (least positive) to 10 (most positive). Moreover, we collect analyst EPS forecast data and match the latest EPS forecast prior to the call with the first EPS forecast post call. This allows us to flag whether or not

[34] As an alternative metric of complexity, we also calculate the Gunning (1952) Fog index, which is a function of the number of words per sentence (length of a sentence) and the share of complex words (words with more than two syllables) relative to total words. Using this measure does not change any of our results.

the EPS forecast was revised upwards and to calculate for each call the percentage of analysts with a positive revision.

We further collect quarterly balance sheet statistics as well as firms' market capitalizations from Compustat to calculate the book-to-market ratio, the natural logarithm of total assets and Tobin's Q as additional firm characteristics.

Table 1: Descriptive statistics

Variable	Obs.	Mean	Std. Dev.	Min	P10	P50	P90	Max
Panel A: Firm-quarter data								
$NonAnswer$	21,191	.074	.015	.029	.056	.073	.093	.14
$NonAnswer^{\phi}$	21,191	.17	.037	.048	.12	.16	.21	.38
$Negativity$	21,191	.028	.0071	.0075	.02	.028	.037	.074
$Uncertainty$	21,191	.016	.0056	0	.0089	.015	.023	.056
$EarnSurp$	21,191	5.7	2.8	1	2	6	10	10
$FF3 - CAR_{0;1}$	21,191	.0004	.052	-.16	-.06	.00033	.062	.15
$FF5 - CAR_{0;1}$	21,191	.00035	.049	-.15	-.057	.00039	.059	.14
$IV_{-1;1}$	19,432	.95	.065	.79	.88	.95	1	1.2
ΔIV	19,431	-.007	.011	-.044	-.021	-.006	.0045	.026
BTM	21,191	.41	.36	-3.2	.1	.34	.78	17
$Ln(Assets)$	21,191	9.5	1.2	5.8	8	9.4	11	14
Q	21,191	2.2	1.4	.63	1.1	1.7	3.8	36
$Numbers$	21,191	.012	.0058	0	.0052	.011	.02	.046
$Complexity$	21,191	.007	.0038	0	.0027	.0064	.012	.035
$\%Positive\ Revisions$	21,129	45	16	0	26	44	66	100
Panel B: Q&A-level data								
$NonAnswer$	621,696	.057	.051	0	0	.05	.12	.67
$NonAnswer^{\phi}$	621,696	.13	.18	0	0	.099	.29	6
$IsFollowUp_q$	621,696	.66	.47	0	0	1	1	1
$Negativity_q$	621,696	.028	.036	0	0	.018	.07	1
$ForwardSentiment_q$	621,696	.11	.065	0	.012	.11	.19	1

Notes: **Panel A** shows descriptive statistics for our firm-quarter level data with language measures aggregated over all Q&As within an earnings call. **Panel B** provides summary statistics for the language measures on the dimension of individual Q&As. $NonAnswer$ ($NonAnswer^{\phi}$) is the ratio of trigrams in our non-answer glossary (weighted by loadings) to the total words. $Negativity$ and $Uncertainty$ are the ratio of negative and uncertain words to the total words. The list of negative and uncertain words are from Loughran and McDonald (2011) word-lists. $EarnSurp$ represents the grouping of all firms in deciles of earnings surprise (defined as the difference between the actual and the consensus forecast earnings as a ratio to the share price 5 trading days before the announcement). $FF3\_CAR_{0;1}$ and $FF5\_CAR_{0;1}$ use Fama-French three (1993) and five (2015) factor model returns respectively. $IV_{-1;1}$ and ΔIV are the change in option's implied volatility around the earnings call as defined in Section 4.1. BTM defined as total Common/Ordinary Equity divided by the market value of equity. $ln(Assets)$ is the natural logarithm of total assets. Q is the Tobin's Q. $Numbers$ is the share of numbers to the sum of total words and numbers calculated as in Zhou (2018). $Complexity$ is the complexity measure derived from the Loughran and McDonald (2019) complexity word list. $\%PositiveRevisions$ is the share of analysts that revise their EPS forecast upwards. $IsFollowUp_q$ is a dummy variable that equals one if a question is a follow up question by the same analyst. $Tone_q$ is the positivity minus negativity sentiment of the question calculated by word count of the corresponding word-lists provided by Loughran and McDonald (2011). $ForwardSentiment_q$ measures the ratio of forward-looking words in a question according to the word-lists provided by Bozanic et al. (2018) and Matsumoto et al. (2011). All return variables are truncated at the 1/99% percentiles.

Descriptive statistics Table 1 presents descriptive statistics for the variables in this analysis. Our metric $NonAnswer$ and $NonAnswer^{\phi}$ are similar in magnitude with the other sentiments from the dictionary approach, i.e., $Negativity$ and $Uncertainty$. In our sample, $Negativity$ and $Uncertainty$ metrics show an average of 2.8% and 1.6%, which are comparable to estimates found in the literature (see, e.g., (Price et al., 2012) and (Dzielinski et al., 2021)). One might expect a strong correlation between the $NonAnswer$ metric and other sentiment metrics, in particular $Uncertainty$. Yet, as Table A1 shows, $NonAnswer$ only correlates with other textual measures very weakly, highlighting that the measure captures a new dimension of precise information sharing.[35]

4.2 How do non-answers affect stock returns?

We attempt to see whether markets react to non-answers and explain cumulative abnormal stock returns for the day of an earnings conference call. For this purpose, we model cumulative abnormal returns of firm i with management m around the days of the earnings call in quarter t as follows:

$$\begin{aligned} CAR_{imt} = \alpha &+ \beta_1 \cdot NonAnswer_{imt} + \beta_2 \cdot EarnSurp_{imt} \\ &+ \beta_3 \cdot Negativity_{imt} + \beta_4 \cdot Uncertainty_{imt} \\ &+ \theta \cdot X_{imt} + \mu_i + \nu_m + \gamma_t + \epsilon_{imt}. \end{aligned} \tag{5}$$

CAR_{imt} represents the cumulative abnormal return for the initial, short-term reaction ($CAR^{0;1}_{i,t}$) for firm i with management m in quarter t. $NonAnswer$ is the main variable of interest generated from our glossary, which measures management's degree of non-answers in the call in quarter t. If our glossary generates a valid measure for not conveying information, we should observe lower abnormal stock returns for earnings calls with a high $NonAnswer$ and thus expect a negative coefficient for β_1.

We control for three important variables to ensure that outside factors do not affect $NonAnswer$. First, we include a metric that captures investors' expectations of future earnings. As is standard in the literature, we measure the difference between analysts' expectations about earnings and realized earnings as earnings surprise and, for a given point in time, cluster all firms into 10 different groups, $EarnSurp$, with a larger number indicating a more positive earnings surprise.

Second, we control for two variables that have been shown to impact returns after an earnings conference call. One of these measures is $Negativity$, defined as the ratio of negative words to total words used in management's answers. In line with Price et al.

[35] We present some examples for answers with high (low) $Uncertainty$ and $NonAnswer$ in Appendix A.3.

(2012), we expect a negative coefficient for β_3. We also control for the vagueness of managements' language, *Uncertainty*. As Dzielinski et al. (2021) show that uncertainty in managements' answers to investors' questions leads to lower stock returns, we expect a negative coefficient for β_4.

Finally, we remove all observable and unobservable firm-specific time-constant variation by including firm fixed effects, μ_i, as well as time (quarter-year) fixed effects, γ_t, to control for firm-constant factors and common trends of abnormal returns in a given quarter, respectively. We further include CEO fixed effects, ν_m, in order to absorb a manager-specific component, which neither the current and future performance of the company nor strategic incentives can explain (Davis et al., 2015). This enables us to separate the effect of $NonAnswer$ from personal specific unobservable time-constant characteristics. To account for autocorrelations of the errors, we employ two-way clustering (Cameron et al., 2011) and cluster standard errors at the firm and time dimensions.

Table 2 and Table 3 display the results of the regression model outlined in Equation 5 for $CAR_{i,t}^{0;1}$ using $NonAnswer$ and the loading-weighted $NonAnswer^{\phi}$. In both tables, we observe a negative and highly significant coefficient, highlighting the negative effect that not answering to analysts' questions has on short-term cumulative abnormal returns. These results are in line with our expectation, provided our glossary measures non-answers. Note that this result also holds if we control for earnings surprises, other characteristics of management language, as well as industry, firm and CEO fixed effects and common time trends by quarter-year fixed effects.[36] Moreover, the coefficients of all control variables are in line with our expectations and with the existing literature: we find a positive and highly significant coefficient for the earnings surprise group, i.e. a greater difference in the actual earnings and earnings expected by analysts leads to more positive abnormal returns. We further obtain a negative point estimate for the tone measure, and, in line with Dzielinski et al. (2021), a negative coefficient for the uncertainty measure.

4.3 Which non-answer symptom drives stock returns?

We derive our glossary from two different symptoms of non-answers, rejecting an answer and blathering. To shed light on which of these violations of the "Gricean norms" is more informative in the earnings call context, we derive a separate glossary for each of the two symptoms and repeat the baseline empirical analysis. The results are shown in Table 4. We find that both symptoms on their own lead to negative abnormal returns. However, the blathering-based glossary appears stronger on its own and dominates the rejecting-derived glossary when combining the two variables in one regression. One potential

[36] As a robustness test, we present cross-sectional Fama-MacBeth regressions in Appendix A.4, which confirm our main result.

Table 2: Management $NonAnswer$ and abnormal returns ($CAR_{0;1}$)

	$FF3-CAR_{0;1}$			$FF5-CAR_{0;1}$		
	(1)	(2)	(3)	(4)	(5)	(6)
$NonAnswer$	-0.080***	-0.105***	-0.111**	-0.073***	-0.095***	-0.101**
	(-3.26)	(-4.01)	(-2.62)	(-3.04)	(-3.10)	(-2.52)
$Negativity$	-0.415***	-0.421***	-0.617***	-0.361***	-0.483***	-0.569***
	(-6.54)	(-6.53)	(-7.62)	(-5.98)	(-7.29)	(-7.28)
$Uncertainty$	-0.093	-0.041	-0.108	-0.073	-0.090	-0.111
	(-1.31)	(-0.56)	(-1.24)	(-1.11)	(-1.15)	(-1.32)
$Numbers$		-0.190***	-0.279***	-0.171**	-0.226***	-0.269***
		(-2.70)	(-2.90)	(-2.65)	(-2.75)	(-3.05)
$Complexity$		0.168	0.398***	0.184*	0.229*	0.366**
		(1.64)	(2.69)	(1.98)	(1.98)	(2.64)
Constant	0.007	0.006	0.130***	0.006	0.092***	0.121***
	(1.60)	(1.21)	(4.82)	(1.50)	(4.86)	(4.70)
Observations	21191	21035	20182	21191	21191	20182
R^2	0.045	0.048	0.131	0.044	0.080	0.130
FirmControls	Yes	Yes	Yes	Yes	Yes	Yes
QuarterYear FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	No	Yes	Implied	No	Implied	Implied
Firm FE	No	No	Yes	No	Yes	Yes
CEO FE	No	No	Yes	No	No	Yes

Notes: OLS regressions for Equation (5). The dependent variable is the abnormal returns over the Fama-French three (1993) and five (2015) factor model returns cumulated from the day of the earnings call to the day after it, $FF3-CAR_{0;1}$ ($FF5-CAR_{0;1}$). $NonAnswer$ is the ratio of trigrams in our non-answer glossary to the total words. $Negativity$ and $Uncertainty$ are the ratio of negative and uncertain words to the total words. List of negative and uncertain words are from Loughran and McDonald (2011) word-lists. *EarnSurp* represents the grouping of all firms in deciles of earnings surprise (defined as the difference between the actual and the consensus forecast earnings as a ratio to the share price 5 trading days before the announcement). BTM defined as total Common/Ordinary Equity divided by the market value of equity. $ln(Assets)$ is the natural logarithm of total assets. Q is the Tobin's Q. $Numbers$ is the share of numbers to the sum of total words and numbers calculated as in Zhou (2018). *Complexity* is the complexity measure derived from the Loughran and McDonald (2019) complexity word list. t-statistics are given in parentheses. Standard errors are clustered in the firm and quarter level. ***, **, * indicate significance at the 1%, 5% and 10% levels.

Table 3: Management $NonAnswer$ and abnormal returns ($CAR_{0;1}$)

	$FF3-CAR_{0;1}$			$FF5-CAR_{0;1}$		
	(1)	(2)	(3)	(4)	(5)	(6)
$NonAnswer^{\phi}$	-0.038***	-0.045***	-0.047***	-0.034***	-0.042***	-0.042***
	(-3.88)	(-4.43)	(-3.01)	(-3.61)	(-3.85)	(-2.92)
$Negativity$	-0.410***	-0.413***	-0.613***	-0.357***	-0.478***	-0.565***
	(-6.48)	(-6.42)	(-7.56)	(-5.92)	(-7.18)	(-7.21)
$Uncertainty$	-0.085	-0.034	-0.106	-0.067	-0.086	-0.109
	(-1.21)	(-0.47)	(-1.22)	(-1.02)	(-1.10)	(-1.29)
$Numbers$		-0.182**	-0.275***	-0.167**	-0.223***	-0.265***
		(-2.59)	(-2.88)	(-2.58)	(-2.71)	(-3.03)
$Complexity$		0.164	0.398***	0.178*	0.226*	0.366**
		(1.60)	(2.70)	(1.91)	(1.96)	(2.64)
Constant	0.007*	0.005	0.129***	0.006	0.091***	0.121***
	(1.69)	(1.09)	(4.82)	(1.56)	(4.82)	(4.69)
Observations	21191	21035	20182	21191	21191	20182
R^2	0.045	0.048	0.131	0.044	0.080	0.130
FirmControls	Yes	Yes	Yes	Yes	Yes	Yes
QuarterYear FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	No	Yes	Implied	No	Implied	Implied
Firm FE	No	No	Yes	No	Yes	Yes
CEO FE	No	No	Yes	No	No	Yes

Notes: OLS regressions for Equation (5). The dependent variable is the abnormal returns over the Fama-French three (1993) and five (2015) factor model returns cumulated from the day of the earnings call to the day after it, $FF3-CAR_{0;1}$ ($FF5-CAR_{0;1}$). $NonAnswer^{\phi}$ is the ratio of trigrams in our non-answer glossary weighted by loadings to the total words. $Negativity$ and $Uncertainty$ are the ratio of negative and uncertain words to the total words. List of negative and uncertain words are from Loughran and McDonald (2011) word-lists. $EarnSurp$ represents the grouping of all firms in deciles of earnings surprise (defined as the difference between the actual and the consensus forecast earnings as a ratio to the share price 5 trading days before the announcement). BTM defined as total Common/Ordinary Equity divided by the market value of equity. $ln(Assets)$ is the natural logarithm of total assets. Q is the Tobin's Q. $Numbers$ is the share of numbers to the sum of total words and numbers calculated as in Zhou (2018). $Complexity$ is the complexity measure derived from the Loughran and McDonald (2019) complexity word list. t-statistics are given in parentheses. Standard errors are clustered in the firm and quarter level. ***, **, * indicate significance at the 1%, 5% and 10% levels.

explanation for this finding lies in the difference of the questions on which we observe the different symptoms of a non-answer. For example, Gow et al. (2021) find that managers directly refuse to answer questions with greater uncertainty and questions asking for proprietary information, both types of questions for which financial markets might already expect to receive a non-answer.

Note that an immediate use of the MNIR attribute measures does not provide a clean measure of non-answers in general. This is not surprising for the blathering measure, as it only captures non-answers for financial firms. The approach to measuring the refusal of an answer in Gow et al. (2021) suffers from insufficient power, especially when aggregating several questions within a call, c.f. Table 4.

4.4 Can random glossaries produce similar results?

In order to show that negative abnormal returns are indeed due to wording that indicates non-answers, we run a Monte Carlo simulation by randomly drawing 1,364 trigrams from the training set 1000 times. For each of these placebo dictionaries, we run regressions as in column 2 of Table 2. The distribution of t-statistics for the $NonAnswer$ coefficient is roughly normal and centered around zero, as shown in Figure 3.[37] The t-statistic for $NonAnswer$ for our original glossary is -4.27 (see Table 2, column 2). This clearly shows that it would be extremely unlikely to produce economically significant results by randomly drawing a glossary from the universe of trigrams.

4.5 How do non-answers affect expected volatility?

Dodging a question retards the information flow and hinders the reduction of informational asymmetries between the management and investors. We therefore would expect to observe that investor uncertainty is reduced more after earnings calls with low $NonAnswer$. To this extent, we analyze the short-term change as well as the abnormal change in implied volatility, similar to the analysis of abnormal returns above. Table 5 shows the results.

We observe a higher post-earnings call implied volatility for earnings calls with high $NonAnswer$, i.e., investors are willing to pay a higher premium in order to insure themselves against stock price changes after conference calls in which managers more frequently avoid answering questions.

[37]We consider the distribution of the t-statistic from the 1,000 regression coefficients as we are not only after the effect of the glossary, but also the precision of the point estimate for each draw (glossary).

Table 4: Different symptoms of management $NonAnswer$

	$FF5 - CAR_{0;1}$				
	(1)	(2)	(3)	(4)	(5)
$NonAnswer$	-0.091***				
	(-3.64)				
$NonAnswer^{Blathering}$		-0.099***		-0.112***	
		(-3.61)		(-3.07)	
$NonAnswer^{Rejecting}$			-0.083*	0.049	
			(-1.82)	(0.76)	
$Rejecting$					0.001
					(0.41)
$Negativity$	-0.388***	-0.390***	-0.371***	-0.391***	-0.366***
	(-6.51)	(-6.56)	(-6.22)	(-6.56)	(-6.10)
$Uncertainty$	-0.060	-0.064	-0.061	-0.072	-0.079
	(-0.85)	(-0.90)	(-0.90)	(-1.07)	(-1.11)
Constant	0.004	0.004	-0.001	0.004	-0.003
	(0.91)	(0.88)	(-0.32)	(0.85)	(-0.78)
Observations	21035	21035	21035	21035	21035
R^2	0.046	0.046	0.046	0.046	0.045
FirmControls	Yes	Yes	Yes	Yes	Yes
QuarterYear FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes

Notes: OLS regressions for Equation (5). The dependent variable is the abnormal returns over the Fama-French three (1993) factor model returns cumulated from the day of the earnings call to the day after it, $FF3 - CAR_{0;1}$. $NonAnswer^{Blathering}$ ($NonAnswer^{Rejecting}$) is the ratio of trigrams in our non-answer glossary derived from the symptom blathering (rejecting) to the total words. $Negativity$ and $Uncertainty$ are the ratio of negative and uncertain words to the total words. List of negative and uncertain words are from Loughran and McDonald (2011) word-lists. $EarnSurp$ represents the grouping of all firms in deciles of earnings surprise (defined as the difference between the actual and the consensus forecast earnings as a ratio to the share price 5 trading days before the announcement). BTM defined as total Common/Ordinary Equity divided by the market value of equity. $ln(Assets)$ is the natural logarithm of total assets. Q is the Tobin's Q. $Numbers$ is the share of numbers to the sum of total words and numbers calculated as in Zhou (2018). $Complexity$ is the complexity measure derived from the Loughran and McDonald (2019) complexity word list. t-statistics are given in parentheses. Standard errors are clustered in the firm and quarter level. ***, **, * indicate significance at the 1%, 5% and 10% levels.

Table 5: Management $NonAnswer$ and option implied volatility ($\Delta\ IV$ and $IV_{-1;1}$)

	ΔIV				$IV_{-1;1}$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$NonAnswer$	0.020**	0.024***			0.077*	0.096**		
	(2.48)	(2.97)			(1.78)	(2.21)		
$NonAnswer^{\phi}$			0.007**	0.008***			0.026	0.028*
			(2.47)	(2.67)			(1.62)	(1.72)
$Negativity$		0.047***		0.045**		0.351***		0.339***
		(2.66)		(2.50)		(3.61)		(3.49)
$Uncertainty$		0.032*		0.031*		0.205**		0.207**
		(1.82)		(1.81)		(2.06)		(2.07)
$Numbers$		0.042*		0.039		0.193		0.182
		(1.77)		(1.65)		(1.54)		(1.45)
$Complexity$		0.043		0.042		0.009		0.004
		(1.45)		(1.43)		(0.05)		(0.02)
Constant	-0.014***	-0.017***	-0.014***	-0.017***	0.943***	0.925***	0.944***	0.929***
	(-7.62)	(-8.19)	(-7.55)	(-8.03)	(95.08)	(82.98)	(96.05)	(84.42)
Observations	19275	19275	19275	19275	19276	19276	19276	19276
R^2	0.152	0.153	0.151	0.153	0.181	0.183	0.181	0.183
FirmControls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
QuarterYear FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: OLS regressions with the dependent variable $IV_{-1;1}(\Delta IV)$ in the columns 1-4 (5-8) indicating the change in an option's implied volatility around the earnings call as defined in Section 4.1. $NonAnswer$ ($NonAnswer^{\phi}$) is the ratio of trigrams in our non-answer glossary (weighted by loadings) to the total words. $Negativity$ and $Uncertainty$ are the ratio of negative and uncertain words to the total words. List of negative and uncertain words are from Loughran and McDonald (2011) word-lists. $Numbers$ is the share of numbers to the sum of total words and numbers calculated as in Zhou (2018). *Complexity* is the complexity measure derived from the Loughran and McDonald (2019) complexity word list. Firm control variables include *EarnSurp*, BTM, $ln(Assets)$, and Q. *EarnSurp* represents the grouping of all firms in deciles of earnings surprise (defined as the difference between the actual and the consensus forecast earnings as a ratio to the share price 5 trading days before the announcement) to 5 negative and 5 (zero and) positive groups. BTM defined as total Common/Ordinary Equity divided by the market value of equity. $ln(Assets)$ is the natural logarithm of total assets. Q is the Tobin's Q. t-statistics are given in parentheses. Standard errors are clustered in the firm and quarter level. ***, **, * indicate significance at the 1%, 5% and 10% levels.

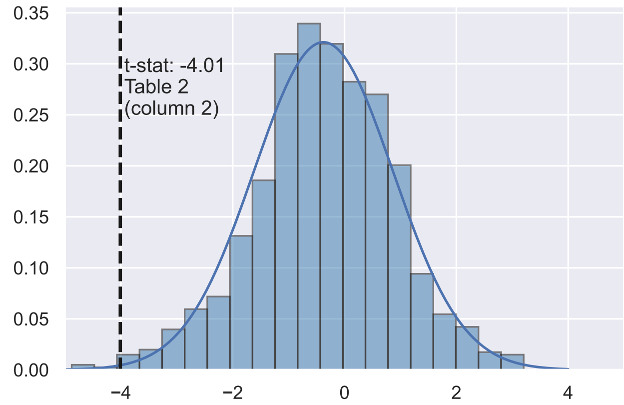


Figure 3: Histogram of t-statistics for the $NonAnswer$ coefficient from Equation 5 with 1000 word-lists of 1,364 trigrams randomly selected from the universe of trigrams. The dashed line shows the t-statistic of a regression with our non-answer glossary (column 2 of Table 2).

4.6 How analysts respond to non-answers?

Next, we investigate analysts' responses to non-answers to see whether or not non-answers affect analysts' expectations. In particular, we track analysts' EPS forecasts before and after an earnings call and calculate the percentage of analysts that increase their EPS estimate after the call.

Results are shown in Table 6 for the unweighted (columns 1-2) and weighted non-answer score (column 3-4). We find that calls with more non-answers are significantly more often followed by a non-positive revision of the EPS estimate.

4.7 Which questions receive non-answers?

We analyze which questions managers avoid precisely responding to during earnings calls. It is reasonable to assume that the management is more likely not to answer disadvantageous and tougher questions, as they want to evade critical questions.[38] We proxy for critical questions in two ways. First, we calculate the negativity sentiment for each question, assuming that a more negative tone reflects a more critical question. Second, we flag whether a question is a follow-up question by the same analyst.[39]

It is also likely that the management may not be able to answer questions about

[38]See, e.g., Mayew (2008) or Cohen et al. (2020) who show that managers try to avoid unfavorable questions by not allowing unfavorable analysts to ask a question.

[39]The second proxy is based on Clayman (1993), who argues that analysts have the capacity to recognize and counter evasive answers by asking a follow-up question, so that follow-up questions by the same analyst are usually more critical.

Table 6: Analysts EPS estimate around the call

	$\%Positive\ Revisions$			
	(1)	(2)	(3)	(4)
$NonAnswer$	-24.038*	-16.691		
	(-1.71)	(-1.40)		
$NonAnswer^{\phi}$			-12.838**	-8.624*
			(-2.44)	(-1.94)
$Negativity$	-99.697***	-115.567***	-98.904***	-114.560***
	(-3.48)	(-4.47)	(-3.48)	(-4.50)
$Uncertainty$	9.789	-45.543	12.820	-43.463
	(0.31)	(-1.63)	(0.40)	(-1.56)
Constant	42.942***	44.093***	43.280***	44.231***
	(15.23)	(18.11)	(15.69)	(18.84)
Observations	21129	20973	21129	20973
R^2	0.075	0.112	0.075	0.112
FirmControls	Yes	Yes	Yes	Yes
QuarterYear FE	Yes	Yes	Yes	Yes
Industry FE	No	Yes	No	Yes

Notes: This table shows the results for OLS regression with the dependent variable $\%PositiveRevisions$, measuring the share of analysts with an upward revision of the EPS forecast after an earnings call. $NonAnswer$ ($NonAnswer^{\phi}$) is the ratio of trigrams in our non-answer glossary (weighted by loadings) to the total words. $Negativity$ and $Uncertainty$ are the ratio of negative and uncertain words to the total words. List of negative and uncertain words are from Loughran and McDonald (2011) word-lists. Firm controls include $EarnSurp$, BTM, $ln(Assets)$ and Tobin's Q. $EarnSurp$ represents the grouping of all firms in deciles of earnings surprise (defined as the difference between the actual and the consensus forecast earnings as a ratio to the share price 5 trading days before the announcement) to 5 negative and 5 (zero and) positive groups. BTM defined as total Common/Ordinary Equity divided by the market value of equity. $ln(Assets)$ is the natural logarithm of total assets. Q is the Tobin's Q. t-statistics are given in parentheses. Standard errors are clustered in the firm and quarter level. ***, **, * indicate significance at the 1%, 5% and 10% levels.

future outcomes. Thus, we add the ratio of forward-looking phrases in a question as an additional dimension that may result in a non-answer and then regress $NonAnswer$ on these metrics. Note that by analyzing questions and answers within the same earnings conference call, we can control for many observable and unobservable factors, such as management or firm characteristics.

Table 7: $NonAnswer$ in response to follow-up questions.

	$NonAnswer$			$NonAnswer^{\phi}$		
	(1)	(2)	(3)	(4)	(5)	(6)
$IsFollowUp_q$	0.0016***	0.0017***	0.0021***	0.0079***	0.0081***	0.0087***
	(10.75)	(11.64)	(14.32)	(16.37)	(16.84)	(18.00)
$Negativity_q$		0.0191***	0.0178***		0.0324***	0.0304***
		(9.49)	(8.84)		(4.57)	(4.29)
$ForwardSentiment_q$			0.0328***			0.0489***
			(27.79)			(11.97)
Observations	621696	621696	621696	621696	621696	621696
R^2	0.066	0.066	0.068	0.050	0.050	0.050
EarningsCall FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table shows the results for OLS regression with the dependent variable $NonAnswer$ ($NonAnswer^{\phi}$) in columns 1 to 3 (4 to 6). $IsFollowUp_q$ is a dummy variable that equals one if a question is a follow up question by the same analyst. $Negativity_q$ is the negativity sentiment of the question calculated by word count of the corresponding word-lists provided by Loughran and McDonald (2011). $ForwardSentiment_q$ measures the ratio of forward-looking words in a question according to the word-lists provided by Bozanic et al. (2018) and Matsumoto et al. (2011). All the specifications control for earnings call fixed effects. t-statistics are given in parentheses. Standard errors are clustered at the earnings call level. ***, **, * indicate significance at the 1%, 5% and 10% levels.

Table 7 shows the regression results for $NonAnswer$ (columns 1 - 3) and $NonAnswer^{\phi}$ (columns 4 - 6). We find that $NonAnswer$ is greater in responses to follow up questions, in line with our expectation. In addition, we observe that questions with a more negative tone are associated with a higher value for $NonAnswer$, indicating that managers more often dodge critical questions. Finally, we find higher values for $NonAnswer$ in response to forward looking questions.

5 Conclusion

The asymmetric distribution of information is considered a key friction in economics. While question and answer (Q&A) settings are designed to remove information asymmetries, the addressee of a question does not necessarily provide the requested information. Building on a large textual dataset of questions and answers and employing a supervised machine learning framework, we generate a glossary that can identify non-answers in earnings Q&A.

Using a multinomial inverse regression (Taddy, 2013b), we identify a glossary of 1,364 trigrams such as 'back to you', 'do not know', 'hard to predict', etc., which are frequently used to refrain from answering a question in a concise and factual manner. The glossary is derived from earnings conference calls, where investors and analysts can directly question senior managers' during Q&A sessions. However, as non-answers do not contain any context- or industry-specific vocabulary, the glossary is applicable to a broad Q&A context, such as political interviews or senate hearings.

We provide evidence for the plausibility and economic relevance of the glossary. In particular, we apply the glossary to market reactions after earnings conference calls for a large sample of firms over a span of 16 years. In regressions with multiple control variables, we show a strong negative impact for the measure derived from our glossary, i.e., not answering analysts' questions leads on average to negative abnormal stock returns after an earnings call. We also link our measure to option implied volatilities after earnings conference calls and show that investor uncertainty increases if management fails to provide the information requested in the call. Both results are in line with the theoretical asset pricing literature, which suggests that higher uncertainty translates to larger risk premia (Andrei and Hasler, 2014).

We further observe that financial analysts provide less often a positive update of their EPS estimates after calls with high non-answers and document that non-answers are observed more prevalently for tougher and more critical questions (Mayew, 2008; Cohen et al., 2020). Within the same conference call, *NonAnswer* is higher for managements' responses to follow-up questions by the same analyst and for managements' responses to more negative questions. We also observe higher *NonAnswer* for questions with forward-looking sentences, i.e., questions that refer to (potentially unknown) future outcomes.

A non-answer does not require a specific context. As both our method and glossary are free of financial context, we believe that the measure is applicable to other fields with a Q&A setup.

References

Andrei, D. and Hasler, M. (2014). Investor Attention and Stock Market Volatility. *Review of Financial Studies*, 28(1):33–72.

Barth, A., Mansouri, S., Woebbeking, F., and Zörgiebel, S. (2021). How to talk down your stock performance. Working paper.

Bekkerman, R. and Allan, J. (2004). Using bigrams in text categorization. Technical report, Technical Report IR-408, Center of Intelligent Information Retrieval, UMass Amherst.

Blau, B. M., DeLisle, J. R., and Price, S. M. K. (2015). Do sophisticated investors interpret earnings conference call tone differently than investors at large? Evidence from short sales. *Journal of Corporate Finance*, 31:203–219.

Bloomfield, R. (2008). Discussion of "annual report readability, current earnings, and earnings persistence". *Journal of Accounting and Economics*, 45(2-3):248–252.

Bonsall, S. B., Leone, A. J., Miller, B. P., and Rennekamp, K. (2017). A plain english measure of financial reporting readability. *Journal of Accounting and Economics*, 63(2-3):329–357.

Bozanic, Z., Roulstone, D. T., and Van Buskirk, A. (2018). Management earnings forecasts and other forward-looking statements. *Journal of Accounting and Economics*, 65(1):1–20.

Brockman, P., Li, X., and Price, S. M. (2015). Differences in conference call tones: Managers vs. analysts. *Financial Analysts Journal*, 71(4):24–42.

Buller, D. B. and Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory*, 6(3):203–242.

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.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics*, 29(2):238–249.

Carlin, B. I., Longstaff, F. A., and Matoba, K. (2014). Disagreement and asset prices. *Journal of Financial Economics*, 114(2):226 – 238.

Clayman, S. (1993). Reformulating the question: A device for answering/not answering questions in news interviews and press conferences. *Text - Interdisciplinary Journal for the Study of Discourse*, 13.

Cohen, L., Lou, D., and Malloy, C. J. (2020). Casting conference calls. *Management Science*, 66(11):5015–5039.

Cook, R. D. et al. (2007). Fisher lecture: Dimension reduction in regression. *Statistical Science*, 22(1):1–26.

Dave, K., Lawrence, S., and Pennock, D. M. (2003). Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. In *Proceedings of the 12th international conference on World Wide Web*, pages 519–528.

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(2):639–673.

Dzielinski, M., Wagner, A. F., and Zeckhauser, R. J. (2021). CEO clarity. *HKS Working Paper No. RWP17-017*.

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.

Fama, E. F. and French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1):1–22.

Gentzkow, M., Shapiro, J. M., and Taddy, M. (2019). Measuring group differences in high-dimensional choices: Method and application to congressional speech. *Econometrica*, 87(4):1307–1340.

Gow, I. D., Larcker, D. F., and Zakolyukina, A. A. (2021). Non-answers during conference calls. *Journal of Accounting Research*, 59(4):1349–1384.

Grice, P. H. (1989). *Studies in the Way of Words*. Harvard University Press., Cambridge, MA.

Guay, W., Samuels, D., and Taylor, D. (2016). Guiding through the fog: Financial statement complexity and voluntary disclosure. *Journal of Accounting and Economics*, 62(2-3):234–269.

Gunning, R. (1952). *The technique of clear writing*. McGraw-Hill, New York.

Harvey, C. R. (2016). Harvey's hypertextual finance glossary. *URL: https://people. duke. edu/˜ charvey/Classes/wpg/bfglosm. htm [Version 8 November 2016.].*

Hastie, T., Tibshirani, R., Friedman, J., and Franklin, J. (2009). *The Elements of Statistical Learning. Data Mining, Inference, and Prediction*. Springer, 2nd edition.

Henry, E. (2008). Are Investors Influenced By How Earnings Press Releases Are Written? *Journal of Business Communication*, 45(4):363–407.

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.

Jelinek, F. (1991). Up from trigrams!-the struggle for improved language models. In *Second European Conference on Speech Communication and Technology*.

Lee, J. (2015). Can investors detect managers' lack of spontaneity? adherence to predetermined scripts during earnings conference calls. *The Accounting Review*, 91(1):229–250.

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, 45(2-3):221–247.

Liu, L. X., Shu, H., and Wei, K. J. (2017). The impacts of political uncertainty on asset prices: Evidence from the bo scandal in china. *Journal of Financial Economics*, 125(2):286 – 310.

Loughran, T. and McDonald, B. (2011). When is a Liability not a Liability? *Journal of Finance*, 66(1):35 – 65.

Loughran, T. and McDonald, B. (2014). Measuring readability in financial disclosures. *Journal of Finance*, 69(4):1643–1671.

Loughran, T. and McDonald, B. (2019). Measuring firm complexity. Working paper.

Loughran, T. and McDonald, B. (2020). Textual analysis in finance. *Annual Review of Financial Economics*, 12:357–375.

Mansouri, S. (2021). Does firm's silence drive media's attention away? *Available at SSRN.*

Matsumoto, D., Pronk, M., and Roelofsen, E. (2011). What makes conference calls useful? the information content of managers' presentations and analysts' discussion sessions. *The Accounting Review*, 86(4):1383–1414.

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

McCornack, S. A., Levine, T. R., Solowczuk, K. A., Torres, H. I., and Campbell, D. M. (1992). When the alteration of information is viewed as deception: An empirical test of information manipulation theory. *Communication Monographs*, 59(1):17–29.

Pak, A. and Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. In *LREc*, volume 10, pages 1320–1326.

Pástor, L. and Veronesi, P. (2013). Political uncertainty and risk premia. *Journal of Financial Economics*, 110(3):520 – 545.

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.

Rogers, J. L., Skinner, D. J., and Buskirk, A. V. (2009). Earnings guidance and market uncertainty. *Journal of Accounting and Economics*, 48(1):90 – 109.

Rogers, T. and Norton, M. I. (2011). The artful dodger: Answering the wrong question the right way. *Journal of Experimental Psychology: Applied*, 17(2):139 – 147.

Taddy, M. (2013a). The gamma lasso. *arXiv preprint arXiv:1308.5623*.

Taddy, M. (2013b). Multinomial inverse regression for text analysis. *Journal of the American Statistical Association*, 108(503):755–770.

Taddy, M. (2015). Distributed multinomial regression. *The Annals of Applied Statistics*, 9(3):1394–1414.

Zhou, D. (2018). Do numbers speak louder than words? Technical report.

Zuckerberg, M. Facebook, social media privacy, and the use and abuse of data. Joint hearing before the COMMITTEE ON COMMERCE, SCIENCE, AND TRANSPORTATION UNITED STATES SENATE and the COMMITTEE ON THE JUDICIARY UNITED STATES SENATE, One Hundred Fifteenth Congress, Second Session.

A Appendix

A.1 The Non-Answer Glossary

A lists of all trigrams in the glossary with their corresponding loading ϕ. A machine readable version of the glossary is available at econlinguistics.org

Token	ϕ
back.to.you	20.8
not.sure.i	13.9
early.to.tell	13.8
going.to.let	13.3
of.my.head	13.0
top.of.my	12.3
on.that.one	12.1
to.answer.that	12.1
off.the.top	12.0
get.back.to	11.9
do.you.want	11.6
answer.that.question	11.3
want.to.take	11.3
hard.for.me	10.7
its.too.early	10.4
to.wait.and	9.6
front.of.me	9.3
not.know.the	9.0
wait.and.see	8.8
if.you.would	8.8
too.early.to	8.7
talk.about.that	8.6
thank.you.for	8.6
comment.on.that	8.5
want.to.talk	8.2
come.back.to	7.9
have.to.get	7.9
tell.you.what	7.9
do.not.disclose	7.8
to.get.back	7.8
and.the.team	7.8
by.the.time	7.7
am.going.to	7.7
i.am.going	7.6
im.going.to	7.6
have.to.wait	7.5
of.the.world	7.5
i.really.do	7.5
im.not.sure	7.4
talk.about.it	7.3
yes.i.do	7.2
to.be.honest	7.2
and.im.not	7.0
not.see.anything	7.0
of.the.question	6.9
number.of.people	6.9
think.there.was	6.9
to.play.out	6.8
year.that.we	6.8
to.tell.you	6.7
on.the.ground	6.7
outside.the.us	6.7
in.front.of	6.6
that.number.is	6.6
the.second.and	6.6
towards.the.end	6.6
hard.to.predict	6.5
going.to.say	6.5
what.we.know	6.5

Token	ϕ
i.can.say	6.4
not.think.it	6.4
not.think.i	6.4
well.you.know	6.4
work.to.do	6.4
think.it.has	6.3
maybe.a.little	6.3
early.in.the	6.3
in.the.coming	6.3
were.going.through	6.3
not.know.if	6.3
know.i.think	6.2
year.and.so	6.2
we.can.make	6.1
i.have.got	6.1
next.year.but	6.1
but.we.certainly	6.1
to.be.done	6.1
not.think.we	6.0
to.get.it	6.0
hard.to.say	6.0
something.we.have	6.0
from.that.standpoint	6.0
doing.a.lot	6.0
think.we.need	6.0
not.know.that	6.0
i.can.give	5.9
at.this.stage	5.9
the.second.half	5.9
said.that.i	5.9
think.what.i	5.9
and.well.be	5.9
think.its.going	5.9
not.think.you	5.9
lot.of.work	5.9
the.first.half	5.8
theres.a.little	5.8
well.have.to	5.8
the.only.thing	5.7
i.would.probably	5.7
be.a.bit	5.7
have.a.better	5.6
to.the.next	5.6
think.its.really	5.6
quarter.and.it	5.5
the.first.part	5.5
we.would.say	5.5
to.say.we	5.5
on.this.call	5.5
take.a.little	5.5
not.know.what	5.4
early.part.of	5.4
first.thing.i	5.4
we.got.a	5.3
can.give.you	5.3
it.might.be	5.3
said.we.would	5.3

Token	ϕ
made.a.lot	5.3
your.first.question	5.2
only.thing.i	5.2
we.get.into	5.2
a.lot.in	5.1
do.not.know	5.1
go.ahead.and	5.1
quarter.but.i	5.1
we.get.through	5.1
just.a.little	5.1
answer.your.question	5.1
all.the.things	5.1
the.time.we	5.1
the.first.two	5.0
so.im.not	5.0
we.are.right	5.0
second.half.of	5.0
there.is.nothing	5.0
much.as.we	5.0
focused.on.it	5.0
i.said.i	5.0
we.said.we	5.0
just.say.that	4.9
to.take.that	4.9
done.a.lot	4.9
have.to.see	4.9
to.see.what	4.9
this.is.going	4.9
you.want.to	4.9
in.any.way	4.9
go.back.and	4.9
want.to.comment	4.8
this.year.i	4.8
and.his.team	4.8
and.then.ill	4.8
would.say.i	4.8
of.the.us	4.8
yes.let.me	4.8
would.say.this	4.8
were.investing.in	4.8
year.so.i	4.7
you.know.i	4.7
with.our.customers	4.7
of.an.impact	4.7
going.to.make	4.7
into.the.year	4.7
and.we.havent	4.7
would.say.its	4.7
not.know.how	4.7
first.half.of	4.6
little.bit.higher	4.6
be.a.little	4.6
just.trying.to	4.6
be.part.of	4.6
of.the.economy	4.6
this.year.but	4.6
the.year.so	4.5

Token	ϕ
of.the.third	4.5
take.some.time	4.5
we.will.get	4.5
to.see.how	4.5
last.year.i	4.5
second.quarter.but	4.5
we.got.to	4.5
through.the.first	4.5
talked.about.it	4.5
the.year.i	4.4
so.its.hard	4.4
for.next.year	4.4
the.answer.is	4.4
by.the.end	4.4
need.to.do	4.4
give.us.a	4.4
answer.to.that	4.4
probably.going.to	4.4
the.first.question	4.4
know.that.we	4.4
to.the.end	4.3
answer.is.yes	4.3
year.and.we	4.3
the.people.that	4.3
its.a.little	4.3
as.we.get	4.3
to.see.if	4.3
want.to.go	4.3
the.year.and	4.3
first.quarter.is	4.3
to.go.back	4.3
know.that.i	4.3
bit.more.than	4.3
one.way.or	4.2
of.the.people	4.2
you.know.its	4.2
try.to.get	4.2
first.part.of	4.2
the.next.quarter	4.2
i.would.be	4.2
i.will.let	4.2
as.good.as	4.2
we.hope.to	4.2
i.have.seen	4.1
as.we.speak	4.1
customers.that.we	4.1
i.cannot.give	4.1
not.think.there	4.1
the.year.that	4.1
i.think.well	4.1
a.little.higher	4.1
is.the.first	4.1
cannot.give.you	4.1
the.number.that	4.1
not.sure.that	4.1
the.answer.to	4.1

Token	ϕ
and.then.well	4.1
get.a.lot	4.1
couple.of.weeks	4.1
for.us.its	4.1
i.would.like	4.0
at.least.for	4.0
im.not.going	4.0
do.not.think	4.0
the.top.of	4.0
think.is.going	4.0
the.next.two	4.0
think.you.would	4.0
i.tried.to	4.0
better.than.we	4.0
were.pleased.with	4.0
to.make.it	4.0
to.talk.about	4.0
to.go.through	3.9
an.area.that	3.9
it.takes.a	3.9
my.guess.is	3.9
to.come.out	3.9
not.think.its	3.9
quarter.and.i	3.9
at.the.beginning	3.9
but.having.said	3.9
to.get.there	3.9
to.comment.on	3.9
we.havent.seen	3.9
the.bank.of	3.9
say.we.have	3.9
in.a.little	3.9
of.our.customers	3.9
in.i.think	3.9
want.to.add	3.8
the.way.they	3.8
as.we.work	3.8
youre.referring.to	3.8
to.take.some	3.8
where.we.want	3.8
you.know.what	3.8
so.i.wouldnt	3.8
think.that.was	3.8
will.take.a	3.8
on.the.call	3.8
of.these.things	3.8
got.to.be	3.7
to.the.fourth	3.7
the.second.part	3.7
feel.very.comfortable	3.7
way.i.would	3.7
first.quarter.but	3.7
want.to.give	3.7
do.not.give	3.7
continue.to.invest	3.7
year.and.then	3.7

Token	ϕ
back.half.of	3.7
say.that.i	3.7
things.that.i	3.7
the.back.half	3.7
to.work.through	3.7
think.what.were	3.7
think.that.its	3.7
later.in.the	3.7
i.know.you	3.7
i.gave.you	3.7
a.really.good	3.7
think.at.the	3.7
we.might.have	3.6
know.what.the	3.6
but.i.wouldnt	3.6
half.of.the	3.6
get.into.the	3.6
over.a.period	3.6
pleased.with.the	3.6
on.the.second	3.6
when.i.look	3.6
time.and.i	3.6
think.you.can	3.6
we.know.we	3.6
think.all.of	3.6
fair.to.say	3.6
into.the.second	3.6
last.year.that	3.6
just.add.to	3.6
to.happen.in	3.5
what.were.going	3.5
with.the.regulators	3.5
but.i.cannot	3.5
the.thing.that	3.5
second.part.of	3.5
we.have.spent	3.5
it.is.today	3.5
will.say.that	3.5
think.were.going	3.5
like.to.see	3.5
have.to.make	3.5
year.and.that	3.5
i.think.whats	3.5
we.can.get	3.5
our.customers.and	3.5
give.you.a	3.5
have.a.great	3.5
the.competitive.environment	3.5
still.going.to	3.5
into.the.fourth	3.5
what.we.said	3.5
have.got.to	3.5
think.there.are	3.5
are.seeing.the	3.5
a.little.better	3.5
its.one.of	3.5
but.you.know	3.4
and.the.first	3.4

Token	ϕ
the.right.way	3.4
we.have.worked	3.4
the.business.so	3.4
was.going.to	3.4
that.we.thought	3.4
do.not.break	3.4
for.the.year	3.4
it.looks.like	3.4
i.should.say	3.4
think.as.i	3.4
quarter.of.last	3.4
cannot.tell.you	3.4
it.will.take	3.4
it.is.hard	3.4
thats.something.that	3.4
no.i.think	3.3
want.to.say	3.3
this.year.and	3.3
used.to.be	3.3
not.think.theres	3.3
going.to.work	3.3
of.next.year	3.3
give.you.more	3.3
well.let.me	3.3
going.to.need	3.3
it.could.be	3.3
about.the.fact	3.3
the.third.and	3.3
i.will.say	3.3
as.soon.as	3.3
going.to.happen	3.3
would.just.say	3.3
trying.to.be	3.3
the.next.year	3.3
were.not.seeing	3.3
our.clients.are	3.2
quarter.that.we	3.2
be.a.lot	3.2
to.help.us	3.2
turn.it.over	3.2
i.would.have	3.2
thought.we.would	3.2
think.that.were	3.2
think.this.is	3.2
well.be.able	3.2
to.answer.your	3.2
an.opportunity.for	3.2
question.i.think	3.2
is.a.business	3.2
going.to.do	3.2
were.excited.about	3.2
the.first.and	3.2
fourth.quarter.and	3.2
we.can.see	3.2
whats.going.to	3.2
our.businesses	

Token	ϕ
.and	3.2
around.the.globe	3.2
there.might.be	3.2
think.that.would	3.2
think.you.should	3.2
right.way.to	3.2
and.kind.of	3.2
think.they.are	3.2
been.focused.on	3.2
were.seeing.some	3.2
and.were.going	3.2
what.we.thought	3.2
of.the.year	3.2
fourth.quarter.as	3.2
to.take.a	3.2
way.i.think	3.1
not.think.the	3.1
i.think.i	3.1
the.us.and	3.1
on.that.front	3.1
until.we.get	3.1
said.we.have	3.1
of.the.first	3.1
happen.in.the	3.1
i.have.said	3.1
not.think.were	3.1
the.other.things	3.1
on.the.last	3.1
just.looking.at	3.1
i.cannot.tell	3.1
the.commercial.side	3.1
that.it.would	3.1
i.wouldnt.expect	3.1
for.many.years	3.1
a.good.job	3.1
to.be.much	3.1
think.that.theres	3.1
of.the.changes	3.1
have.to.go	3.1
be.much.more	3.1
that.right.now	3.1
is.not.something	3.1
for.the.second	3.1
good.about.it	3.1
that.is.going	3.1
we.really.havent	3.1
things.that.are	3.0
need.to.get	3.0
to.come.up	3.0
i.think.thats	3.0
do.not.feel	3.0
other.things.t	

Token	ϕ
hat	3.0
know.if.you	3.0
a.great.question	3.0
going.to.get	3.0
we.get.to	3.0
i.guess.what	3.0
we.are.going	3.0
its.hard.to	3.0
that.its.going	3.0
an.awful.lot	3.0
the.i.think	3.0
you.a.little	3.0
continue.to.drive	3.0
have.made.a	3.0
is.a.lot	3.0
in.the.next	3.0
in.new.york	2.9
i.think.were	2.9
but.we.havent	2.9
talked.about.a	2.9
a.couple.years	2.9
of.years.ago	2.9
happy.with.the	2.9
third.and.fourth	2.9
have.been.pretty	2.9
quarter.i.would	2.9
lot.of.time	2.9
think.it.will	2.9
they.have.got	2.9
it.is.something	2.9
in.the.process	2.9
products.and.services	2.9
think.it.is	2.9
next.few.quarters	2.9
parts.of.our	2.9
guess.i.would	2.9
about.a.year	2.9
got.a.very	2.9
we.will.take	2.9
get.a.little	2.9
the.things.we	2.9
in.the.growth	2.9
going.to.start	2.9
thats.really.what	2.9
say.that.were	2.9
other.parts.of	2.9
think.thats.what	2.9
in.that.category	2.9
but.right.now	2.9
not.a.big	2.9
quarter.last.year	2.9
think.thats.the	2.9
this.quarter.i	2.8

Token	ϕ
i.think.they	2.8
but.i.think	2.8
number.that.we	2.8
expect.to.be	2.8
we.will.see	2.8
thinking.about.it	2.8
of.your.question	2.8
were.still.in	2.8
do.not.want	2.8
that.business.and	2.8
of.people.that	2.8
would.say.it	2.8
really.want.to	2.8
tell.you.is	2.8
and.fourth.quarter	2.8
i.think.it	2.8
that.could.be	2.8
think.i.would	2.8
actions.that.we	2.8
think.it.was	2.8
the.way.i	2.8
there.is.going	2.8
of.things.that	2.8
for.some.time	2.8
the.month.of	2.8
both.sides.of	2.8
right.now.that	2.8
would.add.to	2.8
each.one.of	2.8
to.do.something	2.7
first.and.second	2.7
spent.a.lot	2.7
us.i.think	2.7
think.you.know	2.7
the.next.few	2.7
the.united.states	2.7
been.a.little	2.7
be.happy.to	2.7
would.have.expected	2.7
couple.of.months	2.7
fourth.quarter.so	2.7
and.theres.a	2.7
the.consumer.side	2.7
continue.to.focus	2.7
might.have.been	2.7
still.in.the	2.7
other.part.of	2.7
not.want.to	2.7
this.quarter.it	2.7
well.we.are	2.7
would.say.in	2.7
it.i.think	2.7

Token	ϕ
way.we.think	2.7
its.not.going	2.7
of.the.second	2.7
its.going.to	2.7
part.of.your	2.7
with.a.lot	2.7
it.is.still	2.7
little.bit.from	2.7
for.the.next	2.7
able.to.do	2.7
spend.a.lot	2.7
i.think.this	2.6
very.pleased.w	
ith	2.6
a.whole.lot	2.6
said.i.think	2.6
bit.on.the	2.6
with.us.and	2.6
to.work.on	2.6
well.and.we	2.6
you.know.a	2.6
that.a.lot	2.6
think.youre.go	
ing	2.6
we.made.a	2.6
they.want.to	2.6
think.that.you	2.6
will.tell.you	2.6
a.sense.of	2.6
well.i.would	2.6
to.deal.with	2.6
it.that.way	2.6
not.see.it	2.6
on.the.consumer	2.6
and.the.second	2.6
yes.i.would	2.6
continue.to.ma	
ke	2.6
need.to.be	2.6
terms.of.how	2.6
quarter.it.was	2.6
of.our.growth	2.6
very.difficult	
.to	2.6
in.fact.i	2.6
that.i.mean	2.6
from.our.persp	
ective	2.6
one.thing.i	2.6
year.or.so	2.6
that.we.might	2.6
have.done.that	2.5
have.seen.it	2.5
deal.with.the	2.5
at.this.time	2.5
and.i.know	2.5
talk.a.little	2.5
said.we.are	2.5
what.we.need	2.5
the.last.two	2.5
year.or.two	2.5
of.this.year	2.5
us.so.we	2.5

Token	ϕ
i.think.there	2.5
in.the.numbers	2.5
in.the.second	2.5
and.so.well	2.5
just.going.to	2.5
think.its.fair	2.5
but.i.would	2.5
as.strong.as	2.5
think.that.is	2.5
into.the.first	2.5
for.us.but	2.5
beginning.of.t	
he	2.5
wanted.to.make	2.5
will.probably.	
be	2.5
the.economy.is	2.5
confident.that	
.we	2.5
to.sort.of	2.4
end.of.this	2.4
where.we.need	2.4
i.mean.i	2.4
i.think.from	2.4
the.business.in	2.4
i.guess.i	2.4
its.a.good	2.4
i.agree.with	2.4
think.that.if	2.4
part.of.what	2.4
again.i.think	2.4
that.we.made	2.4
think.its.a	2.4
this.year.that	2.4
were.thinking.	
about	2.4
the.last.years	2.4
try.to.do	2.4
the.pace.of	2.4
think.the.way	2.4
quarter.but.we	2.4
year.but.i	2.4
in.europe.and	2.4
theres.still.a	2.4
we.expect.it	2.4
fourth.quarter	
.of	2.4
bit.of.an	2.4
things.that.we	2.4
think.we.do	2.4
would.say.is	2.4
think.we.had	2.4
our.clients.and	2.4
i.want.to	2.4
the.latter.part	2.4
i.said.earlier	2.4
of.last.year	2.4
think.we.have	2.4
the.year.to	2.4
pick.up.in	2.4
something.that	
.we	2.3
so.i.cannot	2.3

Token	ϕ
i.will.tell	2.3
to.make.sure	2.3
a.year.or	2.3
are.not.seeing	2.3
i.would.not	2.3
really.focused	
.on	2.3
little.bit.abo	
ut	2.3
would.probably	
.be	2.3
going.to.have	2.3
business.and.t	
hat	2.3
what.we.expect	2.3
have.said.that	2.3
and.we.certain	
ly	2.3
first.quarter.	
and	2.3
us.to.do	2.3
say.that.we	2.3
the.right.thing	2.3
i.think.just	2.3
to.go.into	2.3
a.business.that	2.3
were.very.plea	
sed	2.3
is.an.area	2.3
third.quarter.	
so	2.3
a.little.more	2.3
a.question.of	2.3
you.know.if	2.3
right.now.but	2.3
think.we.were	2.3
were.trying.to	2.3
in.the.fourth	2.3
see.a.little	2.3
the.world.and	2.2
the.benefits.of	2.2
had.a.little	2.2
two.or.three	2.2
in.the.early	2.2
you.think.of	2.2
see.how.that	2.2
business.that.	
we	2.2
the.question.is	2.2
in.the.right	2.2
thats.why.i	2.2
not.something.	
that	2.2
to.see.it	2.2
lot.of.things	2.2
a.bit.in	2.2
the.fourth.qua	
rter	2.2
into.next.year	2.2
would.say.a	2.2
the.process.of	2.2
i.think.its	2.2
to.give.you	2.2

Token	ϕ
that.i.think	2.2
have.a.good	2.2
like.that.but	2.2
the.last.quart	
er	2.2
right.now.i	2.2
last.year.and	2.2
i.mean.we	2.2
to.see.us	2.2
on.the.corpora	
te	2.2
most.important	
.thing	2.2
in.the.economy	2.2
it.is.going	2.2
in.q.and	2.2
thats.a.good	2.2
not.going.to	2.2
quarter.i.think	2.2
make.sure.we	2.2
the.end.of	2.2
will.see.that	2.2
was.a.good	2.2
for.a.second	2.2
last.year.but	2.2
just.in.terms	2.2
that.is.probab	
ly	2.2
want.to.see	2.2
little.bit.bet	
ter	2.2
are.things.that	2.2
those.things.a	
re	2.2
last.two.years	2.2
and.well.see	2.2
i.just.think	2.1
a.good.question	2.1
going.to.take	2.1
is.a.great	2.1
would.love.to	2.1
thats.a.little	2.1
us.we.have	2.1
outside.of.the	2.1
i.think.again	2.1
going.to.keep	2.1
lot.of.people	2.1
as.you.said	2.1
sure.that.were	2.1
i.think.we	2.1
i.can.tell	2.1
want.to.make	2.1
we.will.look	2.1
year.ago.and	2.1
think.were.in	2.1
i.would.just	2.1
time.but.we	2.1
give.you.the	2.1
the.us.we	2.1
around.the.wor	
ld	2.1
last.year.so	2.1
like.that.we	2.1

Token	ϕ
mean.i.think	2.1
the.year.we	2.1
continue.to.do	2.1
were.a.little	2.1
to.drive.that	2.1
we.need.to	2.1
to.be.careful	2.1
so.i.guess	2.1
year.and.i	2.0
i.think.about	2.0
to.make.a	2.0
things.we.have	2.0
late.in.the	2.0
make.sure.that	2.0
focused.on.that	2.0
i.said.before	2.0
second.and.thi	
rd	2.0
this.year.so	2.0
the.economy.and	2.0
that.well.be	2.0
i.wouldnt.say	2.0
go.back.to	2.0
have.gone.thro	
ugh	2.0
we.thought.we	2.0
right.i.think	2.0
in.the.united	2.0
through.the.ye	
ar	2.0
think.we.can	2.0
right.now.so	2.0
of.our.busines	
ses	2.0
the.beginning.	
of	2.0
half.of.this	2.0
awful.lot.of	2.0
business.and.i	2.0
a.bit.more	2.0
would.like.to	2.0
impact.in.the	2.0
during.the.cou	
rse	2.0
the.way.you	2.0
year.in.the	2.0
trying.to.do	2.0
i.think.what	2.0
every.one.of	1.9
the.last.three	1.9
fourth.quarter	
.is	1.9
so.we.expect	1.9
the.high.end	1.9
add.to.that	1.9
latter.part.of	1.9
over.a.year	1.9
little.bit.low	
er	1.9
the.growth.rate	1.9
see.some.of	1.9
can.tell.you	1.9
of.the.fourth	1.9

Token	ϕ
think.it.would	1.9
have.a.number	1.9
just.want.to	1.9
sense.for.us	1.9
it.will.continue	1.9
quarter.so.that	1.9
us.and.i	1.9
and.we.expect	1.9
but.we.feel	1.9
types.of.things	1.9
when.we.get	1.9
talked.about.we	1.9
think.we.are	1.9
what.i.said	1.9
the.good.news	1.9
we.would.certainly	1.9
to.get.that	1.9
will.look.at	1.9
of.the.best	1.9
as.we.looked	1.9
thing.that.i	1.9
look.i.think	1.9
a.very.competitive	1.9
based.on.what	1.9
first.quarter.i	1.9
to.respond.to	1.9
businesses.that.we	1.8
year.so.we	1.8
yes.so.i	1.8
have.been.saying	1.8
its.kind.of	1.8
thats.what.were	1.8
think.about.it	1.8
it.makes.sense	1.8
i.think.for	1.8
next.couple.of	1.8
in.the.sense	1.8
to.invest.in	1.8
were.going.to	1.8
got.a.lot	1.8
have.been.investing	1.8
i.think.people	1.8
sure.that.we	1.8
business.for.us	1.8
were.focused.on	1.8
i.would.look	1.8
give.you.an	1.8
of.the.growth	1.8
i.do.think	1.8
the.third.quarter	1.8
and.i.think	1.8
going.to.give	1.8
of.our.business	1.8
at.a.time	1.8
last.year.in	1.8
number.of.things	1.8
we.go.into	1.8
i.think.that	1.8
going.to.come	1.8
to.think.that	1.7
and.i.would	1.7
we.saw.that	1.7
of.the.next	1.7
really.do.not	1.7
number.of.years	1.7
expect.it.to	1.7
i.just.want	1.7
probably.a.little	1.7
that.were.looking	1.7
well.i.do	1.7
want.to.do	1.7
so.well.have	1.7
would.say.that	1.7
for.us.and	1.7
our.customer.base	1.7
we.have.tried	1.7
say.is.that	1.7
i.would.say	1.7
little.bit.more	1.7
i.would.describe	1.7
kinds.of.things	1.7
to.make.that	1.7
the.business.we	1.7
are.a.lot	1.7
come.up.with	1.7
of.the.quarter	1.7
the.first.quarter	1.7
in.the.uk	1.7
have.tried.to	1.7
you.go.back	1.7
the.growth.that	1.7
fourth.quarter.but	1.7
course.of.the	1.7
second.quarter.and	1.6
the.last.year	1.6
area.where.we	1.6
that.continues.to	1.6
we.have.mentioned	1.6
do.not.get	1.6
year.i.think	1.6
initiatives.that.we	1.6
think.you.will	1.6
think.that.we	1.6
we.saw.some	1.6
think.there.is	1.6
us.and.we	1.6
to.our.customers	1.6
want.to.be	1.6
could.be.a	1.6
well.we.have	1.6
because.i.think	1.6
we.are.still	1.6
i.said.we	1.6
that.i.would	1.6
not.a.lot	1.6
the.prior.year	1.6
want.to.get	1.6
everything.that.we	1.6
we.didnt.have	1.5
the.quarter.but	1.5
to.get.into	1.5
yes.i.think	1.5
going.on.there	1.5
the.course.of	1.5
on.the.first	1.5
was.a.little	1.5
we.tried.to	1.5
time.we.have	1.5
going.to.continue	1.5
the.other.part	1.5
but.we.think	1.5
in.the.last	1.5
right.now.and	1.5
have.done.a	1.5
are.a.little	1.5
so.a.lot	1.5
a.little.less	1.5
think.we.will	1.5
have.been.talking	1.5
that.would.be	1.5
thing.i.would	1.5
not.think.that	1.5
i.think.youll	1.5
there.as.well	1.5
quarter.we.have	1.5
in.that.regard	1.5
i.do.believe	1.5
last.year.we	1.5
what.i.would	1.5
to.say.that	1.5
in.the.back	1.5
the.things.that	1.4
quarter.and.then	1.4
i.think.is	1.4
yes.i.mean	1.4
making.sure.that	1.4
it.tends.to	1.4
in.the.year	1.4
of.the.business	1.4
of.the.products	1.4
for.this.year	1.4
i.think.theres	1.4
trying.to.make	1.4
be.the.case	1.4
i.think.youre	1.4
and.the.growth	1.4
all.the.time	1.4
said.that.the	1.4
in.the.third	1.4
the.second.quarter	1.4
expect.us.to	1.4
and.we.want	1.3
we.have.got	1.3
i.think.with	1.3
fourth.quarter.i	1.3
look.at.this	1.3
business.in.the	1.3
things.i.think	1.3
we.feel.we	1.3
that.we.saw	1.3
let.me.just	1.3
we.saw.in	1.3
i.think.all	1.3
need.to.make	1.3
a.lot.to	1.3
know.we.do	1.3
comment.on.the	1.3
we.are.continuing	1.3
i.think.as	1.3
we.needed.to	1.3
have.i.think	1.3
consistent.with.what	1.3
the.part.of	1.3
the.next.couple	1.3
in.the.first	1.3
it.would.be	1.3
so.were.going	1.3
we.are.actually	1.3
the.one.thing	1.3
i.think.on	1.3
be.able.to	1.3
i.mean.its	1.3
been.talking.about	1.2
things.that.were	1.2
would.tell.you	1.2
terms.of.where	1.2
the.institutional.side	1.2
are.continuing.to	1.2
going.through.the	1.2
as.i.said	1.2
first.quarter.of	1.2
think.what.we	1.2
i.would.tell	1.2
would.think.about	1.2
coming.out.of	1.2
do.not.really	1.2
have.kind.of	1.2
that.is.really	1.2
had.a.lot	1.2
to.start.to	1.2
over.the.next	1.2
really.hard.to	1.2
not.really.have	1.1
to.see.the	1.1
side.i.think	1.1
is.a.little	1.1
end.of.the	1.1
to.come.back	1.1
what.we.saw	1.1
four.or.five	1.1
look.at.it	1.1
for.our.clients	1.1
going.to.be	1.1
the.ones.that	1.1
i.think.you	1.1
were.not.going	1.1
well.i.think	1.1
up.a.little	1.1
think.you.have	1.1
a.little.bit	1.1
we.have.done	1.1
to.think.about	1.1
way.to.think	1.1
of.the.things	1.0
think.as.we	1.0
we.have.already	1.0
the.future.but	1.0
things.like.that	1.0
and.you.know	1.0
is.hard.to	1.0
do.not.like	1.0
to.see.some	1.0
trying.to.get	1.0
over.the.course	1.0
i.would.think	1.0
if.you.want	1.0
we.were.going	1.0
have.a.lot	1.0
we.would.like	1.0
are.going.to	0.9
on.the.institutional	0.9
in.the.us	0.9
very.focused.on	0.9
very.hard.to	0.9
think.thats.a	0.9
we.feel.very	0.9
a.year.ago	0.9
but.let.me	0.9
quarter.so.i	0.9
that.were.going	0.9
go.out.and	0.9
we.have.never	0.8
for.us.to	0.8
at.the.end	0.8
have.got.a	0.8
or.may.not	0.8
a.position.to	0.7
we.really.do	0.7

Token	ϕ
of.those.things	0.7
think.if.you	0.7
couple.of.years	0.7
i.know.that	0.7
may.or.may	0.7
just.give.you	0.7
so.i.think	0.7
going.to.go	0.6
the.short.answer	0.6
we.talked.about	0.6
we.sit.here	0.6
so.we.havent	0.6
this.point.i	0.6
part.of.the	0.6
stuff.like.that	0.6
they.are.going	0.6
we.havent.really	0.6
rates.are.going	0.6
as.we.sit	0.6
of.factors.that	0.6
figure.out.what	0.6
sit.here.today	0.6
business.so.i	0.6
figure.out.how	0.5
you.a.sense	0.5
depends.on.what	0.5
what.that.means	0.5
that.level.of	0.5
that.we.think	0.5
on.an.annual	0.5
are.not.going	0.5
an.annual.basis	0.5
for.us.i	0.5
going.to.try	0.5
is.going.to	0.5
its.not.something	0.5
well.see.how	0.5
and.how.much	0.5
with.the.fed	0.5
we.can.provide	0.5
whether.or.not	0.5
how.much.of	0.5
long.as.we	0.5
have.come.down	0.5
you.can.get	0.5
for.the.future	0.5
in.interest.rates	0.5
a.few.things	0.5
can.look.at	0.5
the.net.interest	0.5
i.think.the	0.5
what.is.going	0.4
talk.about.the	0.4
in.response.to	0.4
have.given.you	0.4
end.up.with	0.4
like.that.so	0.4
that.you.could	0.4
in.a.position	0.4
in.the.pipeline	0.4
the.next.months	0.4
that.will.come	0.4
do.not.believe	0.4
theyre.going.to	0.4
over.time.so	0.4
a.quarterly.basis	0.4
we.put.out	0.4
its.fair.to	0.4
may.not.be	0.4
a.great.deal	0.4
money.market.funds	0.4
to.figure.out	0.4
step.back.and	0.4
in.a.good	0.4
to.put.a	0.4
but.we.would	0.4
thats.not.a	0.4
we.know.that	0.4
given.that.we	0.4
to.continue.to	0.4
give.you.some	0.4
of.the.industry	0.4
take.into.account	0.4
the.federal.reserve	0.4
i.would.guess	0.4
at.the.moment	0.4
thats.a.very	0.4
any.sort.of	0.4
at.the.last	0.4
something.in.the	0.4
tell.you.that	0.4
the.mortgage.business	0.4
a.level.of	0.4
some.kind.of	0.4
much.of.that	0.4
those.two.things	0.3
and.i.believe	0.3
seem.to.be	0.3
going.forward.but	0.3
on.a.quarterly	0.3
to.go.in	0.3
having.said.that	0.3
come.in.and	0.3
to.the.market	0.3
it.this.way	0.3
with.regard.to	0.3
net.interest.margin	0.3
at.that.level	0.3
that.people.are	0.3
the.extent.we	0.3
this.point.but	0.3
you.get.into	0.3
in.the.mortgage	0.3
that.theres.a	0.3
not.believe.that	0.3
the.timing.of	0.3
that.we.know	0.3
be.i.think	0.3
the.ccar.process	0.3
at.some.point	0.3
a.lot.of	0.3
in.the.investment	0.3
i.would.put	0.3
likely.to.be	0.3
at.this.point	0.3
you.can.look	0.3
so.thats.one	0.3
you.know.that	0.3
would.think.that	0.3
through.the.end	0.3
mentioned.that.we	0.3
out.i.think	0.3
and.i.guess	0.3
that.go.into	0.3
the.equity.side	0.3
in.our.numbers	0.3
for.a.number	0.3
going.to.look	0.3
year.but.we	0.3
one.is.the	0.3
would.be.very	0.3
you.know.it	0.3
point.in.time	0.3
on.the.revenue	0.3
to.see.in	0.3
think.about.this	0.3
in.this.case	0.3
this.point.in	0.2
ahead.of.the	0.2
also.have.a	0.2
top.of.that	0.2
lot.of.different	0.2
the.longer.term	0.2
in.that.range	0.2
i.think.over	0.2
get.to.a	0.2
but.in.terms	0.2
good.question.i	0.2
on.top.of	0.2
do.not.need	0.2
go.into.the	0.2
when.you.see	0.2
of.the.range	0.2
expectation.is.that	0.2
each.and.every	0.2
and.over.time	0.2
what.happens.with	0.2
the.same.kind	0.2
a.bunch.of	0.2
we.believe.the	0.2
one.of.those	0.2
the.percentage.of	0.2
i.mentioned.in	0.2
it.depends.on	0.2
thats.what.i	0.2
have.to.take	0.2
trying.to.figure	0.2
we.have.given	0.2
we.said.in	0.2
of.the.way	0.2
the.change.in	0.2
the.things.i	0.2
not.trying.to	0.2
back.into.the	0.2
see.that.we	0.2
like.that.and	0.2
our.view.is	0.2
outlook.for.the	0.2
in.my.remarks	0.2
saw.this.quarter	0.2
going.to.change	0.2
same.kind.of	0.2
to.come.down	0.2
is.not.going	0.2
we.would.do	0.2
have.looked.at	0.2
quarter.of.the	0.2
the.industry.and	0.2
of.that.business	0.2
the.size.of	0.2
mentioned.in.my	0.2
it.relates.to	0.2
would.have.to	0.2
that.we.still	0.2
this.point.we	0.2
any.kind.of	0.2
thing.that.we	0.2
i.think.at	0.2
the.rate.of	0.2
this.is.one	0.2
do.not.necessarily	0.2
we.tend.to	0.2
the.revenue.side	0.1
is.to.get	0.1
i.believe.that	0.1
seems.to.be	0.1
are.talking.about	0.1
years.we.have	0.1
an.environment.where	0.1
in.the.company	0.1
as.you.know	0.1
terms.of.what	0.1
in.the.future	0.1
and.so.forth	0.1
have.a.pretty	0.1
see.what.the	0.1
is.one.of	0.1
we.believe.that	0.1
that.kind.of	0.1
i.look.at	0.1
i.mentioned.that	0.1
to.add.to	0.1
in.my.prepared	0.1
thats.going.to	0.1
know.we.have	0.1
by.the.way	0.1
that.we.want	0.1
when.it.comes	0.1
change.in.the	0.1
parts.of.the	0.1
you.know.we	0.1
and.the.reason	0.1
i.guess.the	0.1
as.it.relates	0.1
talked.about.that	0.1
regard.to.the	0.1
other.thing.that	0.1
is.going.on	0.1
how.we.think	0.1
to.try.to	0.1
to.try.and	0.1
at.the.time	0.1
as.much.as	0.1
and.it.really	0.1

A.2 Correlation between sentiment measures

Table A1: Correlation between sentiment measures

	1.	**2.**	**3.**	**4.**	**5.**	**6.**
1. *NonAnswer*						
2. *NonAnswer*$^{\phi}$	0.918					
3. *Negativity*	-0.131	-0.071				
4. *Uncertainty*	0.040	0.071	0.200			
5. *Numbers*	-0.097	-0.057	0.200	0.106		
6. *Complexity*	-0.088	-0.086	0.133	-0.050	0.021	

Notes: This table presents Pearson correlations between the sentiment variables used in our analysis. *NonAnswer* (*NonAnswer*$^{\phi}$) is the ratio of trigrams in our non-answer glossary (weighted by loadings) to the total words. *Negativity* and *Uncertainty* are the ratio of negative and uncertain words to the total words. List of negative and uncertain words are from Loughran and McDonald (2011) word-lists. *Numbers* is the share of numbers to the sum of total words and numbers calculated as in Zhou (2018). *Complexity* is the complexity measure derived from the Loughran and McDonald (2019) complexity word list.

A.3 Uncertain language versus non-answers: examples

A.3.1 Uncertainty without non-answers

Q: “Before i hit on the second question, again the variable piece on that, is that a 4% to 5% number? or how should we think about that?”

A: “That’s probably within the range. maybe 3% to 5%, just depending on the quarter.”

Q: “Okay. And is that the plan still to sort of bring the quarterly capacity up to about 345 million with the expansion in hillview and cammis?”

A: “Yes, depending how the demand goes, yes, we could do that or more or less depending what we needed.”

Q: “First off, what percentage of business is through (inaudible) for oem (ph)?”

A: “Our u.s. business, maybe about 60% through distribution; our international, probably maybe half, distribution.”

A.3.2 Non-answers without uncertainty

Q: “Then in terms of your promotional spending in the marketplace, did it tweak up just slightly sequentially from the second quarter level?”

A: “I actually don’t know offhand. i don’t think so. i don’t know that, we can get back to you on that.”

Q: “What’s your sense on timing? how do you expect it to play out what are the negative – (multiple speakers).”

A: “Too early to tell. i think it’s too early to tell. i wouldn’t take a guess at it at this point.”

Q: “Regionally, do you see one area being more rational than the other on those, rick?”

A: “Yes, but i really don’t want to comment on this call about that.”

A.3.3 Mixture of non-answers and uncertainty

Q: "John, on the 9 billion units that you mentioned, how much of that is actually cans?"

A: "Good question. the majority of it. i don't have the numbers off the top of my head. i would say approximately 75%. but i'd be certainly happy to get back to you on that."

Q: "And i – if i – i think the logical inference from the way you're describing a little bit of a change in business model there is that we'll see, what? greater seasonality, if you will, in – or maybe greater volatility, a wider range of margins through the year in international that gives you the ability to do more when you're doing well, but you'll have some extra staffing costs in softer quarters?"

A: "Well, it could be. we'll have to wait and see, yes."

A.4 Cross-sectional Fama-MacBeth regression

Table A2: Fama-MacBeth Regressions

	$FF3-CAR_{0;1}$				$FF5-CAR_{0;1}$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$NonAnswer$	-0.110***	-0.100***			-0.098***	-0.092***		
	(-2.82)	(-3.92)			(-2.93)	(-3.80)		
$NonAnswer^{\phi}$			-0.051***	-0.075***			-0.046***	-0.070***
			(-4.29)	(-2.97)			(-4.41)	(-2.95)
cons	0.010**	0.010**	0.010***	0.012**	0.009**	0.009**	0.009***	0.011**
	(2.33)	(2.25)	(3.14)	(2.39)	(2.39)	(2.12)	(3.21)	(2.32)
$Negativity$		-0.446***		-0.442***		-0.407***		-0.403***
		(-5.79)		(-5.78)		(-5.78)		(-5.77)
$Uncertainty$		-0.035		-0.031		-0.048		-0.046
		(-0.42)		(-0.37)		(-0.61)		(-0.57)
$Numbers$		-0.092		-0.098		-0.081		-0.087
		(-1.31)		(-1.34)		(-1.21)		(-1.26)
$Complexity$		0.194		0.284**		0.164		0.249**
		(1.26)		(2.36)		(1.14)		(2.22)
Observations	21191	21191	21191	21191	21191	21191	21191	21191
R^2	0.007	0.096	0.008	0.097	0.006	0.094	0.007	0.095
FirmControls	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Fama-MacBeth cross-sectional regressions for Equation (5). The dependent variable is the abnormal returns over the Fama-French three (1993) and five (2015) factor model returns cumulated from the day of the earnings call to the day after it, $FF3-CAR_{0;1}$ ($FF5-CAR_{0;1}$). $NonAnswer$ ($NonAnswer^{\phi}$) is the ratio of trigrams in our non-answer glossary (weighted by loadings) to the total words. $Negativity$ and $Uncertainty$ are the ratio of negative and uncertain words to the total words. List of negative and uncertain words are from Loughran and McDonald (2011) word-lists. $Numbers$ is the share of numbers to the sum of total words and numbers calculated as in Zhou (2018). $Complexity$ is the complexity measure derived from the Loughran and McDonald (2019) complexity word list. Firm controls include $EarnSurp$, BTM, $ln(Assets)$ and Tobin's Q. t-statistics are given in parentheses. Standard errors are clustered in the firm and quarter level. ***, **, * indicate significance at the 1%, 5% and 10% levels.

A.5 Distribution of Non-Answers and Non-Answers over time

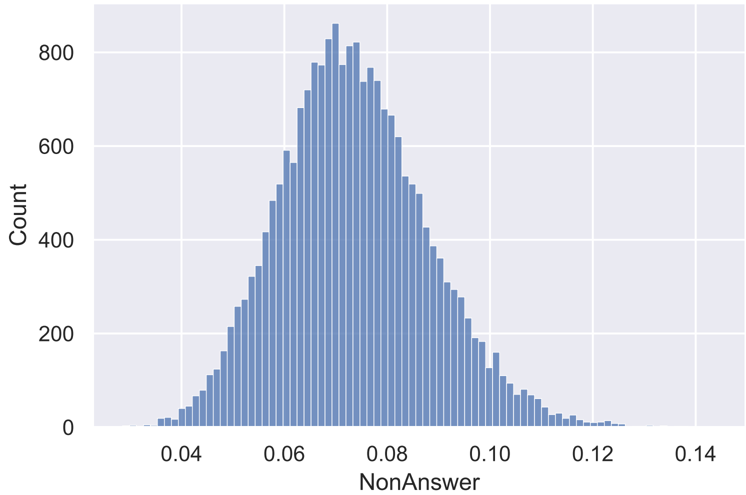


Figure A1: Distribution of *NonAnswer* for the validation set of S&P 500 earnings calls.

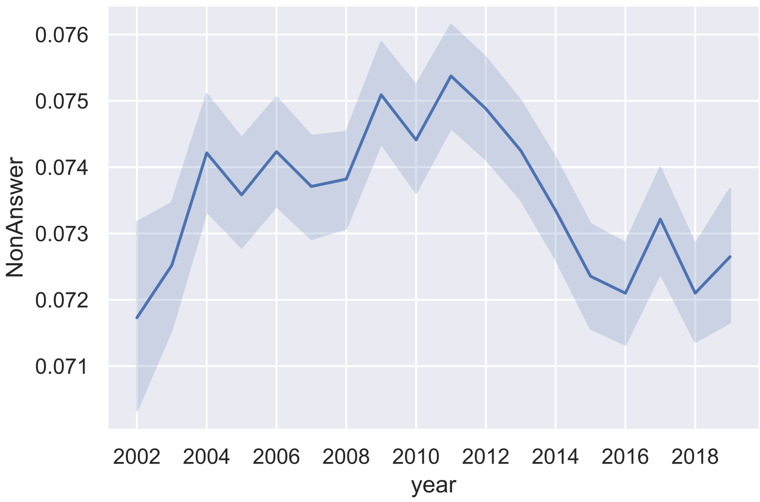


Figure A2: Average *NonAnswer* over time for the validation set of S&P 500 earnings calls. The shaded area reflects the 95% confidence band.

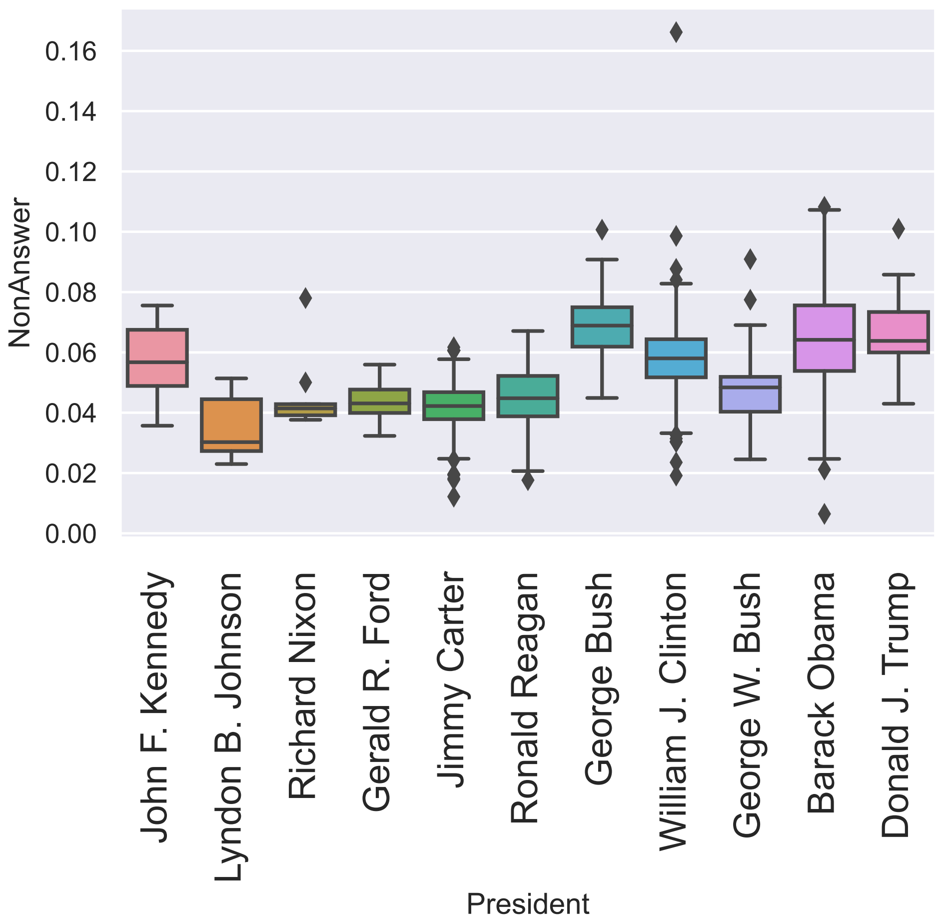
A.6 Alternative Q&A settings

A non-answer does not require a specific context. As both our method and glossary are free of financial context, we believe that the measure is applicable to other fields with a question and answers setup. In order to corroborate this claim, consider Mark Zuckerberg's responses to the US Senate during the Cambridge Analytica hearing as anecdotal evidence. His response "Senator, [...] I can certainly have my team get back to you on any specifics there that I don't know, sitting here today." is clearly a non-answer, and would have been identified as such by the glossary method.

In this Appendix, we briefly explore textual data of presidential interviews as another structured Q&A setting. Starting in 1864 with an interview with Abraham Lincoln, we analyze the answers of roughly 900 presidential interviews, which were collected by UCSB's American Presidency Project, see www.presidency.ucsb.edu.

The presidential data shows substantial variation in $NonAnswer$ (c.f. Figure A3), a necessary condition for the measure to be informative. Particularly high non-answer scores are found in interviews with President Clinton at the time when sexual assault allegations surfaced that later became the basis for an impeachment charge of perjury.

As an example for a high $NonAnswer$ presidential response, consider telephone interview with Morton Kondracke, in which Bill Clinton responded to the question: *"Okay. Let me just ask you one more question about this. You said in a statement today that you had no improper relationship with this intern. What exactly was the nature of your relationship with her?"* with the words *"Well, let me say, the relationship's not improper, and I think that's important enough to say. But because the investigation is going on and because I don't know what is out—what's going to be asked of me, I think I need to cooperate, answer the questions, but I think it's important for me to make it clear what is not. And then, at the appropriate time, I'll try to answer what is. But let me answer, it is not an improper relationship, and I know what the word means."*



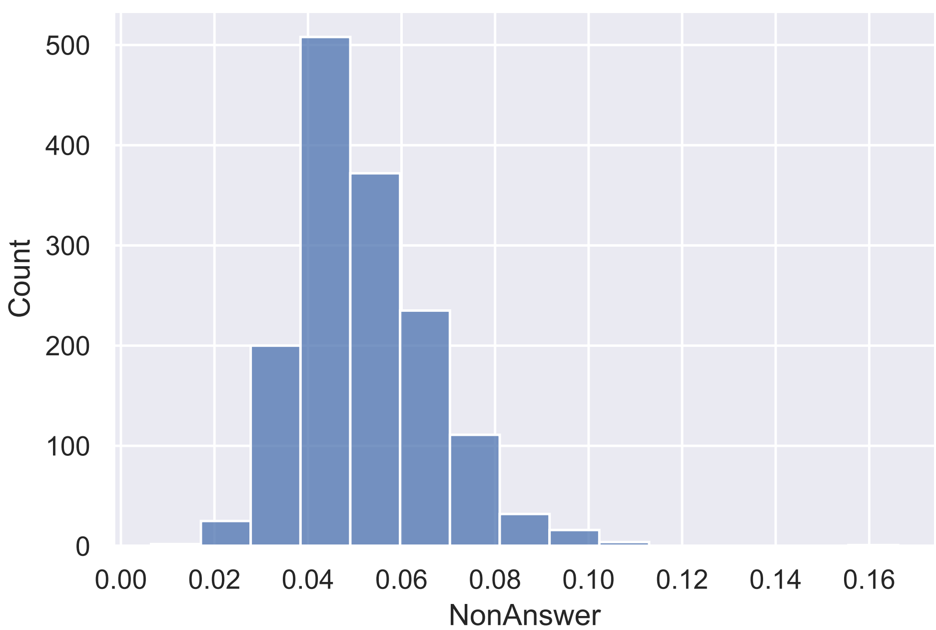


Figure A3: Presidents